Electron. Commun. Probab. **21** (2016), no. 63, 1–9.
DOI: 10.1214/16-ECP4644
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Convergence in density in finite time windows and the Skorohod representation

Hermann Thorisson*

Abstract

According to the Dudley-Wichura extension of the Skorohod representation theorem, convergence in distribution to a limit in a separable set is equivalent to the existence of a coupling with elements converging a.s. in the metric. A density analogue of this theorem says that a sequence of probability densities on a general measurable space has a probability density as a pointwise lower limit if and only if there exists a coupling with elements converging a.s. in the discrete metric. In this paper the discrete-metric theorem is extended to stochastic processes considered in a widening time window. The extension is then used to prove the separability version of the Skorohod representation theorem. The paper concludes with an application to Markov chains.

Keywords: Skorohod representation; convergence in distribution; convergence in density; widening time window.
AMS MSC 2010: Primary 60B10, Secondary 60G99.
Submitted to ECP on October 19, 2015, final version accepted on August 15, 2016.

1 Introduction

Let $X_1, X_2, \dots, X$ be random elements in a general space $(E, \mathcal{E})$ with distributions $P_1, P_2, \dots, P$. Let $f_1, f_2, \dots, f$ be the densities of $P_1, P_2, \dots, P$ with respect to some measure λ on $(E, \mathcal{E})$. Note that such a measure λ always exists, we could for instance take $\lambda = P + \sum_{n=1}^{\infty} 2^{-n} P_n$. If

$$\liminf_{n\to\infty} f_n = f \quad \text{a.e. } \lambda$$

we write

$$X_n \to X \text{ in density as } n \to \infty.$$

Note that f_n/f is defined almost everywhere P. It is the Radon-Nikodym derivative dP_n/dP of the absolutely continuous part of P_n with respect to P. Thus convergence in density does not depend on λ and is equivalent to

$$\liminf_{n\to\infty} dP_n/dP = 1 \quad \text{a.e. } P.$$

*Department of Mathematics, University of Iceland, Dunhaga 5, 107 Reykjavik, Iceland. E-mail: hermann@hi.is

In general, $\liminf_{n\to\infty} f_n = f$ a.e. λ is weaker than $\lim_{n\to\infty} f_n = f$ a.e. λ and stronger than convergence in total variation. However, if $(E, \mathcal{E})$ is discrete (that is, if E is countable and $\mathcal{E} = 2^E =$ the class of all subsets of E) then these three modes of convergence are equivalent and simplify to

$$\lim_{n\to\infty} \mathbf{P}(X_n = x) = \mathbf{P}(X = x), \quad x \in E;$$

see Theorems 6.1 and 7.1 in Chapter 1 of [12].

Let $(\hat{X}_1, \hat{X}_2, \dots, \hat{X})$ denote a coupling of $X_1, X_2, \dots, X$; this means that the random elements $\hat{X}_1, \hat{X}_2, \dots, \hat{X}$ are defined on a common probability space and have the marginal distributions $P_1, P_2, \dots, P$. In a 1995 paper [11], Section 5.4, this author showed that convergence in density is equivalent to the existence of a coupling converging in the discrete metric:

Theorem 0. *It holds that*

$$X_n \to X \text{ in density as } n \to \infty$$

if and only if there exists a coupling $(\hat{X}_1, \hat{X}_2, \dots, \hat{X})$ of $X_1, X_2, \dots, X$ such that for some random variable N taking values in $\mathbb{N} = \{1, 2, \dots\}$,

$$\hat{X}_n = \hat{X}, \quad n \geqslant N. \tag{1.1}$$

This density result is analogous to the Skorohod representation theorem which says that convergence in distribution on a complete separable metric E with $\mathcal{E}$ the Borel sets (a Polish space) is equivalent to the existence of a coupling converging a.s. in the metric. Skorohod proved this theorem in the 1956 paper [10], Dudley removed the completeness assumption in the 1968 paper [7], and Wichura showed in the 1970 paper [13] that it is enough that the limit probability measure P is concentrated on a separable Borel set; for historical notes, see [8]. Theorem 0 was rediscovered by Sethuraman [9] in 2002. For recent developments going beyond separability and considering convergence in probability, see the series of papers [2]–[6] by Berti, Pratelli and Rigo.

In the present paper we extend Theorem 0 to stochastic processes considered in a widening time window. The main result, Theorem 2.1, is established in Section 2 while Section 3 contains corollaries elaborating on that result. In Section 4, we show how this yields a new proof of the separability version of the Skorohod representation theorem. Section 5 concludes with an application to Markov chains.

2 Convergence in a widening time window

In this section we consider continuous-time stochastic processes without restriction on state space or paths. Also we allow the state space to vary with time and include infinity in the time set. Discrete-time processes are considered at the end of the section.

Let $(E^t, \mathcal{E}^t)$, $t \in [0, \infty]$, be a family of measurable spaces. Let H be a non-empty subset of the product set $\{(z^s)_{s\in[0,\infty]} : z^s \in E^s, s \in [0, \infty]\}$ and let $\mathcal{H}$ be the smallest σ-algebra on H making the maps taking $(z^s)_{s\in[0,\infty]} \in H$ to $z^t \in E^t$ measurable for all $t \in [0, \infty]$. For $t \in [0, \infty)$, let $(H^t, \mathcal{H}^t)$ be the image space of $(H, \mathcal{H})$ under the map taking $(z_s)_{s\in[0,\infty]} \in H$ to $(z_s)_{s\in[0,t)}$.

If $\mathbf{Z} = (Z^s)_{s\in[0,\infty]}$ is a random element in $(H, \mathcal{H})$ write $\mathbf{Z}^t = (Z^s)_{s\in[0,t)}$ for a segment of $\mathbf{Z}$ in a finite time window of length $t \in [0, \infty)$. Note that $\mathbf{Z}^t$ is a random element in $(H^t, \mathcal{H}^t)$. We also write $\mathbf{Z}^t$ for a random element in $(H^t, \mathcal{H}^t)$ even if no $\mathbf{Z}$ is present.

According to the following theorem, convergence in density in all finite time windows is the distributional form of discrete-metric convergence in a widening time window.

(Note that the coupling in this theorem is not a full coupling of $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}$ but only a coupling of $\mathbf{Z}_1^{t_1}, \mathbf{Z}_2^{t_2}, \dots, \mathbf{Z}$. Extensions to a full coupling are considered in the next section.)

Theorem 2.1. *Let $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}$ be random elements in $(H, \mathcal{H})$ where $(H, \mathcal{H})$ is as above. Then*

$$\forall t \in [0, \infty): \quad \mathbf{Z}_n^t \to \mathbf{Z}^t \text{ in density as } n \to \infty \tag{2.1}$$

if and only if there exists a sequence of numbers $0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \to \infty$ and a coupling $(\hat{\mathbf{Z}}_1^{t_1}, \hat{\mathbf{Z}}_2^{t_2}, \dots, \hat{\mathbf{Z}})$ of $\mathbf{Z}_1^{t_1}, \mathbf{Z}_2^{t_2}, \dots, \mathbf{Z}$ such that for some $\mathbb{N}$-valued random variable N,

$$\hat{\mathbf{Z}}_n^{t_n} = \hat{\mathbf{Z}}^{t_n}, \quad n \geqslant N. \tag{2.2}$$

Proof. First, assume existence of the coupling. Fix $t \in [0, \infty)$, take $m \in \mathbb{N}$ such that $t_m \geqslant t$, and note that then (2.2) yields $\hat{\mathbf{Z}}_n^t = \hat{\mathbf{Z}}^t$ for $n \geqslant \max\{N, m\}$. Use this and the fact that (1.1) implies convergence in density to obtain (2.1).

Conversely assume that (2.1) holds. With $t \in [0, \infty)$ and $n \in \mathbb{N}$, let Q be the distribution of $\mathbf{Z}$, let Q^t be the distribution of $\mathbf{Z}^t$, let Q_n^t be the distribution of $\mathbf{Z}_n^t$, let f_n^t be the density of $\mathbf{Z}_n^t$ with respect to some measure λ^t on $(H^t, \mathcal{H}^t)$, and let ν_n^t be the measure on $(H^t, \mathcal{H}^t)$ with density $g_n^t := \inf_{i \geqslant n} f_i^t$. Due to the assumption (2.1), g_n^t increases to a density of $\mathbf{Z}^t$ as $n \to \infty$. Thus by monotone convergence, the measures ν_n^t increase setwise to Q^t,

$$\nu_1^t \leqslant \nu_2^t \leqslant \cdots \nearrow Q^t, \quad t \in [0, \infty).$$

Thus there are numbers $1 = n_0 < n_1 < n_2 < \dots$ such that

$$0 \leqslant Q^k - \nu_{n_k}^k \leqslant 2^{-k}, \quad k \in \mathbb{N} \cup \{0\}.$$

For $A \in \mathcal{H}$ and $\mathbf{z}^k \in H^k$, let $q_k(A \,|\, \mathbf{z}^k)$ be the conditional probability of the event $\{\mathbf{Z} \in A\}$ given $\mathbf{Z}^k = \mathbf{z}^k$. Then

$$Q(A) = \int q_k(A \,|\, \cdot)\, dQ^k, \quad A \in \mathcal{H}.$$

Since $\nu_{n_k}^k \leqslant Q^k$ the measure $\nu_{n_k}^k$ is absolutely continuous with respect to Q^k. Thus we can extend $\nu_{n_k}^k$ from $(H^k, \mathcal{H}^k)$ to a measure ν_k on $(H, \mathcal{H})$ by

$$\nu_k(A) := \int q_k(A \,|\, \cdot)\, d\nu_{n_k}^k, \quad A \in \mathcal{H}.$$

The last three displays yield

$$0 \leqslant Q - \nu_k \leqslant 2^{-k}, \quad k \in \mathbb{N} \cup \{0\}.$$

Let h_k be a density of ν_k with respect to Q. For integers $k < m$ let $\nu_{k,m}$ be the measure with density $\min_{k \leqslant j \leqslant m} h_j$ with respect to Q. Partition H into sets $A_k, \dots, A_m \in \mathcal{H}$ such that $\min_{k \leqslant j \leqslant m} h_j = h_i$ on A_i and thus

$$\nu_{k,m}(\cdot \cap A_i) = \nu_i(\cdot \cap A_i), \quad k \leqslant i \leqslant m.$$

Now define $t_n = k$ if $n_k \leqslant n < n_{k+1}$. The last two displays yield

$$0 \leqslant Q - \nu_{t_n,m} = \sum_{i=t_n}^{m} \Big(Q(\cdot \cap A_i) - \nu_i(\cdot \cap A_i) \Big) \leqslant \sum_{i=t_n}^{\infty} 2^{-i} = 2^{-t_n+1}.$$

Let μ_n be the measure with density $\inf_{t_n \leqslant i < \infty} h_i$ with respect to Q and send $m \to \infty$ to obtain $0 \leqslant Q - \mu_n \leqslant 2^{-t_n+1}$. Thus the μ_n increase setwise to Q,

$$0 =: \mu_0 \leqslant \mu_1 \leqslant \mu_2 \leqslant \cdots \nearrow Q. \tag{2.3}$$

Let μ_n^k be the marginal of μ_n on $(H^k, \mathcal{H}^k)$. Note that $\nu_{n_k}^k$ is the marginal of ν_k on $(H^k, \mathcal{H}^k)$ and that $\mu_n \leqslant \nu_{t_n}$ and $\nu_{n_{t_n}}^{t_n} \leqslant \nu_n^{t_n}$ (since $n_{t_n} \leqslant n$). Thus $\mu_n^{t_n} \leqslant \nu_n^{t_n}$. Now ν_n^t has density $\inf_{i \geqslant n} f_i^t$ and Q_n^t has density f_n^t and thus $\nu_n^t \leqslant Q_n^t$. Since $\mu_n^{t_n} \leqslant \nu_n^{t_n}$ this yields

$$\mu_n^{t_n} \leqslant Q_n^{t_n}, \quad n \in \mathbb{N}. \tag{2.4}$$

Keep in mind (2.3) and (2.4) throughout the following coupling construction.

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space supporting the following collection of independent random elements with distributions to be specified below:

$$N,\ \mathbf{V}_1,\ \mathbf{V}_2, \ldots,\ \mathbf{W}_1,\ \mathbf{W}_2, \ldots$$

Let N be $\mathbb{N}$-valued with distribution function (see (2.3))

$$\mathbf{P}(N \leqslant n) = \mu_n(H), \quad n \in \mathbb{N}.$$

Let $\mathbf{V}_n$ be a random element in $(H, \mathcal{H})$ with distribution (see (2.3))

$$\frac{\mu_n - \mu_{n-1}}{\mathbf{P}(N = n)} \qquad \text{(arbitrary distribution if } \mathbf{P}(N = n) = 0\text{).}$$

Let $\mathbf{W}_n$ be a random element in $(H^{t_n}, \mathcal{H}^{t_n})$ with distribution (see (2.4))

$$\frac{Q_n^{t_n} - \mu_n^{t_n}}{\mathbf{P}(N > n)} \qquad \text{(arbitrary distribution if } \mathbf{P}(N > n) = 0\text{).}$$

Put $\hat{\mathbf{Z}} = \mathbf{V}_N$ to obtain that $\hat{\mathbf{Z}}$ has the same distribution as $\mathbf{Z}$,

$$\mathbf{P}(\hat{\mathbf{Z}} \in \cdot) = \sum_{n=1}^{\infty} \mathbf{P}(\mathbf{V}_n \in \cdot)\mathbf{P}(N = n) = \sum_{n=1}^{\infty} (\mu_n - \mu_{n-1}) = Q.$$

Put $\hat{\mathbf{Z}}_n^{t_n} = \mathbf{V}_N^{t_n}$ on $\{N \leqslant n\}$ and $\hat{\mathbf{Z}}_n^{t_n} = \mathbf{W}_n$ on $\{N > n\}$ to obtain that $\hat{\mathbf{Z}}_n^{t_n}$ has the same distribution as $\mathbf{Z}_n^{t_n}$,

$$\begin{aligned}
\mathbf{P}(\hat{\mathbf{Z}}_n^{t_n} \in \cdot) &= \sum_{k=1}^{n} \mathbf{P}(\mathbf{V}_k^{t_n} \in \cdot)\mathbf{P}(N = k) + \mathbf{P}(\mathbf{W}_n \in \cdot)\mathbf{P}(N > n) \\
&= \sum_{k=1}^{n} (\mu_k^{t_n} - \mu_{k-1}^{t_n}) + (Q_n^{t_n} - \mu_n^{t_n}) = Q_n^{t_n}.
\end{aligned}$$

By definition $\hat{\mathbf{Z}} = \mathbf{V}_N$ and thus $\hat{\mathbf{Z}}^{t_n} = \mathbf{V}_N^{t_n}$. Also by definition, $\hat{\mathbf{Z}}_n^{t_n} = \mathbf{V}_N^{t_n}$ on $\{N \leqslant n\}$. Thus $\hat{\mathbf{Z}}_n^{t_n} = \hat{\mathbf{Z}}^{t_n}$ when $n \geqslant N$, that is, (2.2) holds. □

If $\mathbf{Z} = (Z^1, Z^2, \ldots, Z^\infty)$ write $\mathbf{Z}^k = (Z^1, Z^2, \ldots, Z^k)$ for a segment in a finite time window of length $k \in \mathbb{N} \cup \{0\}$. The following is a discrete-time version of Theorem 2.1.

Corollary 2.2. *Let $\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}$ be random elements in some product space $(E^1, \mathcal{E}^1) \otimes (E^2, \mathcal{E}^2) \otimes \cdots \otimes (E^\infty, \mathcal{E}^\infty)$. Then*

$$\forall k \in \mathbb{N}: \quad \mathbf{Z}_n^k \to \mathbf{Z}^k \text{ in density as } n \to \infty$$

if and only if there exists a sequence of integers $0 \leqslant k_1 \leqslant k_2 \leqslant \cdots \to \infty$ and a coupling $(\hat{\mathbf{Z}}_1^{k_1}, \hat{\mathbf{Z}}_2^{k_2}, \ldots, \hat{\mathbf{Z}})$ of $\mathbf{Z}_1^{k_1}, \mathbf{Z}_2^{k_2}, \ldots, \mathbf{Z}$ such that for some $\mathbb{N}$-valued random variable N,

$$\hat{\mathbf{Z}}_n^{k_n} = \hat{\mathbf{Z}}^{k_n}, \quad n \geqslant N.$$

Proof. Apply Theorem 2.1 to $(Z_1^{\lfloor 1+s \rfloor})_{s \in [0,\infty]}, (Z_2^{\lfloor 1+s \rfloor})_{s \in [0,\infty]}, \ldots, (Z^{\lfloor 1+s \rfloor})_{s \in [0,\infty]}$. (Or repeat the proof of Theorem 2.1 with t and t_n replaced by k and k_n.) □

3 Extensions to a full coupling

The coupling in Theorem 2.1 is not a full coupling of $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}$ but only a coupling of $\mathbf{Z}_1^{t_1}, \mathbf{Z}_2^{t_2}, \dots, \mathbf{Z}$. However, in the discrete-time case of Corollary 2.2, if we restrict all but the infinite-time state space to be discrete, then there is the following simple extension of the coupling. It will be used in Section 4 to establish the separability version of the Skorohod representation theorem.

Corollary 3.1. *Let $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}$ be random elements in the product space $(E^1, \mathcal{E}^1) \otimes (E^2, \mathcal{E}^2) \otimes \cdots \otimes (E^\infty, \mathcal{E}^\infty)$ where $(E^1, \mathcal{E}^1), (E^2, \mathcal{E}^2), \dots$ are discrete and $(E^\infty, \mathcal{E}^\infty)$ is some measurable space. Then*

$$\forall k \in \mathbb{N}: \quad \mathbf{Z}_n^k \to \mathbf{Z}^k \text{ in density as } n \to \infty$$

if and only if there exists a coupling $(\hat{\mathbf{Z}}_1, \hat{\mathbf{Z}}_2, \dots, \hat{\mathbf{Z}})$ of $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}$ such that, for some $\mathbb{N}$-valued random variable N and integers $0 \leqslant k_1 \leqslant k_2 \leqslant \cdots \to \infty$,

$$\hat{\mathbf{Z}}_n^{k_n} = \hat{\mathbf{Z}}^{k_n}, \quad n \geqslant N.$$

Proof. Due to Corollary 2.2, we only need to show that $(\hat{\mathbf{Z}}_1^{k_1}, \hat{\mathbf{Z}}_2^{k_2}, \dots, \hat{\mathbf{Z}})$ can be extended to a coupling of $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}$. For that purpose set, for $n \in \mathbb{N}$ and $\mathbf{i}^{k_n} \in E^1 \times E^2 \times \cdots \times E^{k_n}$,

$$Q_{n,\mathbf{i}^{k_n}} = \text{the conditional distribution of } \mathbf{Z}_n \text{ given } \{\mathbf{Z}_n^{k_n} = \mathbf{i}^{k_n}\}. \tag{3.1}$$

Let the probability space $(\Omega, \mathcal{F}, \mathbf{P})$ supporting $\hat{\mathbf{Z}}_1^{k_1}, \hat{\mathbf{Z}}_2^{k_2}, \dots, \hat{\mathbf{Z}}$, N be large enough to also support random elements in $(E^1, \mathcal{E}^1) \otimes (E^2, \mathcal{E}^2) \otimes \cdots \otimes (E^\infty, \mathcal{E}^\infty)$,

$$\mathbf{V}_{n,\mathbf{i}^{k_n}}, \quad n \in \mathbb{N}, \ \mathbf{i}^{k_n} \in E^1 \times E^2 \times \cdots \times E^{k_n},$$

that are independent, independent of $(\hat{\mathbf{Z}}_1^{k_1}, \hat{\mathbf{Z}}_2^{k_2}, \dots, \hat{\mathbf{Z}}, N)$, and such that

$$\mathbf{V}_{n,\mathbf{i}^{k_n}}^{k_n} = \mathbf{i}^{k_n} \text{ and } \mathbf{V}_{n,\mathbf{i}^{k_n}} \text{ has distribution } Q_{n,\mathbf{i}^{k_n}}.$$

Note that $\mathbf{V}^{k_n}_{n,\hat{\mathbf{Z}}_n^{k_n}} = \hat{\mathbf{Z}}_n^{k_n}$. Thus we can extend $\hat{\mathbf{Z}}_n^{k_n}$ to a $\hat{\mathbf{Z}}_n$ by setting $\hat{\mathbf{Z}}_n := \mathbf{V}_{n,\hat{\mathbf{Z}}_n^{k_n}}$. Then

$$\mathbf{P}(\hat{\mathbf{Z}}_n \in \cdot) = \sum_{\mathbf{i}^{k_n}} \mathbf{P}(\mathbf{V}_{n,\mathbf{i}^{k_n}} \in \cdot) \mathbf{P}(\hat{\mathbf{Z}}_n^{k_n} = \mathbf{i}^{k_n}) = \sum_{\mathbf{i}^{k_n}} Q_{n,\mathbf{i}^{k_n}}(\cdot) \mathbf{P}(\hat{\mathbf{Z}}_n^{k_n} = \mathbf{i}^{k_n}).$$

Since $\hat{\mathbf{Z}}_n^{k_n}$ has the same distribution as $\mathbf{Z}_n^{k_n}$ we obtain from this and (3.1) that $\hat{\mathbf{Z}}_n$ has the same distribution as $\mathbf{Z}_n$, as desired. □

In Corollary 3.1 we obtained a full coupling of $\mathbf{Z}_1$, $\mathbf{Z}_2, \dots$, $\mathbf{Z}$ in the discrete-time case by restricting Z_n^k and Z^k to a discrete state space for $k \in \mathbb{N}$ but without restricting the state space of Z_n^∞ and Z^∞. We shall now much weaken this restriction at the expense of putting a restriction on Z_n^∞ and Z^∞.

Corollary 3.2. *Let $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}$ be random elements in the product of Polish spaces $(E^1, \mathcal{E}^1) \otimes (E^2, \mathcal{E}^2) \otimes \cdots \otimes (E^\infty, \mathcal{E}^\infty)$. Then the coupling $(\hat{\mathbf{Z}}_1^{k_1}, \hat{\mathbf{Z}}_2^{k_2}, \dots, \hat{\mathbf{Z}})$ in* Corollary 2.2 *can be extended to a coupling $(\hat{\mathbf{Z}}_1, \hat{\mathbf{Z}}_2, \dots, \hat{\mathbf{Z}})$ of $\mathbf{Z}_1$, $\mathbf{Z}_2, \dots$, $\mathbf{Z}$.*

Proof. Set $(G, \mathcal{G}) = (E^1, \mathcal{E}^1) \otimes (E^2, \mathcal{E}^2) \otimes \cdots \otimes (E^\infty, \mathcal{E}^\infty)$. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be the probability space supporting the random elements $\hat{\mathbf{Z}}_1^{k_1}$, $\hat{\mathbf{Z}}_2^{k_2}, \dots$, $\hat{\mathbf{Z}}$, N in Corollary 2.2. Since a countable product of Polish spaces is Polish, there exist probability kernels $Q_n(\cdot \mid \cdot)$, $n \in \mathbb{N}$, such that $Q_n(A \mid \mathbf{z}^{k_n})$ is the conditional probability of $\{\mathbf{Z}_n \in A\}$ given $\mathbf{Z}_n^{k_n} = \mathbf{z}^{k_n}$,

$A \in \mathcal{G}$ and $\mathbf{z}^{k_n} \in E^1 \times E^2 \times \cdots \times E^{k_n}$. According to the Ionescu-Tulcea extension theorem (see [1], Section 2.7.2), the set function defined, with $A \in \mathcal{F}, A^1, A^2 \ldots \in \mathcal{G}$ and $n \in \mathbb{N}$, by

$$\begin{aligned}&\tilde{\mathbf{P}}(A \times A^1 \times \cdots \times A^n \times E^{n+1} \times \cdots \times E^\infty)\\&\quad = \int_A \mathbf{P}(d\omega) \int_{A^1} Q_1\big(d\mathbf{z}_1 \,|\, \mathbf{Z}_1^{k_1}(\omega)\big) \ldots \int_{A^n} Q_n\big(d\mathbf{z}_n \,|\, \mathbf{Z}_n^{k_n}(\omega)\big)\end{aligned}$$

extends to a probability measure $\tilde{\mathbf{P}}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ where $\tilde{\Omega} = \Omega \times G \times G \times \ldots$ and $\tilde{\mathcal{F}} = \sigma(\mathcal{F} \times \mathcal{G} \times \mathcal{G} \times \ldots)$. Note that $Q_n\big(\{\mathbf{z}_n : \hat{\mathbf{Z}}_n^{k_n} \neq \mathbf{z}_n^{k_n}\} \mid \hat{\mathbf{Z}}_n^{k_n}\big) = 0$ a.s. $\mathbf{P}$ for all $n \in \mathbb{N}$ which implies that $\tilde{\mathbf{P}}\big(\bigcup_{n=1}^\infty \{(\omega, \mathbf{z}_1, \mathbf{z_2}, \ldots) : \hat{\mathbf{Z}}_n^{k_n}(\omega) \neq \hat{\mathbf{z}}_n^{k_n}\}\big) = 0$. Delete this $\tilde{\mathbf{P}}$ null set from $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbf{P}})$ to obtain a probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbf{P}})$ such that if $(\omega, \mathbf{z}_1, \mathbf{z_2}, \ldots) \in \hat{\Omega}$ then $\mathbf{z}_n$ is restricted to satisfy $\mathbf{z}_n^{k_n} = \hat{\mathbf{Z}}_n^{k_n}(\omega)$. Now extend $\hat{\mathbf{Z}}_n^{k_n}$ to a $\hat{\mathbf{Z}}_n$ as follows: for $(\omega, \mathbf{z}_1, \mathbf{z_2}, \ldots) \in \hat{\Omega}$ and $n \in \mathbb{N}$ put $\hat{\mathbf{Z}}_n(\omega, \mathbf{z}_1, \mathbf{z_2}, \ldots) := \mathbf{z}_n$. Due to $\mathbf{z}_n^{k_n} = \hat{\mathbf{Z}}_n^{k_n}(\omega)$, this definition transfers $\hat{\mathbf{Z}}_n^{k_n}$ consistently from $(\Omega, \mathcal{F}, \mathbf{P})$ to $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbf{P}})$. Finally transfer $\hat{\mathbf{Z}}$ to $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbf{P}})$ by putting $\hat{\mathbf{Z}}(\omega, \mathbf{z}_1, \mathbf{z_2}, \ldots) := \hat{\mathbf{Z}}(\omega)$ for $(\omega, \mathbf{z}_1, \mathbf{z_2}, \ldots) \in \hat{\Omega}$. □

The final corollary extends Corollary 3.2 to continuous time.

Corollary 3.3. *Let $\mathbf{Z}_1$, $\mathbf{Z}_2, \ldots$, $\mathbf{Z}$ be random elements in $(D, \mathcal{D}) \otimes (E, \mathcal{E})$ where $(D, \mathcal{D}) = (D[0,\infty), \mathcal{D}[0,\infty))$ is the Skorohod space of a Polish space and $(E, \mathcal{E})$ is Polish. Then the coupling $(\hat{\mathbf{Z}}_1^{t_1}, \hat{\mathbf{Z}}_2^{t_2}, \ldots, \hat{\mathbf{Z}})$ in* Theorem 2.1 *can be extended to a coupling $(\hat{\mathbf{Z}}_1, \hat{\mathbf{Z}}_2, \ldots, \hat{\mathbf{Z}})$ of $\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}$.*

Proof. The Skorohod space $(D, \mathcal{D})$ is Polish and thus the product $(H, \mathcal{H}) = (D, \mathcal{D}) \otimes (E, \mathcal{E})$ is Polish. Proceed as in the proof of Corollary 3.2 referring to Theorem 2.1 rather than Corollary 2.2, replacing $(G, \mathcal{G})$ by $(H, \mathcal{H})$ and k_n by t_n, and with $A \in \mathcal{H}$ and $\mathbf{z}^{t_n} \in H^{t_n}$. □

4 The Skorohod representation

In this section let E be a metric space with metric d and $\mathcal{E}$ its Borel subsets. Recall that X_n is said to converge to X in distribution as $n \to \infty$ if for all bounded continuous functions h from E to $\mathbb{R}$,

$$\int h\, dP_n \to \int h\, dP, \quad n \to \infty.$$

Recall also that $A \in \mathcal{E}$ is called a P-continuity set if $P(\partial A) = 0$ where ∂A denotes the boundary of A, and that by the Portmanteau Theorem (Theorem 11.1.1 in [8]) convergence in distribution is equivalent to

$$P_n(A) \to P(A) \text{ as } n \to \infty \text{ for all } P\text{-continuity sets } A. \tag{4.1}$$

We shall now use Corollary 3.1 to prove the Skorohod representation theorem in the separable case.

Theorem 4.1. *Let X_1, $X_2, \ldots$, X be random elements in a metric space E equipped with its Borel subsets $\mathcal{E}$. Further, let X take values almost surely in a separable subset $E_0 \in \mathcal{E}$. Then*

$$X_n \to X \text{ in distribution as } n \to \infty \tag{4.2}$$

if and only if there is a coupling $(\hat{X}_1, \hat{X}_2, \ldots, \hat{X})$ of X_1, $X_2, \ldots$, X such that

$$\hat{X}_n \to \hat{X} \text{ pointwise as } n \to \infty. \tag{4.3}$$

Proof. Let d be the metric. We begin with basic preliminaries. First, assume existence of the coupling and let h be a bounded continuous function. Then (4.3) yields that $h(\hat{X}_n) \to h(\hat{X})$ pointwise as $n \to \infty$ and by bounded convergence, $\int h\,dP_n \to \int h\,dP$ as $n \to \infty$. Thus (4.2) holds.

Conversely, assume from now on that (4.2), and thus (4.1), holds. For each $\epsilon > 0$, any separable Borel set can be covered by countably many E-balls of diameter $< \epsilon$. Note that for every $y \in E$ and $r > 0$, $\partial\{x \in E : d(y,x) < r\} \subseteq \partial\{x \in E : d(y,x) = r\}$ and that the set on the right-hand side has P-mass 0 except for countably many radii r. Thus the covering sets below may be taken to be P-continuity sets. Moreover, since $\partial(A \cap B) \subseteq \partial A \cup \partial B$ for all subsets A and B of E, the covering sets can be taken to be disjoint.

Let $A_2, A_3, \ldots$ be disjoint P-continuity sets of diameter < 1 covering E_0 and put $A_1 = E \setminus (A_2 \cup A_3 \cup \ldots)$. Then A_1 is also a P-continuity set since $P(A_1) = 0$ and since ∂A_1 cannot contain interior points of the P-continuity sets $A_2, A_3, \ldots$ Thus $\{A_i : i \in \mathbb{N}\}$ is a partition of E into P-continuity sets. Put $A_{11} = A_1$ and $A_{12} = A_{13} = \cdots = \varnothing$. For $i > 1$, let $A_{i2}, A_{i3}, \ldots$ be disjoint P-continuity subsets of A_i of diameter $< 1/2$ covering $E_0 \cap A_i$ and put $A_{i1} = A_i \setminus (A_{i2} \cup A_{i3} \cup \ldots)$. Then again $\{A_{\mathbf{i}^2} : \mathbf{i}^2 \in \mathbb{N}^2\}$ is a partition of E into P-continuity sets. Continue this recursively in $k \in \mathbb{N}$ to obtain a sequence of partitions $\{A_{\mathbf{i}^k} : \mathbf{i}^k \in \mathbb{N}^k\}$ of E into P-continuity sets such that

$$A_{\mathbf{i}^k},\ \mathbf{i}^k \in (\mathbb{N} \setminus \{1\})^k, \text{ cover } E_0 \text{ and are each of diameter } < 1/k \tag{4.4}$$

and such that the partitions are nested in the sense that for $k \in \mathbb{N}$ and $\mathbf{i}^k \in \mathbb{N}^k$ it holds that $A_{\mathbf{i}^k} = A_{\mathbf{i}^k 1} \cup A_{\mathbf{i}^k 2} \cup \ldots$

After these basic preliminaries, we are now ready to apply Corollary 3.1. Let $\mathbf{Z}_1$, $\mathbf{Z}_2, \ldots, \mathbf{Z}$ be the random elements in $(\mathbb{N}, 2^{\mathbb{N}})^{\mathbb{N}} \otimes (E, \mathcal{E})$ defined as follows (well-defined because the partitions are nested): set $Z_n^\infty = X_n$ and $Z^\infty = X$ and for $k \in \mathbb{N}$

$$\mathbf{Z}_n^k = \mathbf{i}^k \quad \text{if} \quad X_n \in A_{\mathbf{i}^k} \qquad \text{and} \qquad \mathbf{Z}^k = \mathbf{i}^k \quad \text{if} \quad X \in A_{\mathbf{i}^k}.$$

Due to (4.1), we have $\mathbf{P}(\mathbf{Z}_n^k = \mathbf{i}^k) \to \mathbf{P}(\mathbf{Z}^k = \mathbf{i}^k)$ as $n \to \infty$, $\mathbf{i}^k \in \mathbb{N}^k$, $k \in \mathbb{N}$. Thus $\mathbf{Z}_n^k \to \mathbf{Z}^k$ in density as $n \to \infty$ and Corollary 3.1 yields the existence of a coupling $(\hat{\mathbf{Z}}_1, \hat{\mathbf{Z}}_2, \ldots, \hat{\mathbf{Z}})$ of $(\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z})$, an $\mathbb{N}$-valued random variable N, and integers $0 \leqslant k_1 \leqslant k_2 \leqslant \cdots \to \infty$, such that

$$\hat{\mathbf{Z}}_n^{k_n} = \hat{\mathbf{Z}}^{k_n}, \quad n \geqslant N. \tag{4.5}$$

Now define the coupling of $X_1, X_2, \ldots, X$ by setting $\hat{X}_n = \hat{Z}_n^\infty$ and $\hat{X} = \hat{Z}^\infty$. Then (after deleting a null event) we have that $\hat{X} \in E_0$ and that for $k \in \mathbb{N}$

$$\hat{\mathbf{Z}}_n^k = \mathbf{i}^k \quad \text{if} \quad \hat{X}_n \in A_{\mathbf{i}^k} \qquad \text{and} \qquad \hat{\mathbf{Z}}^k = \mathbf{i}^k \quad \text{if} \quad \hat{X} \in A_{\mathbf{i}^k}.$$

Thus $\hat{X}_n \in A_{\hat{\mathbf{Z}}_n^{k_n}}$ and $\hat{X} \in A_{\hat{\mathbf{Z}}^{k_n}}$ for all $n \in \mathbb{N}$. Apply (4.5) to obtain that

$$\text{both } \hat{X}_n \in A_{\hat{\mathbf{Z}}^{k_n}} \text{ and } \hat{X} \in A_{\hat{\mathbf{Z}}^{k_n}} \text{ when } n \geqslant N. \tag{4.6}$$

Finally, apply (4.4): since $\hat{X} \in E_0$ we have that $\hat{\mathbf{Z}}^{k_n} \in (\mathbb{N} \setminus \{1\})^{k_n}$ so $A_{\hat{\mathbf{Z}}^{k_n}}$ has diameter $< 1/k_n$. From this and (4.6) we obtain that

$$d(\hat{X}_n, \hat{X}) < 1/k_n, \quad n \geqslant N.$$

Since $N < \infty$ and $\lim_{n\to\infty} 1/k_n = 0$ this implies that $d(\hat{X}_n, \hat{X}) \to 0$ pointwise, that is, (4.3) holds. □

5 Application to Markov chains

In this final section we shall first apply Corollary 3.1 to discrete time Markov chains with time set $\mathbb{N} \cup \{0\}$, and then apply Corollary 3.3 to continuous time Markov chains with time set $[0, \infty)$.

Theorem 5.1. *Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}$ be discrete time irreducible Markov chains on a countable state space E with initial distributions $\alpha_1, \alpha_2, \ldots, \alpha$ and with transition matrices $M_1, M_2, \ldots, M$. Then*

$$\alpha_n \to \alpha \ \text{ and } \ M_n \to M \ \text{ pointwise as } n \to \infty \tag{5.1}$$

if and only if there exists a coupling $(\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, \ldots, \hat{\mathbf{X}})$ of $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}$ such that for some $\mathbb{N}$-valued random variable N and some sequence of integers $0 \leqslant k_1 \leqslant k_2 \leqslant \cdots \to \infty$,

$$\hat{\mathbf{X}}_n^{k_n} = \hat{\mathbf{X}}^{k_n}, \quad n \geqslant N.$$

Proof. Let $p_n(i,j)$ and $p(i,j)$ be the $(i,j) \in E \times E$ entries of M_n and M. Due to irreducibility, (5.1) holds if and only if for all $k \in \mathbb{N}$ and $i_0, i_1, \ldots \in E$

$$\lim_{n\to\infty} \alpha_n(i_0) p_n(i_0, i_1) \ldots p_n(i_{k-2}, i_{k-1}) = \alpha(i_0) p(i_0, i_1) \ldots p(i_{k-2}, i_{k-1}),$$

that is, if and only if

$$\forall k \in \mathbb{N}: \quad \mathbf{X}_n^{k-1} \to \mathbf{X}^{k-1} \text{ in density as } n \to \infty.$$

The desired result now follows from Corollary 3.1 by taking $Z_n^k = X_n^{k-1}$ and $Z^k = X^{k-1}$ for $k, n \in \mathbb{N}$ and letting Z_n^∞ and Z^∞ be arbitrary fixed states. □

Theorem 5.1 is an immediate consequence of Corollary 3.1 because the finite segments $\mathbf{X}_1^k, \mathbf{X}_2^k, \ldots \mathbf{X}^k$ are discrete. In the continuous time case the finite segments are not discrete so the argument becomes more involved.

Theorem 5.2. *Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}$ be continuous time irreducible nonexplosive Markov chains on a countable state space E with initial distributions $\alpha_1, \alpha_2, \ldots, \alpha$ and intensity matrices $C_1, C_2, \ldots, C$. Then*

$$\alpha_n \to \alpha \ \text{ and } \ C_n \to C \ \text{ pointwise as } n \to \infty \tag{5.2}$$

if and only if there exists a coupling $(\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, \ldots, \hat{\mathbf{X}})$ of $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}$ such that for some $\mathbb{N}$-valued random variable N and some sequence of real numbers $0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \to \infty$,

$$\hat{\mathbf{X}}_n^{t_n} = \hat{\mathbf{X}}^{t_n}, \quad n \geqslant N.$$

Proof. Let $c(i,j)$ be the $(i,j) \in E \times E$ entry of C. Let $c(i) = \sum_{j \neq i} c(i,j)$ be the total intensity in state $i \in E$. For $i \neq j$ let $p(i,j) = c(i,j)/c(i)$ be the jump probability from i to j. Let $Y^0, Y^1, \ldots$ be the states visited by $\mathbf{X}$. Let $S^1, S^2, \ldots$ be the times when $\mathbf{X}$ enters the states $Y^1, Y^2, \ldots$ Let $K(t)$ be the last k such that $S^k < t$. Since $\mathbf{X}$ is nonexplosive, $K(t)$ is a.s. finite. Let $c_n(i)$, $p_n(i,j)$, $Y_n^0, Y_n^1, \ldots$, $S_n^1, S_n^2, \ldots$ and $K_n(t)$ be obtained in the same way from C_n and $\mathbf{X}_n$.

Both $\big(Y_n^0, \ldots, Y_n^{K_n(t)}, S_n^1, \ldots, S_n^{K_n(t)}\big)$ and $\big(Y^0, \ldots, Y^{K(t)}, S^1, \ldots, S^{K(t)}\big)$ take values in the union $A^{(t)} = \bigcup_{k=0}^\infty A^{(t,k)}$ of the disjoint sets

$$A^{(t,k)} = E^{k+1} \times B^{(t,k)} \text{ where } B^{(t,k)} = \{(s_1, \ldots, s_k) : 0 \leqslant s_1 < \cdots < s_k < t\}.$$

Let $\lambda^{(t)}$ be the measure on $A^{(t)}$ defined by $\lambda^{(t)}(A^{(t,k)} \cap \cdot) = \mu^{(t,k)}$ where $\mu^{(t,k)}$ is the product of counting measure on E^{k+1} and Lebesgue measure on $B^{(t,k)}$. On $A^{(t,k)}$, the density $f^{(t)}$ of $\big(Y^0, \dots, Y^{K(t)}, S^1, \dots, S^{K(t)}\big)$ with respect to $\lambda^{(t)}$ is

$$f^{(t)}(i_0, \dots, i_k, s_1, \dots, s_k) = \alpha(i_0)p(i_0, i_1)\dots p(i_{k-1}, i_k)$$
$$c(i_0)\dots c(i_{k-1})e^{-c(i_0)s_1}\dots e^{-c(i_{k-1})(s_k-s_{k-1})}e^{-c(i_k)(t-s_k)}$$

and the density $f_n^{(t)}$ of $\big(Y_n^0, \dots, Y_n^{K_n(t)}, S_n^1, \dots, S_n^{K_n(t)}\big)$ with respect to $\lambda^{(t)}$ is

$$f_n^{(t)}(i_0, \dots, i_k, s_1, \dots, s_k) = \alpha_n(i_0)p_n(i_0, i_1)\dots p_n(i_{k-1}, i_k)$$
$$c_n(i_0)\dots c_n(i_{k-1})e^{-c_n(i_0)s_1}\dots e^{-c_n(i_{k-1})(s_k-s_{k-1})}e^{-c_n(i_k)(t-s_k)}.$$

Note that $\lim_{n\to\infty} c_n(i)e^{-c_n(i)x} = c(i)e^{-c(i)x}$ holds for all $x \geqslant 0$ if and only if $\lim_{n\to\infty} c_n(i) = c(i)$ and if and only if $\liminf_{n\to\infty} c_n(i)e^{-c_n(i)x} = c(i)e^{-c(i)x}$ holds for all $x \geqslant 0$. This and irreducibility implies that (5.2) holds if and only if for all $t \in [0, \infty)$, $\liminf_{n\to\infty} f_n^{(t)} = f^{(t)}$. Now $\big(Y_n^0, \dots, Y_n^{K_n(t)}, S_n^1, \dots, S_n^{K_n(t)}\big)$ and $\big(Y^0, \dots, Y^{K(t)}, S^1, \dots, S^{K(t)}\big)$ are random elements in a common space and $\mathbf{X}_n^t$ and $\mathbf{X}^t$ are random elements in a common space, and since these two spaces are Borel equivalent we obtain that (5.2) holds if and only if

$$\forall t \in [0, \infty) : \quad \mathbf{X}_n^t \to \mathbf{X}^t \text{ in density as } n \to \infty.$$

The theorem now follows from Corollary 3.3 by taking $Z_n^t = X_n^t$ and $Z^t = X^t$ for $t \in [0, \infty)$, $n \in \mathbb{N}$, and letting Z_n^∞ and Z^∞ be arbitrary fixed states. $\square$

References

[1] Ash, R.B. (1972). *Real Analysis and Probability.* Academic Press, New York. MR-0435320

[2] Berti, P., Pratelli, L., Rigo, P. (2007). Skorohod representation on a given probability space. *Prob. Theo. Relat. Fields* **137**, 277–288. MR-2278458

[3] Berti, P., Pratelli, L., Rigo, P. (2010). Skorohod representation theorem via disintegrations. *Sakhya* **72**, 208–220. MR-2658171

[4] Berti, P., Pratelli, L., Rigo, P. (2011). A Skorohod representation theorem for uniform distance. *Prob. Theo. Relat. Fields* **150**, 321–335. MR-2800912

[5] Berti, P., Pratelli, L., Rigo, P. (2013). A Skorohod representation theorem without separability. *Electron. Commun. Probab.* **18**, 1–12. MR-3125256

[6] Berti, P., Pratelli, L., Rigo, P. (2015). Glueing lemmas and Skorohod representations. *Electron. Commun. Probab.* **20**, 1–11. MR-3374303

[7] Dudley, R.M. (1968). Distances of probability measures and random variables. *Annals of Mathematical Statistics* **39**, 1563–1572. MR-0230338

[8] Dudley, R.M. (2002). *Real Analysis and Probability.* Cambridge University Press, Cambridge. MR-1932358

[9] Sethuraman, J. (2002). Some extensions of the Skorohod representation theorem. *Sankhya* **64**, 884–893. MR-1981517

[10] Skorohod, A.V. (1956). Limit theorems for stochastic processes. *Theo. Probab. Appl.* **1**, 261–290. MR-0084897

[11] Thorisson, H. (1995). Coupling methods in probability theory. *Scand. J. Statist.* **22**, 159–182. MR-1339749

[12] Thorisson, H. (2000). *Coupling, Stationarity, and Regeneration.* Springer, New York. MR-1741181

[13] Wichura, M.J. (1970). On the construction of almost uniformly convergent random variables with given weakly convergent image laws. *Annals of Mathematical Statistics* **41**, 284–291. MR-0266275

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS[1])
- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]
- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

[1]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2]EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4]BS: Bernoulli Society http://www.bernoulli-society.org/
[5]Project Euclid: https://projecteuclid.org/
[6]IMS Open Access Fund: http://www.imstat.org/publications/open.htm