On tightness of probability measures on Skorokhod spaces

Michael A. Kouritzin [a,*]

[a] *Department of Mathematical and Statistical Sciences*

University of Alberta, Edmonton, AB T6G 2G1 Canada

Abstract

The equivalences to and the connections between the modulus-of-continuity condition, compact containment and tightness on $D_E[a,b]$ with $a < b$ are studied. The results within are tools for establishing tightness for probability measures on $D_E[a,b]$ that generalize and simplify prevailing results in the cases that E is a metric space, nuclear space dual or, more generally, a completely regular topological space. Applications include establishing weak convergence to martingale problems, the long-time typical behavior of nonlinear filters and particle approximation of cadlag probability-measure-valued processes. This particle approximation is studied herein, where the distribution of the particles is the underlying measure-valued process at an arbitrarily fine discrete mesh of points.

Key words: Tightness, Skorokhod topology, Compact Containment, Modulus of Continuity, Measure-valued process, Exchangeable particles, Strassen-Dudley coupling

2010 MSC: primary 60B05; secondary 60B10

1 Introduction

Suppose that E is a topological space, $a < b$ and $\{P_n\}$, P are Borel probability measures on the space $D_E[a, b]$ of E-valued right-continuous, left-hand-limit (cadlag) functions of $[a, b]$ with the Skorokhod topology (introduced in the separable-metric-space case by Skorokhod [26] and extended to more general spaces in Mitoma [21] and Jakubowski [14] for example). Then, one often deduces that P_n converges weakly to P by first showing the family $\{P_n\}$ is tight, which implies relative compactness on Hausdorff spaces, and then identifying P as the unique limit point. Tightness is also a key property for establishing weak convergence in multiple parameters as in the long time performance of approximate filters (see Budhiraja and Kushner [5]) and for establishing existence of probability measures on $D_E[a, b]$ (see Theorem 13.6 of Billingsley [3]). Previously in [4], we used homeomorphic methods to generalize several basic weak convergence and measure separation results from Ethier and Kurtz [10], Billingsley [3] and Kallianpur and Xiong [15]. However, we did not consider tightness nor the related problems of establishing compact containment and modulus of continuity.

Establishing tightness of probability measures on topological spaces can be a challenging problem. For example, Dawson [6], Perkins [24], Kurtz [18] and Kallianpur and Xiong [15], to name just a few, all spend considerable effort proving tightness for random variables on exotic spaces. Kallianpur and Xiong [15] include good basic material on weak convergence and tightness, from which our present work has benefitted. Indeed, our motivation was to develop tightness criterion in Skorokhod spaces systematically after the important contributions by Bhatt and Karandikar [1] and Kallianpur and Xiong [15] and use these criterion to establish general particle approximation of cadlag probability-measure-valued processes.

Mitoma [21] established a fundamental tightness result for probability measures on cadlag

* Corresponding author.
Email address: michaelk@ualberta.ca (Michael A. Kouritzin).

spaces of tempered distributions $D_{\mathcal{S}'(\mathbb{R}^d)}[0,1]$. Jakubowski [14] then generalized Mitoma's result to any completely regular space E instead of just $\mathcal{S}'(\mathbb{R}^d)$. In particular, Jakubowski assumes that $\{P_\alpha\}$ is a collection of probability measures satisfying the compact containment condition CCC such that $P_\alpha(\widetilde{f})^{-1}$ is tight for a set $\widetilde{\mathbb{F}}$ of continuous functions $\widetilde{f} : D_E[0,1] \to D_{\mathbb{R}}[0,1]$ and establishes conditions on $\widetilde{\mathbb{F}}$ that imply $\{P_\alpha\}$ itself is tight. While CCC can basically be finessed in the tempered distribution setting, it is otherwise often hard to establish directly, meaning considerable work may be needed prior to applying Jakubowski's results.

Herein, we obtain tightness results for $D_E[a,b]$-valued random variables, where E is a completely regular space. Our developments further Jakubowski [14] and show that the assumption E has metrizable compacts, together with the compact containment condition, reduces the tightness problem on $D_E[a,b]$ to the case where E is a separable metric space, which we also deal with. Our results trivially apply to recover Mitoma's result that states test function tightness on $D_{\mathbb{R}}[0,1]$ implies tightness on $D_{\mathcal{S}'(\mathbb{R}^d)}[0,1]$. We also study equivalences to the modulus of continuity condition (MCC) as well as what in addition to MCC must be assumed for tightness. This leads to a nice *function-by-function* method for establishing tightness that avoids explicitly verifying CCC. To illustrate this, we take any cadlag probability-measure-valued process $V = \{V_t,\ t \geq 0\}$ on a Polish space and construct cadlag E-valued Markov processes Ξ^N, whose laws $\mathcal{L}(\Xi^N_t)$ are almost surely equal to V_t at an arbitrarily fine partition of times $\{s^N_n\}$, such that the empirical measures $\dfrac{1}{m_N}\displaystyle\sum_{j=1}^{m_N}\Xi^{j,N}$ of conditionally independent copies of Ξ^N converge in path space in probability to V. Historically, measure-valued processses were constructed as the limit of empirical measures. However, some of the most general construction results now use alternative methods (see e.g. Fitzsimmons [12], Kouritzin and Long [17]). Our application herein shows, at least for probability-measure processes on Polish spaces, that (exchangeable) particle approximation is always available. In additional to its inherent interest, this fact could be important for such things as simulation and establishing support properties of measure-valued processes

(see e.g. Evans and Perkins [11], Liu and Zhou [19]).

Section 2 contains notation and a motivating application. Section 3 provides background on completely regular spaces. Section 4 houses modulus of continuity and compact containment conditions as well as equivalence results. Section 5 has containment and tightness results on completely regular spaces, assuming a modulus of continuity condition. Section 6 deals with minimal modulus of continuity conditions for pathspace tightness when compact containment holds. The appendix contains necessary technical results.

2 Notation and Application

Herein, $\mathbb{N} = \{1, 2, ...\}$ and $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$; $\mathcal{B}(E)$ are the Borel sets; $\mathbb{M}(E)$, $\mathbb{M}_F(E)$, $\mathbb{P}(E)$ are respectively the Borel measures, finite measures and probability measures; and $M(E)$, $B(E)$, $C(E)$, $\overline{C}(E)$ are respectively the Borel measurable, bounded measurable, continuous, and continuous bounded $\mathbb{R}$-valued functions on topological space E. When $g \in B(E)$ and $\mu \in \mathbb{M}_F(E)$, we let $\hat{g}(\mu) = \mu(g) = \int g d\mu$ be a function of μ. Further, when (E, r) is a metric space, we let $BL(E)$ denote the bounded, Lipschitz continuous $\mathbb{R}$-valued functions on E with $\|g\|_{BL} = \sup_x |g(x)| + \sup_{y \neq x} \frac{|g(y)-g(x)|}{r(x,y)}$ and $A^\eta \triangleq \{y \in E : r(x, y) < \eta \text{ for some } x \in A\}$ for any $\eta > 0$ and $A \subset E$. π_t denotes the projection function from $D_E[a, b]$ to E for $t \in [a, b]$ and $J(f) = \{t \in (a, b] : f_{t-} \neq f_t\}$ for $f \in D_E[a, b]$. Finally, $\mathcal{L}(X)$ denotes the law of a random element X and we use the following extended Vinogradov symbol (also used in [16]): Suppose $q(n, m), r(n, m)$ are expressions depending upon two variables n, m. Then,

$$q(n, m) \overset{n}{\ll} r(n, m) \text{ means } \exists\ c_m > 0 \text{ such that } q(n, m) \leq c_m r(n, m) \quad \forall\ n, m.$$

For clarity, c_m depends only on m.

In essence, our results in the sequel are simple tools for establishing tightness for existence and weak convergence results, generalizing prior, heavily-cited results in e.g. Mitoma [21], Jakubowski [14], Ethier and Kurtz [10] and Billingsley [3]. However, we investigate the breadth of particle approximation in measure-valued processes as an illustrative example. We construct our approximations using Theorem 1 below, a version of the Strassen-Dudley coupling theorem naturally expressed in terms of the following Prohorov metric variant.

Definition 1 *Let E be a metric space and $p \geq 1$. Then, we define the p-Prohorov metric on $\mathcal{P}(E)$ to be $\rho_p(P,Q) = \inf\{\epsilon > 0 : P(F) \leq Q(F^\epsilon) + \epsilon^p \ \forall \text{ closed } F\}$.*

The requirement $p \geq 1$ ensures that $\epsilon^p + \delta^p \leq (\epsilon + \delta)^p$ for $\epsilon, \delta > 0$ so one can adapt the standard argument establishing that the Prohorov metric is truly a metric to our setting. Moreover, these p-Prohorov metrics still topologize weak convergence when E is a separable metric space. ρ_p becomes the standard Prohorov metric, denoted ρ herein, when $p = 1$.

Theorem 1 *Let (E,d) be a separable metric space, $p \geq 1$, $\mu \in \mathbb{P}(E)$, and ξ_1 be an E-valued random variable on some probability space $(\Omega_1, \mathcal{F}_1, P_1)$. Then, there exists a non-negative collection $\{g_m\}_{m=1}^\infty \subset BL(E)$ that is closed under multiplication, satisfies $\sup_m \|g_m\|_{BL} \leq 1$ and strongly separates points (s.s.p.); another space $(\Omega_2, \mathcal{F}_2, P_2)$ and an E-valued random variable ξ_2 on $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, P_1 \times P_2)$ such that $\mathcal{L}(\xi_2) = \mu$, $E[g(\xi_2)|\mathcal{F}_1 \otimes \{\emptyset, \Omega_2\}] = E[g(\xi_2)|\sigma(\xi_1) \otimes \{\emptyset, \Omega_2\}]$ for all $g \in B(E)$ and*

$$E\left[|g_m(\xi_2) - g_m(\xi_1)|^p \middle| \mathcal{F}_1 \otimes \{\emptyset, \Omega_2\}\right] \leq 3\left(\frac{\pi^2}{12} m^2 \ \rho_p(\mathcal{L}(\xi_1), \mu)\right)^p \ \forall m \in \mathbb{N}.$$

Proof. As this is part of our motivating application, we refer to results that follow. The existence of the $\{g_m\}_{m=1}^\infty$ follows from (23) and Lemma 5 (to follow). These g_m are uniformly bounded by $\frac{1}{2}$ and also have a Lipschitz constant of $\frac{1}{2}$. Then, $G \doteq (g_1, g_2, ...) : E \to G(E) \subset \left[0, \frac{1}{2}\right]^\infty \subset \mathbb{R}^\infty$ is a homeomorphism by Lemma 4 (to follow). $G(E)$ is a separable metric

space with metric $r(\theta,\zeta)=\frac{12}{\pi^2}\sum_{m=1}^{\infty}\frac{1}{m^2}|\theta^m-\zeta^m|\leq 1$ and we have that

$$r(G(x),G(z))=\frac{12}{\pi^2}\sum_{m=1}^{\infty}\frac{1}{m^2}|g_m(x)-g_m(z)|\leq\frac{6}{\pi^2}\sum_{m=1}^{\infty}\frac{1}{m^2}d(x,z)=d(x,z).$$

Hence, it follows that $G(F^\gamma)\subset G(F)^\gamma$ for any $\gamma>0$, (closed) F and the p-Prohorov distance between the law of $X\doteq G(\xi_1)$ and $Q\doteq\mu G^{-1}$ satisfies $\rho_p(\mathcal{L}(X),Q)\leq\rho_p\left(\mathcal{L}(\xi_1),\mu\right)$. Let $\rho_p\left(\mathcal{L}(\xi_1),\mu\right)>0$ since otherwise we could just take $\xi_2=\xi_1$. Now, we let $n\in\mathbb{N}$ and partition $\left[0,\frac{1}{2}\right]$ into

$$A^1=\left[0,\frac{1}{2n}\right]\text{ and }A^i=\left(\frac{i-1}{2n},\frac{i}{2n}\right]$$

for $i=2,3,...,n$. Next, we consider the n^n disjoint sets of the form

$$B^I=A^{i_1}\times A^{i_2}\times\cdots\times A^{i_n}\times\mathbb{R}^\infty,$$

where $I=(i_1,i_2,...,i_n)$ with $i_j\in\{1,...,n\}$. Hence, the n^n disjoint sets

$$E^I=B^I\cap G(E)\in\mathcal{B}(G(E))=\mathcal{B}(\mathbb{R}^\infty)\cap G(E)$$

partition $G(E)$ and have diameter no more than $\delta_n\doteq\frac{1}{2n}+\frac{6}{\pi^2}\frac{1}{n}$. ($G(E)$ is a Borel set if E is also complete by Parthasarathy [23, Corollary I.3.3]. However, we do not need that fact.) Now, by the proof of Ethier and Kurtz [10, Lemma 3.1.3] (with $S=G(E),N=n^n,E_0=\emptyset,\delta=\delta_n$) for all $\epsilon>\rho_p\left(\mathcal{L}(X),Q\right)$ there is a random variable $Z^\epsilon\in H\doteq(G(E)^{n^n+1}\times[0,1])$ on an independent probability space $(\Omega_2^\epsilon,\mathcal{F}_2^\epsilon,P_2^\epsilon)$ and a measurable mapping $h:G(E)\times H\to G(E)$ such that $Y^\epsilon\doteq h(X,Z^\epsilon)$ satisfies $\mathcal{L}(Y^\epsilon)=Q$ and $\{(r(X,Y^\epsilon)\geq\epsilon+\delta_n\}\subset C\in\mathcal{F}_2^\epsilon$ (see (3.1.28) of [10] and note $E_0=\emptyset$), where $P_2^\epsilon(C)<\epsilon^p$. Setting $\xi_2^\epsilon=G^{-1}(Y^\epsilon)$, we thereby get

$$E[g(\xi_2^\epsilon)|\mathcal{F}_1\otimes\{\emptyset,\Omega_2^\epsilon\}]=\int_{\Omega_2^\epsilon}g(G^{-1}\circ h(X,z))P_2^\epsilon(dz)=E[g(\xi_2^\epsilon)|\sigma(\xi_1)\otimes\{\emptyset,\Omega_2^\epsilon\}]\quad\text{and}$$

$$E\left[\left|g_m\left(\xi_2^\epsilon\right)-g_m\left(\xi_1\right)\right|^p\right|\mathcal{F}_1\otimes\{\emptyset,\Omega_2^\epsilon\}\Big]=E\left[\left|Y_m^\epsilon-X_m\right|^p\right|\mathcal{F}_1\otimes\{\emptyset,\Omega_2^\epsilon\}\Big] \tag{1}$$
$$\leq\left(\frac{\pi^2}{12}\right)^p m^{2p}E\left[r^p(Y^\epsilon,X)\right|\mathcal{F}_1\otimes\{\emptyset,\Omega_2^\epsilon\}\Big]$$
$$<\left(\frac{\pi^2}{12}\right)^p m^{2p}\left((\epsilon+\delta_n)^p+\epsilon^p\right).$$

The result follows by letting n be large enough, ϵ be close enough to $\rho_p\left(\mathcal{L}(X),Q\right)$ and using the fact $\rho_p(\mathcal{L}(X),Q)\leq\rho_p\left(\mathcal{L}(\xi_1),\mu\right)$. □

Now, we establish our compact-time-interval measure-valued-process representation result. An advantage of this coupling approach is that our particles are Markov processes.

Theorem 2 *Let $-\infty<a<b<\infty$; E be a Polish space; and $\{V_t,t\in[a,b]\}$ be a process on $(\Omega,\mathcal{F},P)$ with $D_{\mathbb{P}(E)}[a,b]$-valued paths. Then, there are numbers $m_N\to\infty$ and an enlarged probability space supporting conditionally-independent and identically-distributed E-valued cadlag Markov processes $\{\Xi^{1,N},...,\Xi^{m_N,N}\}$ such that the empirical processes $V_t^N\doteq\frac{1}{m_N}\sum_{i=1}^{m_N}\delta_{\Xi_t^{i,N}}$ converge in probability to V on $D_{\mathbb{P}(E)}[a,b]$ as $N\to\infty$.*

Remark 1 *Here and below, we sometimes use V_t^ω to denote the probability measure at time t and random occurrence $\omega\in\Omega$. The enlarged probability space has the form*

$$\left(\overline{\Omega}\times\Omega,\overline{\mathcal{F}}\otimes\mathcal{F},Q^{V(\omega)}(d\overline{\omega})P(d\omega)\right)$$

and conditionally independent means given $\{\emptyset,\overline{\Omega}\}\otimes\mathcal{F}$. *This* conditional independence *together with the identical distribution of the copies implies that our particles are* exchangeable.

Remark 2 *The proof to follow will reveal: 1) the empirical measures V^N are non-anticipative. 2) m_N can be taken arbitrarily (provided $m_N\to\infty$) if there exist $\beta(\omega),C(\omega)>0$ such that $\rho_p(V_t^\omega,V_s^\omega)\leq C(\omega)|t-s|^{\beta(\omega)}$ a.s. In this case, we would replace the definition of s_k^N below with $s_k^N=a+(b-a)k2^{-N}$ for $k=0,1,...,l(N)=2^N$.*

Remark 3 *The original purpose of this result was to illustrate the methods developed in Sections 4 and 5 of this paper. Notice in the sequel that, after set up, there are basically three steps. Moreover, there is no need to verify any compact containment condition explicitly. Finally, the verification of the modulus of continuity condition is made relatively easy.*

Proof. Set Up Part a: Define particles and empirical measure approximation.
Let $\{g_n\} \subset \mathcal{B}L(E)$ be the collection given by Theorem 1, take $\{v_N\}$ to be positive numbers decreasing to 0 and define the stopping times (with respect to $\mathcal{F}_t^V \doteq \sigma\{V_s, s \leq t\}$)

$$\tau_0^N = a,\ \tau_i^N = \inf\{t > \tau_{i-1}^N : \max_{k \leq N} |V_t(g_k) - V_{\tau_{i-1}^N}(g_k)| > v_N\} \wedge (\tau_{i-1}^N + v_N) \wedge b\ \forall i \geq 1 \quad (2)$$

$$\{s_k^N\}_{k=0}^{l(N)} = \bigcup_{M=1}^{N} \{\tau_i^M\}_{i=1}^{\infty} \quad (3)$$

so $a = s_0^N < s_1^N < \cdots < s_{l(N)}^N = b$ and $\{s_k^N\}_{k=0}^{l(N)} \subset \{s_k^{N+1}\}_{k=0}^{l(N+1)}$. Let $\delta_N = \min_i\{s_i^N - s_{i-1}^N\}$, set $|\rho^N|_p = \left[\sum_{k=0}^{l(N)-1} \left(\rho_p(V_{s_{k+1}^N}, V_{s_k^N})\right)^p\right]^{\frac{1}{p}}$ for some $p \geq 2$ and take numbers m_N such that

$$P\left(\lim_{N\to\infty} \frac{|\rho^N|_p^2\ \delta_N^{-\frac{2}{q}}}{m_N} \leq 1\right) = 1, \quad (4)$$

where $\frac{1}{p} + \frac{1}{q} = 1$. V is a $D_{\mathbb{P}(E)}[a,b]$-valued random variable by [10, Theorem 3.1.7 and Proposition 3.7.1]. Using the *quenched* approach, we fix a $V \in D_{\mathbb{P}(E)}[a,b]$ such that

$$\lim_{N\to\infty} \frac{|\rho^N|_p^2\ \delta_N^{-\frac{2}{q}}}{m_N} \leq 1. \quad (5)$$

Since we will later average over V, we note that each $V_{s_k^N}$ is $\sigma(V)$-measurable. Now, using Theorem 1 for each j,k starting with $k = 0$ and independent $\{\xi_0^{j,N}\}_{j=1}^{m_N}$, we construct independent E-valued Markov chains $\{\xi_{s_k^N}^{j,N},\ k = 0,1,...,l(N)\}_{j=1}^{m_N}$ on some probability space $(\overline{\Omega}, \overline{\mathcal{F}}, Q^V)$ such that $\mathcal{L}(\xi_{s_k^N}^{j,N}) = V_{s_k^N}$ for all $k = 0,1,...,l(N); j = 1,...,m_N$ and

$$E\left\{\left|g_n\left(\xi_{s_{k+1}^N}^{j,N}\right) - g_n\left(\xi_{s_k^N}^{j,N}\right)\right|^p |\mathcal{F}_{s_k^N}^N\right\} \leq 3\left(\frac{\pi^2}{12} n^2\ \rho_p\left(V_{s_k^N}, V_{s_{k+1}^N}\right)\right)^p \quad (6)$$

for all $n \in \mathbb{N}$, $j = 1, ..., m_N$ and $k = 0, 1, ..., l(N) - 1$, where $\mathcal{F}^N_{s^N_k} = \sigma\{\xi^{j,N}_{s^N_0}, ..., \xi^{j,N}_{s^N_k}, j = 1, 2, ..., m_N\}$. In particular, this implies by Jensen's inequality that

$$\left| E\left\{ g_n\left(\xi^{j,N}_{s^N_{k+1}}\right) - g_n\left(\xi^{j,N}_{s^N_k}\right)\right\}\right|^p \leq 3\left(\frac{\pi^2}{12} n^2 \ \rho_p\left(V_{s^N_k}, V_{s^N_{k+1}}\right)\right)^p \tag{7}$$

for all $n \in \mathbb{N}$, $j = 1, ..., m_N$ and $k = 0, 1, ..., l(N) - 1$. Next, we set $t_N = \max\{s^N_n : s^N_n \leq t\}$ for $t \geq a$, define $\Xi^{j,N}_t = \xi^{j,N}_{t_N}$ for $t \geq a$ and recall $V^N_t \doteq \frac{1}{m_N} \sum_{j=1}^{m_N} \delta_{\Xi^{j,N}_t}$. Each $\{V^N_t,\ t \geq a\}$ and each $\{\Xi^{j,N}_t,\ t \geq a\}$ is a cadlag $\{\mathcal{G}^N_t\}$-Markov process that is constant between jumps at $\{s^N_n\}_{n=0}^{l(N)}$ and after b for each j, N, where $\mathcal{G}^N_t \doteq \mathcal{F}_{t_N}$. The expectations here and below are with respect to Q^V, i.e. with V fixed.

Set Up Part b: Explain metric-space compactification of space of probability measures. By Example 1 (to follow), the countable collection $\{\hat{g}_n\}$, defined by $\hat{g}_n(\mu) = \int g_n d\mu$, s.s.p. on $\mathbb{P}(E)$ so $\hat{G} \doteq (\hat{g}_1, \hat{g}_2, ...) : \mathbb{P}(E) \to \hat{G}(\mathbb{P}(E)) \subset \left[0, \frac{1}{2}\right]^\infty \subset \mathbb{R}^\infty$ is a homeomorphism by Lemma 4 (to follow) and $\varrho(\mu, \nu) = \sqrt{\sum_{n=1}^\infty 2^{-n} |\hat{g}_n(\mu) - \hat{g}_n(\nu)|^2}$ defines an alternate metric for weak convergence of probability measures on E. Further, by Ethier and Kurtz [10, Theorem 3.1.7] and Theorem 6 herein, $\mathbb{P}(E)$ has a (Stone-Čech) metric space compactification $\overline{\mathbb{P}(E)}$ such that $\mathbb{P}(E)$ is a Borel subset of $\overline{\mathbb{P}(E)}$ and the homeomorphism $\hat{G}$ can be extended to a homeomorphism $\overline{G} : \overline{\mathbb{P}(E)} \to \overline{G}(\overline{\mathbb{P}(E)}) \subset \mathbb{R}^\infty$. (We are using different metrics for the topology of weak convergence. We just need to know that there is one complete metric, the normal Prohorov metric here, to conclude $\mathbb{P}(E)$ is a Borel subset.) Now, we extend the homeomorphism $\hat{G}$ to a homeomorphism $\tilde{G} : D_{\overline{\mathbb{P}(E)}}[a, b] \to D_{\overline{G}(\overline{\mathbb{P}(E)})}[a, b] \subset D_{\mathbb{R}^\infty}[a, b]$ by

$$\tilde{G}(x)(t) = \overline{G}(x(t)) \ \text{ and } \ \tilde{G}^{-1}(y)(t) = \overline{G}^{-1}(y(t)) \tag{8}$$

(see [10, Problem 3.11.13] for continuity). Then, $D_{\mathbb{P}(E)}[a, b]$ is Borel measurable in $D_{\overline{\mathbb{P}(E)}}[a, b]$ by the development of Ethier and Kurtz [10, Theorem 3.5.6], Parthasarathy [23, Corollary I.3.3] and the argument in Theorem 6. We imposed our Polish space condition in this theorem to obtain this measureability and use [10, Corollary 3.3.2] in lieu of a lot of work below.

Step 1: Establish a modulus of continuity condition for empirical measures.

Define the $D_{\mathbb{R}}[a,b]$-functions $\zeta_t^{N,n} = V_{t_N}(g_n)$ and fix $(g,\zeta^N) \in \{(g_n,\zeta^{N,n})\}_{n=1}^{\infty} \cup \{(g_l+g_n, \zeta^{N,l}+\zeta^{N,n})\}_{l,n=1}^{\infty}$. Given $\eta > 0$, we use (2,3) to choose M such that $\sup_{s,t\in[s_{i-1}^M, s_i^M)} |V_t(g) - V_s(g)| < \eta$ and note $\sup_{s,t\in[s_{i-1}^M, s_i^M)} |\zeta_t^N - \zeta_s^N| = 0$ when $N \leq M$ and $\sup_{s,t\in[s_{i-1}^M, s_i^M)} |\zeta_t^N - \zeta_s^N| \leq \sup_{s,t\in[s_{i-1}^M, s_i^M)} |V_t(g) - V_s(g)| < \eta$ otherwise. Hence, $\{\zeta^N\}$ satisfies the modulus of continuity condition (MCC) defined in Section 4 below as processes. Next, for $a \leq s < t \leq b$

$$
\begin{aligned}
|\hat{g}\left(V_t^N\right) - \hat{g}\left(V_s^N\right) - \zeta_t^N + \zeta_s^N| &= \left| \frac{1}{m_N} \sum_{j=1}^{m_N} \left[\overline{g}\left(\xi_{t_N}^{j,N}\right) - \overline{g}\left(\xi_{s_N}^{j,N}\right) \right] \right| \\
&= \left| \int_{s_N}^{t_N} \frac{1}{\delta_N m_N} \sum_{j=1}^{m_N} \left\{ \overline{g}\left(\Xi_{u+\delta_N}^{j,N}\right) - \overline{g}\left(\Xi_u^{j,N}\right) \right\} du \right| \\
&\leq \frac{|t_N - s_N|^{\frac{1}{q}}}{\delta_N m_N} \left[\int_a^b \left| \sum_{j=1}^{m_N} \left\{ \overline{g}\left(\Xi_{u+\delta_N}^{j,N}\right) - \overline{g}\left(\Xi_u^{j,N}\right) \right\} \right|^p du \right]^{\frac{1}{p}}
\end{aligned} \tag{9}
$$

by Holder's inequality, where $\overline{g}\left(\xi_u^{j,N}\right) = g\left(\xi_u^{j,N}\right) - E[g\left(\xi_u^{j,N}\right)]$ and q satisfies $\frac{1}{p} + \frac{1}{q} = 1$. Next, by the fact $p \geq 2$, Jensen's inequality and the Marcinkiewicz-Zygmund inequality

$$
\begin{aligned}
& E\left[\left[\int_a^b \left| \sum_{j=1}^{m_N} \left\{ \overline{g}\left(\Xi_{u+\delta_N}^{j,N}\right) - \overline{g}\left(\Xi_u^{j,N}\right) \right\} \right|^p du \right]^{\frac{1}{p}} \right] \\
&= E\left[\left[\sum_k \int_{s_{k+1}^N - \delta_N}^{s_{k+1}^N} \left| \sum_{j=1}^{m_N} \left\{ \overline{g}\left(\Xi_{u+\delta_N}^{j,N}\right) - \overline{g}\left(\Xi_u^{j,N}\right) \right\} \right|^p du \right]^{\frac{1}{p}} \right] \\
&= \delta_N^{\frac{1}{p}} E\left[\left[\sum_k \left| \sum_{j=1}^{m_N} \left\{ \overline{g}\left(\xi_{s_{k+1}^N}^{j,N}\right) - \overline{g}\left(\xi_{s_k^N}^{j,N}\right) \right\} \right|^p \right]^{\frac{1}{p}} \right] \\
&\overset{N}{\ll} \delta_N^{\frac{1}{p}} \left\{ \sum_k E\left[\left[\sum_{j=1}^{m_N} \left| \left\{ \overline{g}\left(\xi_{s_{k+1}^N}^{j,N}\right) - \overline{g}\left(\xi_{s_k^N}^{j,N}\right) \right\} \right|^2 \right]^{\frac{p}{2}} \right] \right\}^{\frac{1}{p}} \\
&\leq \delta_N^{\frac{1}{p}} m_N^{\frac{1}{2}} \left\{ \sum_k E\left[\frac{1}{m_N} \sum_{j=1}^{m_N} \left| \left\{ \overline{g}\left(\xi_{s_{k+1}^N}^{j,N}\right) - \overline{g}\left(\xi_{s_k^N}^{j,N}\right) \right\} \right|^p \right] \right\}^{\frac{1}{p}}.
\end{aligned} \tag{10}
$$

Thus, by Jensen's inequality and (6,7,9,10) there are $K_1, K_2 > 0$ independent of δ, N so that

$$
\begin{aligned}
&E\left[\sup_{\substack{a\le s<t\le b\\ t-s\le\delta}}|\hat{g}\big(V_t^N\big)-\hat{g}\big(V_s^N\big)-\zeta_t^N+\zeta_s^N|\right] \qquad (11)\\
&\le K_1\frac{\nu(\delta,N)}{\delta_N^{\frac{1}{q}}m_N^{\frac{1}{2}}}\left\{\sum_k E\left[\frac{1}{m_N}\sum_{j=1}^{m_N}\left|\overline{g}\left(\xi^{j,N}_{s^N_{k+1}}\right)-\overline{g}\left(\xi^{j,N}_{s^N_k}\right)\right|^p\right]\right\}^{\frac{1}{p}}\\
&\le K_2\frac{\nu(\delta,N)|\rho^N|_p}{\delta_N^{\frac{1}{q}}m_N^{\frac{1}{2}}},
\end{aligned}
$$

where $\nu(\delta,N)=\sup_{\substack{a\le s<t\le b\\ t-s\le\delta}}|t_N-s_N|^{\frac{1}{q}}\le(\delta+v_N)^{\frac{1}{q}}$ by (2,3). Now, given $\eta>0$, we set $\delta=\left(\frac{\eta^2}{2K_2}\right)^q$ and then use (5,11) to find an $N_0=N_0(\eta)$ such that

$$
E\left[\sup_{\substack{a\le s<t\le b\\ t-s\le\delta}}|\hat{g}\big(V_t^N\big)-\hat{g}\big(V_s^N\big)-\zeta_t^N+\zeta_s^N|\right]\le\eta^2\quad\forall N\ge N_0. \qquad (12)
$$

MCC for $\{V^N\}$ follows by Markov's inequality and Proposition 14 b (to follow).

Step 2: Establish convergence of finite dimensional distributions.

One has by independence of $\{\Xi^{j,N}\}_{j=1}^{m_N}$ and the fact $\mathcal{L}(\xi^{j,N}_{t_N})=V_{t_N}$ that

$$
\begin{aligned}
E[(\varrho(V_t^N,V_{t_N}))^2]&=\sum_{n=1}^{\infty}2^{-n}E|V_t^N(g_n)-V_{t_N}(g_n)|^2 \qquad (13)\\
&=\sum_{n=1}^{\infty}2^{-n}\frac{1}{m_N^2}\sum_{j=1}^{m_N}\left(E[g_n^2(\xi^{j,N}_{t_N})]-E[g_n(\xi^{j,N}_{t_N})]^2\right)\overset{N}{\ll}m_N^{-1}\to 0.
\end{aligned}
$$

Also, $V_{t_N}\to V_t$ (surely) for $t\in\{b\}\cup J(V)^C$ so convergence in probability,

$$
Q^V(\varrho(V_t^N,V_t)\ge\epsilon)\to 0 \qquad (14)
$$

for any $\epsilon>0$, follows from (13). Then, convergence in probability at multiple times $a\le t_1<t_2<\cdots<t_d\le b$ with each $t_i\in\{b\}\cup(J(V))^C$,

$$
Q^V(\max_i\varrho(V_{t_i}^N,V_{t_i})\ge\epsilon)\to 0\quad\forall\ \ \epsilon>0, \qquad (15)
$$

follows immediately from (14). This is equivalent to convergence of the finite dimensional distributions (fdds) of V^N to those of the deterministic V on $\mathbb{P}(E)$. Convergence of the fdds on $\overline{\mathbb{P}(E)}$ follows from Ethier and Kurtz [10, Corollary 3.3.2].

Step 3: Conclude convergence in probability in $D_{\mathbb{P}(E)}[a,b]$.
Tightness of $\left\{V^N\right\}$ in $D_{\overline{\mathbb{P}(E)}}[a,b]$ with respect to Q^V follows from Step 1 and Theorem 15 d) below. Relative compactness then follows since $D_{\overline{\mathbb{P}(E)}}[a,b]$ is Hausdorff so convergence of the fdds implies $V^N \Rightarrow V$ on $D_{\overline{\mathbb{P}(E)}}[a,b]$. However, $D_{\mathbb{P}(E)}[a,b] \in \mathcal{B}(D_{\overline{\mathbb{P}(E)}}[a,b])$ so $V^N \Rightarrow V$ on $D_{\mathbb{P}(E)}[a,b]$ by [10, Corollary 3.3.2]. Since V is deterministic, we get convergence in probability,

$$\lim_{N\to\infty} Q^V(\tilde{\rho}(V^N, V) > \epsilon) = 0 \ \forall \epsilon > 0, \tag{16}$$

where $\tilde{\rho}$ is the $D_{\mathbb{P}(E)}[a,b]$ metric defined in terms of Prohorov metric ρ in (24) below. Finally, $Q^V(\tilde{\rho}(V^N, V) > \epsilon)$ is (an expectation that is) measurable with respect to $\sigma(V)$ so

$$Q(\tilde{\rho}(V^N, V) > \epsilon) \to 0 \ \forall \epsilon > 0, \tag{17}$$

by dominated convergence, where Q is the probability measure defined on $\overline{\Omega} \times \Omega$ by

$$Q(A) = \int_{\Omega} \int_{\overline{\Omega}} 1_A(\overline{\omega}, \omega) Q^{V(\omega)}(d\overline{\omega}) P(d\omega) \ \ \forall \ A \in \overline{\mathcal{F}} \times \mathcal{F}. \ \square \tag{18}$$

We give a $[0,\infty)$, almost-sure version of the previous result.

Theorem 3 *Let E be a Polish space and $\{V_t, t \geq 0\}$ have $D_{\mathbb{P}(E)}[0,\infty)$ paths. Then, there are numbers $m_N \to \infty$ and conditionally-i.i.d. E-valued cadlag Markov processes on an enlarged probability space $\{\Xi^{1,N}, ..., \Xi^{m_N,N}\}$ such that the empirical processes $V^N \doteq \frac{1}{m_N} \sum_{i=1}^{m_N} \delta_{\Xi^{i,N}} \to V$ a.s. on $D_{\mathbb{P}(E)}[0,\infty)$.*

Proof. Let $\{s_k^N\}_{k=0}^{\infty}$ be as in (2,3) but with $a = 0$ and $b = \infty$, set $\sigma_K = s_K^1$ for $K \in \mathbb{N}$ so $\sigma_K \leq Kv_1$ and take $l(N,K) = \{k : s_k^N = \sigma_K\}$. Let $\delta_{N,K} = \min_{i \leq l(N,K)} \{s_i^N - s_{i-1}^N\}$, set $|\rho^{N,K}|_p = \left[\sum_{k=0}^{l(N,K)-1} \left(\rho_p(V^{\omega}_{s_{k+1}^N}, V_{s_k^N})\right)^p\right]^{\frac{1}{p}}$ for some $p \geq 2$ and take numbers m_N^K such that

$$P\left(\lim_{N\to\infty} \frac{|\rho^{N,K}|_p^2 \, \delta_{N,K}^{-\frac{2}{q}}}{m_N^K} \leq 1\right) = 1, \tag{19}$$

where $\frac{1}{p} + \frac{1}{q} = 1$, so $m_N \doteq \max_{K \leq N} m_N^K$ satisfies $P\left(\lim_{N\to\infty} \frac{|\rho^{N,K}|_p^2 \, \delta_{N,K}^{-\frac{2}{q}}}{m_N} \leq 1, \ \forall K \in \mathbb{N}\right) = 1$. Now, fix V so that

$$\lim_{N\to\infty} \frac{|\rho^{N,K}|_p^2 \, \delta_{N,K}^{-\frac{2}{q}}}{m_N} \leq 1, \ \forall K \in \mathbb{N}. \tag{20}$$

Next, we define $t_N = \max\{s_n^N : s_n^N \leq t\}$, $\Xi_t^{j,N} = \xi_{t_N}^{j,N}$ and $V_t^N \doteq \frac{1}{m_N} \sum_{j=1}^{m_N} \delta_{\Xi_t^{j,N}}$ for $t \geq 0$ on some probability space $(\overline{\Omega}, \overline{\mathcal{F}}, Q^V)$ as in the proof of Theorem 2. Each $\{V_t^N, \ t \geq 0\}$ and each $\{\Xi_t^{j,N}, \ t \geq 0\}$ is a cadlag $\{\mathcal{G}_t^N\}$-Markov process that is constant between jumps at $\{s_n^N\}_{n=0}^{\infty}$ for each j, N, where $\mathcal{G}_t^N \doteq \sigma\{\xi_{s_N}^{j,N} : a \leq s \leq t, j \leq m_N\}$.

Now, we let $T, \eta > 0$ and pick K so that $\sigma_K \geq T$. Using the proof of Theorem 2, we find that $\{V^N|_{[0,\sigma_K]}\}$ is tight with respect to Q^V. Therefore, by the only if part of Corollary 21 (to follow): Given $\eta > 0$, there is a compact set $K_{T,\eta} \subset \mathbb{P}(E)$ and a $\delta > 0$ satisfying

$$\inf_N Q^V(V_t^N \in K_{T,\eta} \ \text{ for } t \in [0,T]) \geq 1 - \eta \text{ and } \sup_N Q^V(w'_\rho(V^N, \delta, T) \geq \eta) \leq \eta,$$

where $w'_\rho(x, \delta, T) = \inf_{\{t_i\}} \max_i \sup_{s,t \in [t_{i-1}, t_i)} \rho(x_t, x_s)$ and $\{t_i\}$ ranges over all partitions of the form $0 = t_0 < t_1 < \cdots < t_{n-1} < T \leq t_n$ with $\min_{i \in \{1,2,\ldots,n\}}(t_i - t_{i-1}) > \delta$ and $n \in \mathbb{N}$. Therefore, $\{V^N\}$ is relatively compact on $D_{\mathbb{P}(E)}[0,\infty)$ by Ethier and Kurtz [10, Theorem 3.7.2]. Next, we get the fdd convergence exactly as Step 2 of the proof of Theorem 2. Hence, $V^N \Rightarrow V$ on $D_{\mathbb{P}(E)}[0,\infty)$. The proof is completed as in the last part of Step 3 of the proof

of Theorem 2 and then taking a subsequence so that almost sure convergence is attained. □

3 Background

The following definition is motivated by [10] p. 113.

Definition 2 *Let $(E, \mathcal{T})$ be a topological space and $\mathcal{G} \subset M(E)$. Then, i) $\mathcal{G}$ separates points (s.p.) if for $x \neq y \in E$ there is a $g \in \mathcal{G}$ with $g(x) \neq g(y)$ and ii) $\mathcal{G}$ strongly separates points (s.s.p.) if, for every $x \in E$ and neighborhood O_x of x, there is a finite collection $\left\{g^1, ..., g^k\right\} \subset \mathcal{G}$ such that $\inf_{y \notin O_x} \max_{1 \leq l \leq k} |g^l(y) - g^l(x)| > 0$.*
Acronym s.s.p. will also be used for strongly separate points *and* strongly separating points *depending upon English context.*

$\mathcal{G}$ s.s.p. means that for any x and neighborhood O_x there are $\varepsilon > 0$ and $\left\{g^1, ..., g^k\right\} \subset \mathcal{G}$ such that $\max_{1 \leq l \leq k} |g^l(y) - g^l(x)| < \varepsilon$ implies $y \in O_x$. Thus, $\mathcal{G}$ s.s.p. implies $\mathcal{G}$ s.p. (in a Hausdorff space) and defines a topology $\mathcal{T}^{\mathcal{G}}$ through the basis

$$\mathbb{B}^{\mathcal{G}} \doteq \{\{y \in E : \max_{1 \leq l \leq k} |g^l(y) - g^l(x)| < \varepsilon\} |\ g^1, ..., g^k \in \mathcal{G},\ \varepsilon > 0,\ x \in E, k \in \mathbb{N}\} \tag{21}$$

on E that is finer than the original topology. This yields the following lemma:

Lemma 4 *Let $(E, \mathcal{T})$ be a Hausdorff space, $\mathcal{G} \subset M(E)$ and $\Gamma(x) \doteq (g(x))_{g \in \mathcal{G}}$. Then, Γ has a continuous inverse $\Gamma^{-1} : \Gamma(E) \subset \mathbb{R}^{\mathcal{G}} \to E$ if and only if $\mathcal{G}$ s.s.p. Γ is an imbedding of E in $\mathbb{R}^{\mathcal{G}}$ if and only if $\mathcal{G} \subset C(E)$ and $\mathcal{G}$ s.s.p.*

Proof. If $\mathcal{G}$ s.s.p., then $\mathcal{G}$ s.p. so Γ^{-1} exists. $\mathcal{T} \subset \mathcal{T}^{\mathcal{G}}$ so Γ^{-1} is continuous. □

Given $\mathcal{G} \subset M(E)$ that may not s.s.p., we still define a topology $\mathcal{T}^{\mathcal{G}}$ (through $\mathbb{B}^{\mathcal{G}}$ or, equivalently, pseudometrics $\max_{1 \leq l \leq k} |g^l(y) - g^l(x)|$) that may differ from $\mathcal{T}$. Clearly, $\mathcal{G}$ s.s.p. on $(E, \mathcal{T}^{\mathcal{G}})$. If $\mathcal{G}$ s.p. then $(E, \mathcal{T}^{\mathcal{G}})$ is Hausdorff. In any case, if $\mathcal{G} = \{g_k\}_{k=0}^{\infty} \subset M(E)$ is countable, then

$$\rho(x,y) \doteq \sum_{k=0}^{\infty} 2^{-k} \left(|g_k(x) - g_k(y)| \wedge 1\right) \tag{22}$$

is a single pseudometric that generates $\mathcal{T}^{\mathcal{G}}$. (See [9] p. 20 for a pseudometric definition.) If, in addition, $\{g_k\}_{k=0}^{\infty}$ s.p., then (22) becomes a metric.

(22) illustrates the importance of reducing a strongly separating collection to *countability.*

Lemma 5 *If $(E, \mathcal{T})$ has a countable basis and $\mathcal{G} \subset C(E)$ s.s.p., then there is a countable collection $\{g_k\}_{k=0}^{\infty} \subset \mathcal{G}$ that s.s.p. Moreover, $\{g_k\}_{k=0}^{\infty}$ can be taken closed under either multiplication or addition if $\mathcal{G}$ is.*

Proof. See Blount and Kouritzin [4]. □

The following classical topology result is the key to our homeomorphic methods.

Theorem 6 *(Stone-Čech compactification) Suppose $(E, \mathcal{T})$ is a Hausdorff space; $\mathcal{G} \subset \overline{C}(E)$ s.s.p. and $G(x) = (g(x))_{g \in \mathcal{G}}$. Then, there exists a compact Hausdorff space $\overline{E} \supset E$ and a homeomorphism $\overline{G} : \overline{E} \to \overline{G}(\overline{E}) \subset \mathbb{R}^{\mathcal{G}}$:*

(1) $\overline{E}$'s subspace topology on E is $\mathcal{T}$;

(2) $\overline{G}(\overline{E})$ is the $\mathbb{R}^{\mathcal{G}}$ closure of $G(E)$;

(3) $\overline{G}|_E = G$;

(4) If E has a countable base and $\mathcal{G} = \{g_k\}$ is taken to be countable, then $\overline{E}$ is metrizable with a metric $\overline{d}$ that can satisfy either: $\overline{d}(x,y) = \sum_{k=1}^{\infty} 2^{-k}(|g_k(x) - g_k(y)| \wedge 1)$ or $\overline{d}(x,y) = \sqrt{\sum_{k=1}^{\infty} 2^{-k}(|g_k(x) - g_k(y)|^2 \wedge 1)}$ for $x, y \in E$;

(5) If E is a complete, separable metric space and $\mathcal{G} = \{g_k\}$ is taken to be countable, then $E, G(E)$ are Borel subsets of $\overline{E}, \overline{G}(\overline{E})$ respectively.

Remark 4 *The Stone-Čech compactification is usually defined with all continuous bounded functions i.e. $\mathcal{G} = \overline{C}(E)$. However, it is important for us to use a countable subclass $\mathcal{G} \subset$*

$\overline{C}(E)$ *at times. Naturally, we only produce continuous extensions for the functions* $\mathcal{G}$*.*

Proof. This theorem follows from Lemma 4, Lemma 5, the development on p. 239 of Munkries [22], and Corollary I.3.3 of Parthasarathy [23]. We explain the final item: By a first application of [23, Corollary I.3.3] $G(E)$ is a Borel measurable subset of $\mathbb{R}^\infty$ and hence of the compact subset $\overline{G}(\overline{E})$. A second application yields $\overline{G}^{-1}(G(E)) = E \in \mathcal{B}(\overline{E})$. □

There is an equivalent way to view s.s.p. that clarifies some developments.

Lemma 7 *Suppose* E *is Hausdorff and* $\mathcal{G} \subset M(E)$*. Then,* $\mathcal{G}$ *s.s.p. if and only if for any net* $\{x_i\}_{i\in I} \subset E$ *and point* $x \in E$*, one has* $g(x_i) \to g(x)$ *for all* $g \in \mathcal{G}$ *implies that* $x_i \to x$ *in* E*.*

Proof. See Blount and Kouritzin [4]. □

Notwithstanding the previous lemma, it is often difficult to determine if a collection of functions s.s.p. The next example from Theorems 6 and 11 (a) of Blount and Kouritzin [4], used to establish Theorem 2, shows that the s.s.p. property can sometimes be inferred.

Example 1 *Suppose* E *is a topological space and* $\mathcal{G} \subset B(E)$ *s.p., s.s.p. and is closed under multiplication. Then, the functions* $\{\phi_g(P) \doteq \int_E g dP : g \in \mathcal{G}\}$ *s.p. and s.s.p. on* $\mathbb{P}(E)$ *if either* $\mathcal{G}$ *is countable or* E *has a countable base and* $\mathcal{G} \subset \overline{C}(E)$*.*

In the sequel, we will often assume that E is a completely regular (CR) space, i.e. Hausdorff with the property that for any closed subset $F \subset E$ and any $x \notin F$, there exists $f \in C(E)$ such that $f(x) = 0$ and $f(y) = 1$ for every $y \in F$. (Note: Some authors do not include the Hausdorff property within the definition of CR but it is convenient for us to.)

There are non-Polish, CR spaces of interest. For example, Mitoma [21], Holley and Stroock [13] and others consider probability measures on cadlag spaces of tempered distributions while Fitzsimmons [12], Meyer and Zhang [20] and others consider probability measures on Lusin spaces. Any CR space E has a collection $\mathcal{D}$ of pseudometrics that determines its

topology. We write $(E, \mathcal{D})$ to denote such a framework. A Hausdorff space E is CR if and only if some $\mathcal{G} \subset C(E)$ s.s.p. Indeed, it follows from Lemma 7 that

$$\mathcal{G} \doteq \left\{ g_{x,d} \doteq \prod_{i=1}^{l} \frac{1}{2}(1 - d_i(\cdot, x_i)) \vee 0 : x_i \in E, d_i \in \mathcal{D}, l \in \mathbb{N} \right\}, \tag{23}$$

is a collection of non-negative, continuous functions bounded by $\frac{1}{2}$ that is closed under multiplication and s.s.p. When $\mathcal{D}$ is a single metric each $g_{x,d}$ has bounded support, is Lipschitz continuous and satisfies $\|g\|_{BL} \leq 1$. The following result will help to explain our assumptions in Theorem 20 to follow.

Proposition 8 *Suppose E is a topological space and there is a countable collection $\mathcal{G} = \{g_k\}_{k=0}^{\infty} \subset C(E)$ that s.p. Then; each compact $K \subset E$ is metrizable; the subspace topologies on K, $\mathcal{T}|_K$ and $\mathcal{T}^{\mathcal{G}}|_K$, are equal; and all compact subsets of $D_E[a,b]$ are metrizable as well.*

Proof. $G \doteq (g_0, g_1, ...) : E \to \mathbb{R}^{\infty}$ is 1-1 and continuous so $G : K \to G(K)$ is a closed map, hence a homeomorphism for any compact K. The remainder follows from Proposition 1.6 vii) of Jakubowski [14] (and time homeomorphism). □

We are interested in cadlag function spaces with the usual Skorokhod topology (see [3], [10], [14], [21], [26] also). For a CR space $(E, \mathcal{D})$, we define $D_E[a,b]$ for some $-\infty < a < b < \infty$ to be the space of all E-valued functions on $[a,b]$ that are right continuous and have left-hand limits (i.e. cadlag) with respect to the topology on E. The pseudometrics

$$\tilde{d}(x,y) = \inf_{\lambda \in \Lambda_{[a,b]}} \left(\operatorname*{ess\,sup}_{a<t<b} |\log \lambda'(t)| \vee \sup_{t \in [a,b]} d\left(x\left(\lambda(t)\right), y\left(t\right)\right) \right) \ \forall d \in \mathcal{D} \tag{24}$$

topologize $D_E[a,b]$ in the Skorokhod sense, where $\Lambda_{[a,b]}$ are the strictly increasing, continuous mappings of $[a,b]$ onto itself. Actually, these spaces are isometrically isomorphic to each other.

Definition 3 *Suppose $(E, \mathcal{D})$ and $(E_0, \mathcal{D}_0)$ are CR. Then, these spaces are isometrically isomorphic if there is a bijective map $\Gamma : E \to E_0$ such that $\mathcal{D} = \{d_0(\Gamma(\cdot), \Gamma(\cdot)) : d_0 \in \mathcal{D}_0\}$.*

Suppose $(E,\mathcal{D})$ is CR, $\rho : [0,1] \overset{\text{onto}}{\to} [a,b]$ is an increasing homeomorphism that is a C^1-diffeomorphism on $(0,1)$ and $\overline{\lambda} = \rho\circ\lambda\circ\rho^{-1}, \overline{x} = x\rho^{-1}, \overline{y} = y\rho^{-1}$. Then, $\Lambda_{[a,b]} = \left\{\overline{\lambda} : \lambda \in \Lambda_{[0,1]}\right\}$

$$\sup_{t\in[a,b]} d\left(\overline{x}\left(\overline{\lambda}(t)\right), \overline{y}\left(t\right)\right) = \sup_{s\in[0,1]} d\left(x\left(\lambda(s)\right), y\left(s\right)\right) \ \forall x, y \in D_E[0,1],\ d \in \mathcal{D} \text{ and}$$
$$\operatorname*{ess\,sup}_{a<t<b} |\log \overline{\lambda}'(t)| = \operatorname*{ess\,sup}_{0<s<1} |\log \lambda'(s) + \log \rho'(\lambda(s)) - \log \rho'(s)| \ \forall \lambda \in \Lambda_{[0,1]}$$

for any $\lambda \in \Lambda_{[0,1]}$. This immediately gives us the following result.

Lemma 9 *Suppose $-\infty < a < b < \infty$ and $\rho(t) = a + t(b-a)$. Then, $\overline{x} = x\rho^{-1}$ defines an isometric isomorphism between CR spaces $D_E[0,1]$ and $D_E[a,b]$.*

Each $D_E[a,b]$ is CR with pseudometrics $\widetilde{\mathcal{D}} \doteq \{\tilde{d} : d \in \mathcal{D}\}$ defining its topology and is isometrically isomorphic to $D_E[0,1]$ so results for $D_E[0,1]$ can extend to $D_E[a,b]$ via isomorphism.

The topology on $D_E[a,b]$ is independent of the choice of pseudometrics $\mathcal{D}$ generating the topology on E by (isometric isomorphism and) Jakubowski [14, Theorem 1.3]. Suppose S is another topological space and $f : E \to S$ is continuous. Then, $\widetilde{f} : D_E[a,b] \to D_S[a,b]$, defined by $\widetilde{f}(x)(t) = f(x(t))$ for $t \in [a,b]$, is also continuous. We also use $f \circ x$ to denote the function $\widetilde{f}(x)$. Let $\mathcal{C}^\delta$ be the collection of all partitions $\{t_i\}_{i=0}^n$ satisfying $a = t_0 < t_1 < \cdots < t_n = b$ with $\min_{1\le i\le n}(t_i - t_{i-1}) > \delta$ and define the moduli of continuity $w_d(x,[c,d)) \doteq \sup_{c\le s<t<d} d(x_t, x_s)$, $w'_d(x,\delta) \doteq \inf_{\{t_i\}\in\mathcal{C}^\delta} \max_i w_d(x,[t_{i-1},t_i))$, $w^{"}_d(x,\delta) \doteq \sup_{\substack{a\le t_1\le t\le t_2<b \\ t_2-t_1\le\delta}} \min\{d(x_t,x_{t_1}), d(x_{t_2},x_t)\}$ for each $d \in \mathcal{D}$. Each $x \to w'_d(x,\delta)$ is measurable by the proof of Lemma 3.6.2 in [10]. Following Billingsley [2, pp. 119-120], one finds

$$w_d(x,[a,a+\delta)) + w^{"}_d(x,\delta) + w_d(x,[b-\delta,b)) \le 3w'_d(x,\delta) \text{ and} \tag{25}$$
$$w'_d\left(x,\frac{\delta}{2}\right) \le w_d(x,[a,a+\delta)) \vee 6w^{"}_d(x,\delta) \vee w_d(x,[b-\delta,b)). \tag{26}$$

For notational ease, we write $w, w', w^{"}$ for $w_d, w'_d, w^{"}_d$ when d is absolute value or Euclidean distance. Now, following Ethier and Kurtz [10, Theorem 3.6.3], we can characterize the

compact sets of $D_E[a,b]$ when E is a complete metric space.

Theorem 10 *Let (E,d) be a complete metric space and $a < b$. Then, $A \subset D_E[a,b]$ has compact closure if and only if: a) There is a compact set $K \subset E$ such that $x(t) \in K$ for all $t \in [a,b]$, $x \in A$; and b) $\lim_{\delta \to 0} \sup_{x \in A} w'_d(x,\delta) = 0$.*

For the general CR case Jakubowski [14, Proposition 1.6 vi)] (and isomorphism) establish:

Proposition 11 *Let $(E,\mathcal{D})$ be a CR space and $a < b$. Then, for any compact $\mathcal{K} \subset D_E[a,b]$ there is a compact $K \subset E$ such that $\mathcal{K} \subset \{x \in D_E[a,b] : x(t) \in K \ \ \forall t \in [a,b]\}$.*

4 Modulus of Continuity and Containment Conditions

Prohorov's theorem states that tightness implies relative compactness for the distributions $P_\alpha \circ (X^\alpha)^{-1}$ when E is Hausdorff (see Theorem 2.2.1 of [15]). It follows from Theorem 10 and Proposition 11 above that the compact containment and the modulus of continuity conditions are important properties for establishing tightness on Skorokhod spaces.

CCC: Suppose E is a topological space. A family of processes $\{(\Omega_\alpha, \mathcal{F}_\alpha, P_\alpha), X^\alpha\}$ with $D_E[a,b]$ paths satisfies the *compact containment condition* if: For all $\varepsilon > 0$, there is a compact $K_\varepsilon \subset E$ such that $\inf_\alpha P_\alpha(X_t^\alpha \in K_\varepsilon,\, a \le t \le b) \ge 1 - \varepsilon$;

It is often a lot of work to verify CCC directly. However, on dual spaces CCC sometimes degenerates to the real-valued case, which may be relatively easy to verify.

Lemma 12 *Let $E = S'$ be the topological dual of a nuclear Fréchet space S. Then, $\{X^\alpha\}$ satisfies CCC if $\{\langle X^\alpha, \varphi\rangle\}$ does for each $\varphi \in S$.*

Proof. This result follows from the Fréchet space version of the uniform boundedness principle (see Dieudonné [8]). The details are given on pp. 263-264 of Jakubowski [14]. □

MCC with $\mathcal{D}$, AMCC with $\mathcal{D}$: Suppose E is CR, $\mathcal{D}$ is a collection of pseudometrics determining the topology on E and $\{X^\alpha\}$ is a family of processes with $D_E[a,b]$ paths. Then, $\{X^\alpha\}$ satisfies the *modulus of continuity condition with* $\mathcal{D}$ if: for every $\eta > 0$ and $d \in \mathcal{D}$, there is a $\delta = \delta_{d,\eta} > 0$ satisfying $\sup_\alpha P_\alpha(w'_d(X^\alpha, \delta) \geq \eta) \leq \eta$, *and* $\{X^\alpha\}$ satisfies the *asymptotic modulus of continuity condition with* $\mathcal{D}$ if: for every $\eta > 0$ and $d \in \mathcal{D}$, there is a $\delta = \delta_{d,\eta} > 0$ and a finite set $A = A_{d,\eta}$ satisfying $\sup_{\alpha \notin A} P_\alpha(w'_d(X^\alpha, \delta) \geq \eta) \leq \eta$.

It can be easier to verify AMCC than MCC. It follows from (25,26) that we can replace $w'_d(X^\alpha, \delta)$ above with $w_d(X^\alpha, [a, +\delta)) + w''_d(X^\alpha, \delta) + w_d(X^\alpha, [b-\delta, b))$. In any event, we often just need MCC to be true for *some* collection of pseudometrics.

MCC (AMCC): A family $\{X^\alpha\}$ with $D_E[a,b]$ paths satisfies the *(asymptotic) modulus of continuity condition* if it satisfies MCC (AMCC) with some topology-determining $\mathcal{D}$.

The following tightness result for Polish spaces is included for later reference.

Theorem 13 *Suppose (E,d) is Polish and $\{Y^\alpha\}$ are cadlag processes that satisfy (CCC) and (AMCC with d). Then, $\{Y^\alpha\}$ is a tight collection of $D_E[a,b]$-valued random variables.*

Proof. Given $\eta > 0$, $n \in \mathbb{N}$, there are $\delta_{n,\eta} > 0$, finite set $A_{n,\eta}$ and compact $K_{n,\eta}$ such that

$$\sup_{\alpha \notin A} P_\alpha(w'_d(Y^\alpha, \delta_{n,\eta}) \geq \eta 2^{-n}) \vee P_\alpha(Y^\alpha_t \notin K_{n,\eta} \text{ for some } t \in [a,b]) \leq \eta 2^{-n-1}. \tag{27}$$

However, using Ulam's theorem, recalling Theorem 10 and skrinking $\delta_{n,\eta} > 0$ if necessary, we can take $A = \emptyset$. Next, letting $\mathcal{K}_\eta = \text{closure}\Big(\bigcap_{n=1}^{\infty} F_{n,\eta}\Big)$, where

$$F_{n,\eta} \doteq \{y \in D_E[a,b] : w'_d(y, \delta_{n,\eta}) < \eta 2^{-n},\ y_t \in K_{n,\eta}\ \forall t \in [a,b]\},$$

one finds by (27) that $P_\alpha(Y^\alpha \notin \mathcal{K}_\eta) \leq \sum_{n=1}^\infty P_\alpha(Y^\alpha \notin F_{n,\eta}) \leq \eta \sum_{n=1}^\infty 2^{-n} = \eta$. Moreover, $\mathcal{K}_\eta$ is compact by Theorem 10 since $0 \leq \sup_{y \in F_{n,\eta}} w'_r(y, \delta_{n,\eta}) \leq \eta 2^{-n}$ for all n and monotonicity imply that $\lim_{\delta \searrow 0} \sup_{y \in \bigcap_{n=1}^\infty F_n} w'_r(y, \delta) = 0$. $\square$

MCC is equivalent to three other conditions that are often easier to verify.

Proposition 14 *Let E be CR and $\{X^\alpha\}$ have $D_E[a,b]$ paths. Then, the following are equivalent to* ***a)*** *$\{X^\alpha\}$ satisfies MCC:*

b) *There is a class $\mathcal{J} \subset \overline{C}(E)$ that s.s.p. such that for each $g \in \mathcal{J}' \triangleq \mathcal{J} \cup \{g_1 + g_2 : g_1, g_2 \in \mathcal{J}\}$ and $\eta > 0$ there are a $\delta > 0$, finite set $\mathcal{A} = \mathcal{A}_{\eta,g}$ and cadlag processes $\{\zeta^\alpha = \zeta^\alpha_{\eta,g}\}$ satisfying $\sup_{\alpha \notin \mathcal{A}} P_\alpha(w'_d(\zeta^\alpha, \delta) \geq \eta) \leq \eta$ and $\sup_{\alpha \notin \mathcal{A}} P_\alpha(\sup_{\substack{a \leq s < t \leq b \\ t-s \leq \delta}} |g(X^\alpha_t) - g(X^\alpha_s) - \zeta^\alpha_t + \zeta^\alpha_s| \geq \eta) \leq \eta$;*

c) *There is a class $\mathcal{G} \subset C(E)$ that s.s.p. and satisfies: i) for each $g_0, ..., g_k \in \mathcal{G}$, $k \in \mathbb{N}_0$, $(g_0, ..., g_k) \circ X^\alpha$ satisfies AMCC in $\mathbb{R}^{k+1}$ with $|\cdot|$, and ii) for any $g \in \mathcal{G}$ and $\eta > 0$ there is a $K = K_{\eta,g} > 0$ such that $\inf_\alpha P_\alpha(\sup_{a \leq t \leq b} |g(X^\alpha_t)| \leq K) > 1 - \eta$; and*

d) *There is a class $\mathcal{G} \subset C(E)$ that s.s.p. and satisfies: i) for each $g_0, ..., g_k \in \mathcal{G}$, $k \in \mathbb{N}_0$, and $a_0, ..., a_k \in \mathbb{R}$, $\sum_{j=0}^k a_j g_j(X^\alpha)$ satisfies AMCC in $\mathbb{R}$ with $|\cdot|$, and ii) for any $g \in \mathcal{G}$ and $\eta > 0$ there is a $K = K_{\eta,g} > 0$ such that $\inf_\alpha P_\alpha(\sup_{a \leq t \leq b} |g(X^\alpha_t)| \leq K) > 1 - \eta$.*

Remark 5 *1) While b-d appear weaker than MCC, they are actually equivalent. 2) It follows from the proof below that if any of a-d are true, then all parts are true with the finite sets $A = \mathcal{A} = \emptyset$ and b) is true with $\mathcal{J}$ being closed under addition. 3) Parts c) and d) are particularly useful when each $\{g \circ X^\alpha\}$ satisfies CCC. 4) One might think that the last conditions in parts c) and d) are some weak compact containment condition that does not appear in a) or b). However, these conditions just ensure that the choice of $\mathcal{G}$ is good. They are not restrictive on their own as one could choose $\mathcal{G}$ so they were true if it were not for the other constraints.*

Proof of Proposition 14. Assuming a) holds, we let $\mathcal{G} \subset \{d(e, \cdot) \wedge 1 : e \in E, d \in \mathcal{D}\}$ s.s.p. Then, for every $\eta \in (0,1)$, one finds by the triangle inequality that

$$\sup_\alpha P_\alpha(\inf_{\{t_i\}} \max_i \sup_{s,t\in[t_{i-1},t_i)} |(g_0(X_t^\alpha),...,g_k(X_t^\alpha)) - (g_0(X_s^\alpha),...,g_k(X_s^\alpha))| \geq \eta)$$
$$\leq \sup_\alpha P_\alpha(\inf_{\{t_i\}} \max_i \sup_{s,t\in[t_{i-1},t_i)} \sqrt{k+1} \max_{0\leq l\leq k} d_l(X_t^\alpha, X_s^\alpha) \geq \eta) \leq \eta\sqrt{k+1}$$

for all $g_l(\cdot) = d_l(e_l,\cdot) \in \mathcal{G}$ and c) holds with $A_\eta = \emptyset$. (The last equation in c) holds trivially since $\mathcal{G}$ is uniformly bounded by 1.) Moreover, assuming c) holds and letting $\mathcal{D} = \{d : d(x,y) = |(g_0(x),...,g_k(x)) - (g_0(y),...,g_k(y))|, k \in \mathbb{N}_0 \text{ with } g_0,...,g_k \in \mathcal{G}\}$, one finds that $\mathcal{D}$ is a collection of pseudometrics determining the topology on E. Now, $(g_0(X^\alpha),...,g_k(X^\alpha))$ satisfies AMCC holds by i) and CCC by ii) so a) holds by Theorems 13 and 10 with $E = \mathbb{R}^{k+1}$.

When b) holds, $g \in \mathcal{J}'$, and $\eta > 0$, one finds

$$|g(X_t^\alpha) - g(X_s^\alpha)| \leq \sup_{\substack{a\leq\sigma<\tau\leq b\\ \tau-\sigma\leq\delta}} |g(X_\tau^\alpha) - g(X_\sigma^\alpha) - \zeta_\tau^\alpha + \zeta_\sigma^\alpha| + |\zeta_t^\alpha - \zeta_s^\alpha| \tag{28}$$

for $t - s \leq \delta$ and some $\{\zeta^\alpha\}$ satisfying AMCC. It follows that there is a $\delta > 0$ such that

$$\begin{aligned}
&\sup_{\alpha\notin\mathcal{A}} P_\alpha(w(g(X^\alpha),[a,a+\delta)) + w''(g(X^\alpha),\delta) + w(g(X^\alpha),[b-\delta,b)) \geq 2\eta) \qquad (29)\\
\leq &\sup_{\alpha\notin\mathcal{A}} P_\alpha(w(\zeta^\alpha,[a,a+\delta)) + w''(\zeta^\alpha,\delta) + w(\zeta^\alpha,[b-\delta,b)) \geq \eta)\\
+ &\sup_{\alpha\notin\mathcal{A}} P_\alpha\left(3 \sup_{\substack{a\leq\sigma<\tau\leq b\\ \tau-\sigma\leq\delta}} |g(X_\tau^\alpha) - g(X_\sigma^\alpha) - \zeta_\tau^\alpha + \zeta_\sigma^\alpha| \geq \eta\right)\\
\leq &\, 2\eta
\end{aligned}$$

so $\{g\circ X^\alpha\}$ (is bounded and) satisfies AMCC and hence is tight in $D_\mathbb{R}[a,b]$ for each $g \in \mathcal{J}'$ by Theorem 13. c) holds with $\mathcal{G} = \mathcal{J}$ by Corollary 23 of the Appendix (with $S = \mathbb{R}^{k+1}$, $Y^\alpha = (g_0(X^\alpha),...,g_k(X^\alpha))$ for $g_0,...,g_k \in \mathcal{J}$ and $\mathcal{G} = \{\pi_0,\pi_1,...,\pi_k\}$ being the projection functions) and Theorem 10. When c) holds, we let $\mathcal{J} \triangleq \{g^K : g \in \mathcal{G}, K > 0\}$, where $g^K(x) \triangleq -K \vee g(x) \wedge K$, to find that $\mathcal{J} \subset \overline{C}(E)$ s.s.p. and $(g_0^K(X^\alpha),...,g_k^K(X^\alpha))$ satisfies AMCC for $g_0^K,...,g_k^K \in \mathcal{J}$. Hence, $g_i^K(X^\alpha) + g_j^K(X^\alpha)$ satisfies AMCC with $|\cdot|$ by Cauchy-Schwarz so b) follows by taking ζ^α as $g_i^K(X^\alpha)$ or $g_i^K(X^\alpha) + g_j^K(X^\alpha)$. c) to d) follows from the Cauchy-Schwarz-based bound

$$P_\alpha\left(\inf_{\{t_i\}}\max_i \sup_{s,t\in[t_{i-1},t_i)}\left|\sum_{j=0}^k a_j g_j(X_t^\alpha)-\sum_{j=0}^k a_j g_j(X_s^\alpha)\right|\geq\eta\right)$$
$$\leq P_\alpha\left(\inf_{\{t_i\}}\max_i \sup_{s,t\in[t_{i-1},t_i)}|(g_0,...,g_k)(X_t^\alpha)-(g_0,...,g_k)(X_s^\alpha)|\geq\frac{\eta}{\sqrt{\sum_{j=0}^k a_j^2}}\right).$$

Finally, it follows from (A.30) of Holley and Stroock [13] that there are (finitely many) $\theta^1,...,\theta^m\in\mathbb{R}^{k+1}$ with each $|\theta^i|=1$ such that $|v|\leq 2\max_{i=1,...,m}|\theta^i\cdot v|$ for all $v\in\mathbb{R}^{k+1}$ and so d) implies c) through the bound

$$P_\alpha(w((g_0,...,g_k)\circ X^\alpha,[a,a+\delta))+w"((g_0,...,g_k)\circ X^\alpha,\delta)+w((g_0,...,g_k)\circ X^\alpha,[b-\delta,b))\geq\eta)$$
$$\leq\sum_{i=1}^m P_\alpha\left(w\left(\sum_{j=0}^k\theta_j^i g_j\circ X^\alpha,[a,a+\delta)\right)+w"\left(\sum_{j=0}^k\theta_j^i g_j\circ X^\alpha,\delta\right)+w\left(\sum_{j=0}^k\theta_j^i g_j\circ X^\alpha,[b-\delta,b)\right)\geq\frac{\eta}{2}\right).\ \square$$

5 Tightness and Containment assuming MCE

Proposition 14 provides equivalent conditions that can be verified in lieu of MCC.

MCE: Let E be CR. Then, cadlag $\{X^\alpha\}$ satisfies the *modulus of continuity equivalence condition* if it satisfies one (hence all) of a-d in Proposition 14.

When E has a countable basis MCE has further equivalences related to tightness:

Theorem 15 *Let E be CR with a countable basis and $\{X^\alpha\}$ be E-valued processes with cadlag paths. Then, the following are equivalent to* ***a)*** *$\{X^\alpha\}$ satisfies MCE:*

b) *There is a collection $\{g_k\}_{k=0}^\infty\subset\overline{C}(E)$ that s.s.p. such that the metric*

$$d(x,y)=\sum_{k=0}^\infty 2^{-k}\left(|g_k(x)-g_k(y)|\wedge 1\right)\quad\forall\ \ x,y\in E$$

generates the topology on E and $\{X^\alpha\}$ satisfies MCC with d;

c) *There is an imbedding $G=(g_0,g_1,....):E\to\mathbb{R}^\infty$ such that $\{g_k\}\subset\overline{C}(E)$ and $\left\{\widetilde{G}(X^\alpha)\right\}$ is tight in $D_{\mathbb{R}^\infty}[a,b]$; and*

d) *$\{X^\alpha\}$ is tight in $D_{\overline{E}}[a,b]$, where $(\overline{E},\overline{d})$ is the Stone-Čech metric-space compactification of E so $\overline{d}$ generates the original topology on E.*

Proof. When a) holds, we use Remark 5, Proposition 14 c), bounding $g \to g^K$ (as in the proof of Proposition 14) and Lemma 5 to find $\{g_k\}_{k=0}^{\infty} \subset \overline{C}(E)$ that s.s.p. and $(g_0, ..., g_k) \circ X^\alpha$ satisfies MCC with $|\cdot|$ for *each* k. b) follows by noting

$$P_\alpha\left(w'_d(X^\alpha,\delta) \geq 2\eta\right) \leq P_\alpha\left(\inf_{\{t_i\}} \max_i \sup_{s,t\in[t_{i-1},t_i)} \sum_{k=0}^{K_\eta} 2^{-k}|g_k(X_t^\alpha) - g_k(X_s^\alpha)| \geq \eta\right)$$

for $K_\eta = \lceil -\log_2 \eta \rceil$ and then using Cauchy-Schwarz. Next, assuming b) holds, we define the imbedding $G : E \to \mathbb{R}^\infty$ by $G(x) = (g_0(x), g_1(x), ...)$ and find $Y^\alpha = \widetilde{G}(X^\alpha)$ satisfies MCC as well as CCC since each g_k is bounded. Hence, c) holds by Theorem 13. Assuming c), using Theorem 6 and noting that $D_{\overline{G(E)}}[a,b]$ is closed in $D_{\mathbb{R}^\infty}[a,b]$, one finds that d) follows by Lemma 24 below (with $S = D_{\overline{E}}[a,b]$, $\widehat{S} = D_{\mathbb{R}^\infty}[a,b]$, $F_n = S$ for all n and G defined by $G(x)(t) = \Gamma(x(t))$ for $x \in S$).

Given $(\overline{E},\overline{d})$ such that $\{X^\alpha\}$ is tight in $D_{\overline{E}}[a,b]$ and $\overline{d}$ induces the original topology on E, we let $\eta > 0$ be arbitrary, set $d = \overline{d}\big|_E$, and find a $\delta > 0$ so that $\sup_\alpha P_\alpha(w'_d(X^\alpha,\delta) \geq \eta) \leq \eta$ by Theorem 10 so a) follows. □

Theorem 15 generalizes an approach of [1]. They used the imbedding $G : E \to \mathbb{R}^\infty$, and assumed conditions implying those in Proposition 14 b). If we can show MCE and have a countable basis like on a separable metric space, then $\{X^\alpha\}$ is tight in $D_{\overline{E}}[a,b]$. Still, the goal is usually to show tightness on $D_E[a,b]$, which is attained by assuming CCC in addition.

Corollary 16 *Let E be CR with a countable basis, $\{X^\alpha\}$ have cadlag paths and MCE holds. Then, $\{X^\alpha\}$ is tight on $D_E[a,b]$ if and only if CCC holds.*

Proof. Let $\overline{E}$ be the Stone-Čech metric space compactification and $\tilde{I} : D_E[a,b] \to D_{\overline{E}}[a,b]$ be the identity imbedding. By CCC there are compact $K_n \subset E$ and closed $F_n = \{x \in$

$D_E[a,b] : x(t) \in K_n, a \le t \le b\}$ such that $\inf_\alpha P(X^\alpha \in F_n) > 1 - \frac{1}{n}$. $\{K_n\}$ are compact in $\overline{E}$, $\{F_n\}$ are closed in $D_{\overline{E}}[a,b]$ and $P(X^\alpha)^{-1}(\tilde{I})^{-1}$ is tight by Theorem 15. Hence, $\{X^\alpha\}$ is tight by Lemma 24 below. Conversely, CCC holds when $\{X^\alpha\}$ is tight by Proposition 11. □

Pointwise containment (c.f. p. 128 of [10]) can often be used with MCC to establish CCC.

PCP: Let (E,d) be a metric space. A family of E-valued processes $\{X^\alpha\}$ satisfies the *pointwise containment property* if: For all $\eta > 0$ and $t \in [a,b]$, there is a compact $K_{\eta,t} \subset E$ such that $\inf_\alpha P_\alpha(X_t^\alpha \in K_{\eta,t}^\eta) \ge 1 - \eta$.

If each X^α has cadlag paths, then $\inf_\alpha P_\alpha(X_t^\alpha \in K_{\eta,t}^\eta) \ge 1 - \eta$ for t in a dense subset of $[a,b]$ that includes $\{b\}$ implies PCP.

PCP is related to the pointwise tight condition (used in e.g. [1]):

PTC: Suppose E is a topological space. A family of E-valued processes $\{X^\alpha\}$ satisfies the *pointwise tight condition* if: For all $\eta > 0$ and $a \le t \le b$, there is a compact $K_{\eta,t} \subset E$ such that $\inf_\alpha P_\alpha(X_t^\alpha \in K_{\eta,t}) \ge 1 - \eta$.

Clearly, PTC implies PCP (on a metric space). However, the proof of Theorem 17 (to follow) also establishes that PCP implies PTC *if* the metric space is complete. We use PCP for generality since we do not always assume completeness.

The combination MCC, PCP holding in the same complete metric implies CCC.

Theorem 17 *Let (E,d) be a metric space, $\{X^\alpha\}$ have $D_E[a,b]$ paths and MCC, PCP both hold with metric d. Then, i) for any $\eta > 0$ there is a closed totally bounded set B_η such that*

$$\inf_\alpha P_\alpha(X_t^\alpha \in B_\eta \text{ for all } t \in [a,b]) \ge 1 - \eta$$

and ii) CCC holds when (E,d) is complete.

Notice that we required MCC in the same metric as PCP rather than just MCE. This fact is what makes CCC often difficult to establish directly and is the reason we avoided showing it in our application. For clarity, we include the following simple example.

Example 2 *Suppose* $E = \mathbb{R}$, ξ *is uniformly distributed on* $[0,1]$ *and* $Y_t^n \doteq \begin{cases} \frac{1}{\xi - t} \wedge n & 0 \leq t < \xi \\ 0 & t \geq \xi \end{cases}$.

Then, $\{Y^n\}$ *does not satisfy CCC nor MCC with Euclidean distance. However, letting* $\mathcal{J} = C_c(\mathbb{R})$*, the continuous functions with compact support, and* $\zeta_t^n = g(Y_t^n)$*, we can verify Proposition 14 b) so* $\{Y^n\}$ *satisfies MCC (with some collection of pseudometrics). Next, by Theorem 15 there is a countable collection* $\{g_k\}$ *that s.s.p. and a new metric* d *(as in that theorem) that generates the Euclidean topology and for which* $\{Y^n\}$ *satisfies MCC. Moreover, it is relatively easy to see that* $\{Y^n\}$ *satisfies PCP with either metric. Therefore by Theorem 17, for any* $\eta > 0$ *there is a closed, totally bounded* $B_\eta \subset (E, d)$ *such that*

$$\inf_n P_n(Y_t^n \in B_\eta \text{ for all } 0 \leq t \leq 1) \geq 1 - \eta.$$

However, (E, d) *is not complete and CCC does not hold.*

Proof of Theorem 17. By isometry, we can just consider the case $a = 0, b = 1$. i) Using MCC and letting $n \in \mathbb{N}$, we take δ so that

$$\sup_\alpha P_\alpha(\inf_{\{t_i\} \in \mathcal{C}^\delta} \max_i \sup_{s,t \in [t_{i-1}, t_i)} d(X_t^\alpha, X_s^\alpha) \geq \eta 2^{-n-1}) \leq \eta 2^{-n-1}. \tag{30}$$

Next, we let $\frac{1}{m} < \delta$, set $\tau_i = \frac{i}{m}$ for $i = 0, 1, ..., m$ and use PCP to find $\{\Gamma_{\eta 2^{-n-1}, \tau_i}\}$ such that

$$\inf_\alpha P_\alpha(X_{\tau_i}^\alpha \in \Gamma_{\eta 2^{-n-1}, \tau_i}^{\eta 2^{-n-1}} \text{ for } i = 0, 1, ..., m) \geq 1 - \eta 2^{-n-1}. \tag{31}$$

Using (30,31) and letting $\Gamma_{\eta,n} = \bigcup_{i=0}^m \Gamma_{\eta 2^{-n-1}, \tau_i}$, one finds that

$$\inf_\alpha P_\alpha(X_t^\alpha \in \Gamma_{\eta,n}^{\eta 2^{-n}} \text{ for all } 0 \leq t \leq 1) \geq 1 - \eta 2^{-n}.$$

Therefore, letting B_η =closure($\bigcap_{n=1}^{\infty} \Gamma_{\eta,n}^{\eta 2^{-n}}$), we have that

$$\inf_\alpha P_\alpha(X_t^\alpha \in B_\eta \text{ for all } 0 \le t \le 1) \ge 1-\eta. \quad \square$$

Corollary 18 *Let (E,d) be a metric space, $\{X^\alpha\}$ have $D_E[a,b]$-paths and MCC, PCP hold with d. Then, there is a countable union of totally bounded sets $E_0 \subset E$ and modifications $\{\widehat{X}^\alpha\}$ of $\{X^\alpha\}$ that are $D_E[a,b]$-valued random variables taking values in the (separable) subset $D_{E_0}[a,b]$.*

Remark 6 *The above Corollary could be thought of as performing two tasks: 1) It establishes that there is a modification whose paths are in this separable subset E_0 and 2) This modified process can be taken to be a $D_E[a,b]$-valued random variable, which is a measurability question (see Ethier and Kurtz [10, Proposition 3.7.1] for example). This second step does not follow a priori from [10] because we have not assumed separability of E.*

Proof. Let $E_0 = \bigcup_{i=1}^{\infty} B_{1/i}$ and $S_0 = \bigcup_{i=1}^{\infty} \left\{x \in D_E[a,b] : x_t \in B_{1/i}, a \le t \le b\right\}$, where $\{B_\eta\}$ are the closed, totally bounded sets from Theorem 17 i) so $P_\alpha(X^\alpha \in S_0) = 1$ and $S_0 \subset D_{E_0}[a,b]$. Now, we fix $e_0 \in E_0$ and take $\widehat{X}^\alpha = X^\alpha$ on $\{X^\alpha \in S_0\}$ and $\widehat{X}_t^\alpha \equiv e_0$ on $\{X^\alpha \in S_0\}^c$. $\square$

6 Tightness and Modulus of Continuity assuming CCC

In this section, we investigate the question: What conditions in addition to CCC yield pathspace tightness? If CCC is true, then we need only verify the conditional form of AMCC.

CMCC: A family of processes $\{X^\alpha\}$ with $D_E[a,b]$-valued paths satisfies the *conditional modulus of continuity condition* if there is a collection of pseudometrics $\mathcal{D}$ on E determining the relative topology on compacts such that: For every $\eta \in (0, \frac{1}{2}]$, $d \in \mathcal{D}$ and compact K satisfying $\inf_\alpha P_\alpha(X_t^\alpha \in K \ \ \forall a \le t \le b) \ge 1-\eta$, there is a $\delta = \delta_{d,\eta,K} > 0$ such that $\sup_\alpha P_\alpha(w'_d(X^\alpha,\delta) \ge \eta | X_t^\alpha \in K \ \ \forall a \le t \le b) \le \eta$.

Actually, there is a convenient condition that implies CMCC:

WMCC: Cadlag family $\{X^\alpha\}$ satisfies the *weak modulus of continuity condition* if there is $\mathcal{G} \subset C(E)$ that s.p. on compacts and $\{g \circ X^\alpha\}$ satisfies AMCC with $|\cdot|$ for all $g \in \mathcal{G}' \triangleq \mathcal{G} \cup \{f + h : f, h \in \mathcal{G}\}$;

WMCC is clearly weaker than the analogous condition in Theorems 3.1, 4.6 of Jakubowski [14], which proved useful in Dawson [6, Section 3.7] and Perkins [24, Section II.4] for showing tightness of measure-valued processes. MCC implies WMCC through Proposition 14 d) and WMCC implies MCC if $\mathcal{G} \subset \overline{C}(E)$ s.s.p. on E through Proposition 14 b).

Motivated by Proposition 14, we have important equivalences to CMCC.

Proposition 19 *Let E be CR and $\{X^\alpha\}$ have $D_E[a,b]$-valued paths. Then, the following are equivalent to* ***a)*** *$\{X^\alpha\}$ satisfies CMCC:*

b) *There is a class $\mathcal{J} \subset C(E)$ that s.p. on compacts such that for each $\eta \in (0, \frac{1}{2}]$, $g \in \mathcal{J}' \triangleq \mathcal{J} \cup \{g_1 + g_2 : g_1, g_2 \in \mathcal{J}\}$ and compact K satisfying $\inf_\alpha P_\alpha(X_t^\alpha \in K \ \ \forall a \le t \le b) \ge 1 - \eta$, there are $\delta = \delta_{\eta,g,K} > 0$, finite set $\mathcal{A} = \mathcal{A}_{\eta,g,K}$ and cadlag processes $\{\zeta^\alpha = \zeta^\alpha_{\eta,g,K}\}$ satisfying i) $\sup_{\alpha \notin \mathcal{A}} P_\alpha(w'(\zeta^\alpha, \delta) \ge \eta | X_t^\alpha \in K \ \ \forall a \le t \le b) \le \eta$ and*

$$ii) \sup_{\alpha \notin \mathcal{A}} P_\alpha(\sup_{\substack{a \le s < t \le b \\ t - s \le \delta}} |g(X_t^\alpha) - g(X_s^\alpha) - \zeta_t^\alpha + \zeta_s^\alpha| \ge \eta | X_t^\alpha \in K \ \ \forall a \le t \le b) \le \eta;$$

c) *There is a class $\mathcal{J} \subset C(E)$ that s.p. on compacts such that for each $\eta \in (0, \frac{1}{2}]$, $g_0, ..., g_k \in \mathcal{J}$ and compact K satisfying $\inf_\alpha P_\alpha(X_t^\alpha \in K \ \ \forall a \le t \le b) \ge 1 - \eta$, one has $\sup_{\alpha \notin \mathcal{A}} P_\alpha(w'((g_0, ..., g_k) \circ X^\alpha, \delta) \ge \eta | X_t^\alpha \in K \ \ \forall a \le t \le b) \le \eta$ for some $\delta = \delta_{\eta,g,K} > 0$ and finite set $\mathcal{A} = \mathcal{A}_{\eta,g,K}$.*

d) *There is a class $\mathcal{J} \subset C(E)$ that s.p. on compacts such that for each $\eta \in (0, \frac{1}{2}]$, $g_0, ..., g_k \in \mathcal{J}$, $a_0, ..., a_k \in \mathbb{R}$ and compact K satisfying $\inf_\alpha P_\alpha(X_t^\alpha \in K \ \ \forall a \le t \le b) \ge 1 - \eta$, one has $\sup_{\alpha \notin \mathcal{A}} P_\alpha(w'(\sum_{j=0}^k a_j g_j \circ X^\alpha, \delta) \ge \eta | X_t^\alpha \in K \ \ \forall a \le t \le b) \le \eta$ for some $\delta = \delta_{\eta,g,K} > 0$ and finite set $\mathcal{A} = \mathcal{A}_{\eta,g,K}$.*

Moreover, WMCC implies a-d.

Proof. Let K satisfy $\inf_\alpha P_\alpha(X_t^\alpha \in K \ \ \forall a \le t \le b) \ge 1-\eta$. Then, a-d) are reduced to a-d) of Proposition 14 with $E = K$ and conditional probability in place of P_α. The only difference is that the pseudometrics and functions are defined on all E instead of just K. Hence, the equivalence of a-d) follows from the proof of Proposition 14 with the following comments: In showing a) is equivalent to c), we take $\mathcal{G} = \{d(e,\cdot) \wedge 1 : e \in E, d \in \mathcal{D}\} \subset \overline{C}(E)$ and note that the definition of $\mathcal{D}$ in terms of $\mathcal{G}$ does not depend upon K either. In the proof of b) implies c), we only need $\{g \circ X^\alpha\}$ satisfy CCC and, in the other direction, we just let $\mathcal{J} = \mathcal{G}$ since the functions need not be bounded. WMCC implies b) by letting $\mathcal{J} = \mathcal{G}$ and $\zeta_t^\alpha = g(X_t^\alpha)$. □

By Proposition 19, one need only verify WMCC or one of the conditions equivalent to CMCC.

CMCE: Let E be CR and $\{X^\alpha\}$ be cadlag. Then, $\{X^\alpha\}$ satisfies the *conditional modulus of continuity equivalence condition* if it satisfies one (hence all) of a-d in Proposition 19.

Remark 7 *Suppose CCC and CMCE are both true, $\eta > 0$, K satisfies $\inf_\alpha P_\alpha(X_t^\alpha \in K \ \ \forall a \le t \le b) \ge 1-\eta$, and $\mathcal{J} \subset C(E)$, $\{\zeta^\alpha\}$ are as in Proposition 19 b). Then, for each $g \in \mathcal{J} \cup \{g_1 + g_2 : g_1, g_2 \in \mathcal{J}\}$ there is a $\delta > 0$ and finite set $\mathcal{A}$ such that*

$$
\begin{aligned}
&\sup_{\alpha \notin \mathcal{A}} P_\alpha(w(g(X^\alpha),[a,a+\delta)) + w''(g(X^\alpha),\delta) + w(g(X^\alpha),[b-\delta,b)) \ge 2\eta) \qquad (32)\\
\le &\sup_{\alpha \notin \mathcal{A}} P_\alpha(w(\zeta^\alpha,[a,a+\delta)) + w''(\zeta^\alpha,\delta) + w(\zeta^\alpha,[b-\delta,b)) \ge \eta) \ge \eta | X_t \in K \ \forall a \le t \le b)\\
&P_\alpha(X_t^\alpha \in K \ \forall a \le t \le b)\\
+ &\sup_{\alpha \notin \mathcal{A}} P_\alpha(\sup_{\substack{a \le s < t \le b \\ t-s \le \delta}} 3|g(X_t^\alpha) - g(X_s^\alpha) - \zeta_t^\alpha + \zeta_s^\alpha| \ge \eta | X_t^\alpha \in K \ \ \forall a \le t \le b) P_\alpha(X_t^\alpha \in K \ \forall a \le t \le b)\\
+ &\sup_\alpha P_\alpha(X_t^\alpha \notin K \text{ for some } t \in [a,b]).
\end{aligned}
$$

Therefore, it follows from (25,26) that there is a $\delta > 0$ such that

$$\sup_{\alpha \mathcal{A}} P_\alpha(w'(g(X^\alpha),\delta) \ge 2\eta) \le 3\eta. \qquad (33)$$

It follows that $\{g \circ X^\alpha\}$ is tight by Ulam's theorem and CCC.

There is another modulus-of-continuity-type condition in use.

MMCC: Cadlag family $\{X^\alpha\}$ satisfies the *mild modulus of continuity condition* if there is $\mathcal{H} \subset M(E)$ whose uniform-convergence-on-compacts closure contains $\overline{C}(E)$ and $\{h \circ X^\alpha\}$ satisfies AMCC for all $h \in \mathcal{H}$.

MMCC with $\mathcal{H} \subset \overline{C}(E)$ was used in Kurtz [18, p. 628–629] to show $D_E[0,\infty)$ tightness in locally compact Polish spaces. For general Polish spaces, it appears in [10, Theorem 3.9.1].

The following result extends Theorem 3.1 i) of [14] while simplifying its proof.

Theorem 20 *Suppose E is CR with metrizable compacts, and $\{X^\alpha\}$ have $D_E[a,b]$-valued paths. Then, the following are equivalent:*

i) *$\{X^\alpha\}$ satisfies CCC and MCE;*

ii) *$\{X^\alpha\}$ satisfies CCC and WMCC;*

iii) *$\{X^\alpha\}$ satisfies CCC and MMCC;*

iv) *$\{X^\alpha\}$ satisfies CCC and CMCE; and*

v) *Each X^α is indistinguishable from a $D_E[a,b]$-valued random variable $\widehat{X}^\alpha$ such that $\left\{\widehat{X}^\alpha\right\}$ is tight on $D_E[a,b]$.*

Proof. MMCC implies Proposition 14 b) hence MCE under CCC by letting $\mathcal{J} = \overline{C}(E)$ and $\zeta_t^\alpha = h(X_t^\alpha)$. Moreover, v) implies CCC and that $\left\{h \circ \widehat{X}^\alpha\right\}$ is tight so it satisfies AMCC for any $h \in \overline{C}(E)$ and MMCC is true. Therefore, we only need link the other four conditions.

We first show iv) implies v). By CCC there exists compact K_n such that $\inf_\alpha P_\alpha(X_t^\alpha \in K_n,\ a \leq t \leq b) \geq 1 - \frac{1}{n}$ for each n and by Remark 7 as well as Lemma 5 there is a countable collection $\mathcal{J} \subset C(E)$ that s.s.p. on each K_n and $\{g_k \circ X^\alpha\}$ is tight for each k, where $\{g_k\}_{k=1}^\infty \triangleq \mathcal{J} \cup \{f + g : f, g \in \mathcal{J}\}$. Next, we define $F_n \triangleq D_{K_n}[a,b]$ (which is closed in $D_E[a,b]$), $S_0 = \bigcup_{n=1}^\infty F_n$, the continuous maps $G \triangleq (g_1, g_2, ...) : E \to \mathbb{R}^\infty$ (so $\widetilde{G} : D_E[a,b] \to$

$D_{\mathbb{R}^\infty}[a,b]$) and $\widehat{G} : D_E[a,b] \to (D_{\mathbb{R}}[a,b])^\infty$ by $\widehat{G}(x) = (\widetilde{g}(x))_{g\in\{g_k\}_{k=1}^\infty}$. Then, $\widehat{G}(X^\alpha)$ is tight in $(D_{\mathbb{R}}[a,b])^\infty$ since each $\widetilde{g}(X^\alpha)$ is tight and by Theorem 22 (with $S = K_n$) of the Appendix $\widehat{G} : F_n \to \widehat{G}(F_n)$ is a homeomorphism for each $n = 1, 2, ...$ Now, $\widehat{G}(F_n) = \{y \in (D_{\mathbb{R}}[a,b])^\infty : (y_1(t), y_2(t), ...) \in G(K_n)\ \forall t \in [a,b]\}$ because $\widehat{G}(F_n) = \widehat{G} \circ (\widetilde{G})^{-1}(D_{G(K_n)}[a,b])$. Moreover, if $\{y^m\} \subset \widehat{G}(F_n)$ satisfies $y^m \to y$ in $(D_{\mathbb{R}}[a,b])^\infty$, then $(y_1^m(t), y_2^m(t), ...) \to (y_1(t), y_2(t), ...)$ for almost all $t \in [a,b]$ (excluding the jump times in (a,b) for some y_i). (See e.g. [2] p. 121 where it is established that $x_t^n \to x_t$ if $x^n \to x$ in $D_{\mathbb{R}}[a,b]$ and x is continuous at t.) Hence, $\widehat{G}(F_n)$ is closed by the closedness of $G(K_n)$ and right continuity. To create $\widehat{X^\alpha}$, we note

$$(X^\alpha)^{-1}(S_0) = \bigcup_{n=1}^{\infty} \bigcap_{t\in\mathbb{Q}\cap[a,b]} (X_t^\alpha)^{-1}(K_n) \tag{34}$$

is measurable by the closedness of each K_n. Fixing $\Delta \in \bigcup_{n=1}^\infty K_n$ and defining

$$\widehat{Y}^\alpha \triangleq \begin{cases} \widehat{G}(X^\alpha) \text{ on } (X^\alpha)^{-1}(S_0) \\ \\ \widehat{G}(\Delta) \qquad \text{otherwise} \end{cases} \quad \text{and } \widehat{X^\alpha} \triangleq \widehat{G}^{-1}(\widehat{Y}^\alpha),$$

one finds that $\left\{\widehat{X^\alpha} = X^\alpha\right\} = \{X^\alpha \in S_0\}$ is measurable, $P\left(\widehat{X^\alpha} = X^\alpha\right) = 1$ and

$$\left\{\widehat{X^\alpha} \in O\right\} = \begin{cases} \bigcup_{n=1}^\infty \left\{\widehat{Y}^\alpha \in \widehat{G}(O \cap F_n)\right\} & \Delta \notin O \\ \\ \bigcup_{n=1}^\infty \left\{\widehat{Y}^\alpha \in \widehat{G}(O \cap F_n)\right\} \cup (X^\alpha)^{-1}(S_0^c) & \Delta \in O \end{cases} \tag{35}$$

for any open $O \subset D_E[a,b]$ and uses the proof of Proposition 3.7.1 of [10] to establish that $\widehat{Y}^\alpha$ is a $(D_{\mathbb{R}}[a,b])^\infty$-valued random variable taking values in $\widehat{G}(S_0)$. Then, observing that $\widehat{G}(O \cap F_n)$ is relatively open in (closed set) $\widehat{G}(F_n)$ by the homeomorphism, we find that $\widehat{X^\alpha}$ is a $D_E[a,b]$-valued random variable. v) follows from Lemma 24 of the Appendix with $P^\alpha = P(\widehat{X^\alpha})^{-1}$, $S = D_E[a,b]$, $\widehat{S} = (D_{\mathbb{R}}[a,b])^\infty$ and $G = \widehat{G}$.

If v) holds, then CCC holds by Proposition 11 and $\{\tilde{g}(\widehat{X}^\alpha)\}$ is tight for all $g \in \overline{C}(E)$ so MCC holds by Theorem 10 and Proposition 14 b). Hence i) holds.

Next, i) implies ii) as explained just after the definition of WMCC.

Finally, Proposition 19 establishes that WMCC implies CMCE. □

The extensions over [14] Theorem 3.1 i) are: the $\{X^\alpha\}$ are not assumed forthright to be $D_E[a,b]$-valued random variables, our WMCC condition is weaker than the condition used there and we give alternatives to WMCC. The $\widehat{G}(D_K[a,b])$ closedness is key to the proofs of our result and Theorem 3.1 i) in [14].

We now extend Theorem 3.7.2 of [10] to the general-metric-space, compact-time-interval case, generalize its MCC, and strengthen its relative compactness conclusion to tightness.

Corollary 21 *Suppose (E,d) is a metric space and $\{X^\alpha\}$ are processes with $D_E[a,b]$-valued paths. Then, each X^α is indistinguishable from a $D_E[a,b]$-valued random variable $\widehat{X}^\alpha$ such that $\left\{\widehat{X}^\alpha\right\}$ is tight on $D_E[a,b]$ if any of (MCC, PCP with d complete), (MMCC, CCC) or (WMCC, CCC) hold and only if (CCC) holds and for every $\eta > 0$ there is $\delta > 0$ so that*

$$\sup_\alpha P_\alpha(w'_d(X^\alpha,\delta) \geq \eta) \leq \eta.$$

Proof. This follows from Theorems 17 and 20 as well as Proposition 11 and Theorem 10. □

Since PCP is less stringent than PTC this corollary generalizes Skorohod's tightness theorem.

7 Appendix - Auxiliary Results

This appendix houses results that were referenced earlier. We first list a basic result which follows from the proof of Theorem 1.7 of [14]. This result and its corollary were used in

Proposition 14 and Theorem 20.

Theorem 22 *Let S be CR, $\mathcal{H}^1 \subset C(S)$ s.s.p. and $\mathcal{H} = \mathcal{H}^1 \cup \{f+g : f,g \in \mathcal{H}^1\}$. Then, $\widehat{G} : D_S[a,b] \to (D_{\mathbb{R}}[a,b])^{\mathcal{H}}$ is an imbedding, where $\widehat{G}(x) = (\widetilde{g}(x))_{g\in\mathcal{H}}$ for $x \in D_S[a,b]$.*

This theorem has a particular corollary used in Proposition 14:

Corollary 23 *Suppose S is CR, $\mathcal{G} \subset C(S)$ is countable and s.s.p., $\mathcal{H} \triangleq \mathcal{G} \cup \{g+h : g,h \in \mathcal{G}\}$, $\{Y^\alpha\}$ is a family of processes such that $\{\widetilde{h}(Y^\alpha)\}$ is tight on $D_{\mathbb{R}}[a,b]$ for all $h \in \mathcal{H}$. Then, $\{\widetilde{H}(Y^\alpha)\}$ is tight on $D_{\mathbb{R}^{\mathcal{H}}}[a,b]$, where $H \triangleq (h)_{h\in\mathcal{H}} : S \to \mathbb{R}^{\mathcal{H}}$.*

Proof. $\{(\widetilde{h}(Y^\alpha))_{h\in\mathcal{H}}\}$ is tight in $(D_{\mathbb{R}}[a,b])^{\mathcal{H}}$ by Proposition 3.2.4 of [10]. Therefore, $\{\widetilde{H}(Y^\alpha)\}$ tight in $D_{\mathbb{R}^{\mathcal{H}}}[a,b]$ by Theorem 22 (with $\mathcal{H}^1 \triangleq \mathcal{G}$) and Lemma 24 (with $F_n = S = D_{\mathbb{R}^{\mathcal{H}}}[a,b]$ and $(G(y))^i(t) = y^i(t)$ so $G(F_n) = (D_{\mathbb{R}}[a,b])^{\mathcal{H}} = \widehat{S}$). ■

The next lemma is used to prove of Proposition 14, Theorems 15, 20 and Corollary 23.

Lemma 24 *Suppose $S, \widehat{S}$ are topological spaces, $\{P^\alpha\}$, $\{Q^\alpha\}$ are collections of probability measures on $S, \widehat{S}$, and for any $n \in \mathbb{N}$ there is an $F_n \in \mathcal{B}(S)$ satisfying $\inf_\alpha P^\alpha(F_n) > 1 - \frac{1}{n}$. If $S_0 \triangleq \bigcup_n F_n$, $G : S_0 \to G(S_0) \subset \widehat{S}$ is a measurable mapping such that $G : F_n \to G(F_n)$ is a homeomorphism for each n, $G(F_n)$ is closed in $\widehat{S}$ for each n, and $Q^\alpha(A) \triangleq P^\alpha G^{-1}(A)$ for all $A \in \mathcal{B}(\widehat{S})$ and each α, then $\{P^\alpha\}$ is tight if $\{Q^\alpha\}$ is. In particular, given $\varepsilon > 0$, there exists a compact $K_\varepsilon \subset S_0$ such that $\inf_\alpha P^\alpha(K_\varepsilon) \geq 1 - \varepsilon$.*

Proof. Fixing $\varepsilon > 0$ and taking compact $K \subset \widehat{S}$ such that $\inf_\alpha Q^\alpha(K) \geq 1 - \frac{\varepsilon}{2}$ as well as F_n such that $\inf_\alpha P^\alpha(F_n) > 1 - \frac{\varepsilon}{2}$, we have that

$$P^\alpha(G^{-1}(K \cap G(F_n))) = Q^\alpha(K \cap G(F_n)) \geq 1 - Q^\alpha(K^c) - P^\alpha(F_n^c) > 1 - \varepsilon$$

for all α. Moreover, using the G^{-1}-continuity and the closedness of $G(F_n)$, one has that $G^{-1}(K \cap G(F_n))$ is compact. □

8 Acknowledgement

The author gratefully acknowledges that this work benefitted from several discussions with Douglas Blount.

References

[1] Bhatt, A. G. and Karandikar, R. L. (1993). Weak convergence to a Markov process: The martingale approach. *Probab. Theory Relat. Fields* **96**, 335–351.

[2] Billingsley, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

[3] Billingsley, P. (1999). *Convergence of Probability Measures*. Wiley, New York.

[4] Blount, D. and Kouritzin, M. A. (2010). On convergence determining and separating classes of functions. *Stoch. Proc. Appl.* **120**, 1898–1907.

[5] Budhiraja, A. and Kushner, H. J. (1999). Approximation and limit results for nonlinear filters over an infinite time interval. *SIAM J. Control Optim.* **37**, 1946–1979.

[6] Dawson, D. A. (1993). *Measure-valued Markov processes*, École d'Été de Probabilités de Saint Flour, 1991, Lecture Notes in Math., vol. **1541**, Springer, New York.

[7] Dawson, D. A. and Perkins, E. A. (1991). *Historical processes*. Mem. Amer. Math. Soc. **93** (454).

[8] Dieudonné, J. (1970). *Treatise on Analysis, volume 2*. Academic Press, , CA.

[9] Dudley, R. M. (1989). *Real Analysis and Probability*. Wadsworth & Brooks/Cole, Pacific Grove, CA.

[10] Ethier, S. and Kurtz, T. (1986). *Markov Processes, Characterization and Convergence*. Wiley, New York.

[11] Evans, S. N. and Perkins, E. A. (1991). Absolute continuity results for superprocesses with some applications. *Trans. Amer. Math. Soc.* **325**, 661-682.

[12] Fitzsimmons, P.J. (1988). Construction and regularity of measure-valued branching processes. *Israel J. Math.* **64**, 337-361.

[13] Holley, R. and Stroock, D.W. (1979). Central limit phenomena of various interacting systems. *Ann. Math.* **110**, 333–393.

[14] Jakubowski, A. (1986). On the Skorokhod topology. *Ann. Inst. Henri Poincaré* **22**, 263–285.

[15] Kallianpur, G. and Xiong, J. (1995). *Stochastic Differential Equations in Infinite Dimensional Spaces*, IMS Lecture Notes-Monograph Series **26**, Institute of Mathematical Statistics.

[16] Kouritzin, M. A. (1996). On the interrelation of almost sure invariance principles for certain stochastic adaptive algorithms and for partial sums of random variables, *J. Theoret. Probab.*, **9**, 811–840.

[17] Kouritzin, M.A. and Long, H. (2008). On extending the classical filtering equations, *Statist. Probab. Lett.*, **78**, 3195-3202.

[18] Kurtz, T. (1975). Semigroups of conditioned shifts and approximation of Markov processes. *Ann. Probab.* **3**, 618–642.

[19] Liu, H. L. and Zhou, X. W. (2012). Compact support property of the λ-Fleming-Viot process with underlying Brownian motion. *Elect. J. Probab.* **17**(73), 1–20.

[20] Meyer, P.A. and Zheng, W.A. (1984). Tightness criteria for laws of semimartingales. *Ann. Inst. Henri Poincaré B* **20**, 353–372.

[21] Mitoma, I. (1983). Tightness of probabilities on $C([0,1];\mathcal{S}')$ and $D([0,1];\mathcal{S}')$. *Ann. Probab.* **11**, 989–999.

[22] Munkres, J. R. (1975). *Topology, A First Course*. Prentice-Hall, Englewood Cliffs New Jersey.

[23] Parthasarathy, K.R. (1967). *Probability Measures on Metric Spaces*. Academic, New York.

[24] Perkins, E. (2002). *Dawson-Watanabe Superprocesses and Measure-valued Diffusions*. Springer Lecture Notes in Mathematics, no. 1781, Lectures on Probability Theory and Statistics, Ecole d'Eté de Probabilités de Saint-Flour XXIX-1999, Ed. P. Bernard, 132–329.

[25] Rudin, W. (1973). *Functional Analysis*. McGraw-Hill, New Delhi.

[26] Skorokhod, A.V. (1956). Limit theorems for Stochastic Processes. *Theory Probab. Appl.* **1**, 261–290.