OPEN ACCESS

University of BRISTOL

Ledger, S. (2016). Skorokhod's M1 topology for distribution-valued processes. Electronic Communications in Probability, 34. DOI: 10.1214/16-ECP4754

Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
10.1214/16-ECP4754

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research
General rights

Skorokhod's M1 topology for distribution-valued processes

Sean Ledger*

Abstract

Skorokhod's M1 topology is defined for càdlàg paths taking values in the space of tempered distributions (more generally, in the dual of a countably Hilbertian nuclear space). Compactness and tightness characterisations are derived which allow us to study a collection of stochastic processes through their projections on the familiar space of real-valued càdlàg processes. It is shown how this topological space can be used in analysing the convergence of empirical process approximations to distribution-valued evolution equations with Dirichlet boundary conditions.

1 Introduction

In [28] Skorokhod introduced four topologies (labelled J1, J2, M1 and M2) on the space of càdlàg paths that take values in a Banach space. The most well-known of these is the J1 topology, which has been used extensively for studying stochastic-process limits [6, 11, 34]. However, the weaker and less popular M1 topology has an important feature: its modulus of continuity vanishes when a monotone function is passed as an argument [34, Ch. 12 (4.7)]. This feature has proven useful in applications such as queuing theory [12, 18, 21, 25, 27, 32, 33], functional statistics [2, 3, 10, 20, 31, 35], scaling limits for random walks [5] and mathematical neuroscience [9].

The purpose of this paper is to extend the M1 topology to collections of càdlàg processes taking values in the space of tempered distributions or, more generally, in the dual of a countably Hilbertian nuclear space (CHNS). Following the work of Itô [14], in which a central limit theorem was developed for distribution-valued processes, the J1 topology was extended to these spaces by Mitoma [24]. These results were extended by Jakubowski to completely regular range spaces [16] and this is the focus of recent work by Kouritzin [19]. The advantage of working in the dual of a CHNS (as opposed to some Hilbertian subspace, as explored in [26]) is that these spaces have a strong finite-dimensional character. Consequently, compactness in the path space can be checked test-function-by-test-function: if A is a collection of càdlàg paths taking values in the dual of a CHNS, then A is J1-compact if and only if $\{f(\phi) = (t \mapsto f_t(\phi)) : f \in A\}$ is J1-compact in the space of real-valued càdlàg paths, for every test function, ϕ, in the

---

*Heilbronn Institute, University of Bristol, United Kingdom. E-mail: sean.ledger@bristol.ac.uk

CHNS [24, Thm. 3.1 & 4.1]. Hence the tightness of a sequence of distribution-valued càdlàg processes can be established by projecting down to the familiar space of real-valued càdlàg processes. It should be mentioned that contributions like [16], [7] and [19] also contain results using real-valued projections for more general spaces, but they require establishing point-wise tightness or existence of a limit.

Our aim is to fill the gap in theory between these two settings and to combine the temporal properties of the M1 topology with the spatial properties of the tempered distributions. In Section 2 we construct the M1 topology for càdlàg paths taking values in the dual of a CHNS (Definition 2.6). The corresponding compactness and tightness criteria to those of [24] are stated and proved in Section 3 (Theorems 3.1, 3.2 and 3.3). Finally, we use our tools on a concrete example in Section 4, specifically we prove the tightness of a sequence of discrete empirical-measure processes that approximate the solution of a stochastic evolution equation with Dirichlet boundary conditions. Here, the mass lost at the boundary is a monotone process, hence the M1 topology offers a simple decomposition trick for controlling the fluctuations in the approximating sequence (Proposition 4.2).

2 Construction of the topology

We refer the reader to [4, 17, 28, 34] for the basic theory of countably Hilbertian nuclear spaces (CHNS) and the standard Skorokhod topologies for Banach range spaces. Our construction will mirror [24] for the J1 topology.

Throughout, E will denote a general CHNS (a specific example being $\mathcal{S}$, the space of rapidly decreasing functions [17, Ex. 1.3.2]). The properties we will use are:

- E is a linear topological space generated by an increasing sequence of Hilbertian semi-norms $\|\cdot\|_0 \leq \|\cdot\|_1 \leq \|\cdot\|_2 \leq \cdots$,
- The closure, E_n, of $(E, \|\cdot\|_n)$ is a separable Hilbert space with $E_0 \supseteq E_1 \supseteq E_2 \supseteq \cdots$ and $E = \bigcap_{n \geq 0} E_n$,
- For every $n \geq 0$, there exists $m > n$ such that the canonical injection $E_m \hookrightarrow E_n$ is Hilbert–Schmidt, that is

$$\sum_{i=1}^{\infty} \|e_i^m\|_n^2 < \infty, \qquad \text{whenever } \{e_i^m\}_{i \geq 1} \text{ is an orthonormal system of } E_m,$$

- The topological duals, $E_{-n} = E_n'$, satisfy $\|\cdot\|_0 \geq \|\cdot\|_{-1} \geq \|\cdot\|_{-2} \geq \cdots$, $E_0 \subseteq E_{-1} \subseteq E_{-2} \subseteq \cdots$ and $E' = \bigcup_{n \geq 0} E_{-n}$,
- The *strong topology* of E' is that generated by the collection of semi-norms $\{p_B : E' \to [0, \infty)\}_{B \in \mathcal{B}}$, where $p_B(f) := \sup_{x \in B} |f(x)|$ and $\mathcal{B}$ is the collection of bounded subsets of E.

Definition 2.1 ($D_{E'}$)**.** *The space of càdlàg paths, $D_{E'}$, is defined to be the collection of functions mapping $[0, T]$ to E' that are right-continuous and have left limits with respect to the strong topology on E'.*

In practice checking Definition 2.1 is straightforward due to [23]. As in the classical case, we define the M1 topology on $D_{E'}$ through a (pseudo-)graph distance on $E' \times [0, T]$. The graph of an element in $D_{E'}$ is formed by joining up its points of discontinuity with intervals (see Figure 1):

Definition 2.2 (Interval)**.** *For f and g in E', define the* interval *between these points to be*

$$[f, g]_{E'} := \{(1 - \lambda) f + \lambda g \in E' : \lambda \in [0, 1]\}.$$

If h_1 and h_2 are two points in the interval, $h_1 = (1 - \lambda_1)f + \lambda_1 g$ and $h_2 = (1 - \lambda_2)f + \lambda_2 g$, then we write $h_1 \leq_{[f,g]_{E'}} h_2$ whenever $\lambda_1 \leq \lambda_2$.
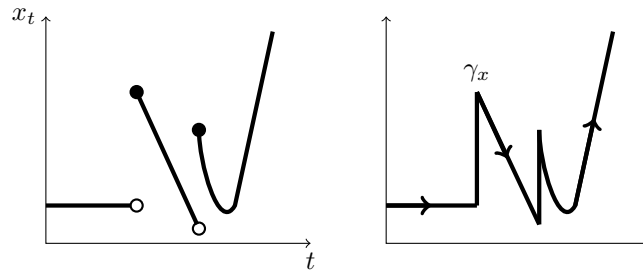
Figure 1: An element of $D_{\mathbb{R}}$ (left) together with its graph, $\gamma_x \subseteq \mathbb{R} \times [0, T]$, and graph ordering (right).

Definition 2.3 (Graph). *The graph of an element $x \in D_{E'}$ is defined to be the closed subset of $E' \times [0, T]$*

$$\gamma_x := \{(z, t) \in E' \times [0, T] : z \in [x(t-), x(t)]_{E'}\}.$$

For $(z_1, t_1), (z_2, t_2) \in \gamma_x$, we say that $(z_1, t_1) \leq_{\gamma_x} (z_2, t_2)$ if either:

$$\text{(i) } t_1 < t_2 \qquad \text{or} \qquad \text{(ii) } t_1 = t_2 \text{ and } z_1 \leq_{[x(t_1-), x(t_1)]_{E'}} z_2.$$

Definition 2.4 (Parametric representation). *A parametric representation, λ_x, of the graph γ_x is a continuous surjection $\lambda_x : [0, 1] \to \gamma_x$ that is non-decreasing with respect to the graph ordering on γ_x. For $x \in D_X$, let the collection of all such parametrisations of γ_x be denoted Λ_x.*

We can define a family of pseudometrics on $D_{E'}$ by using the family of semi-norms, $\{p_B\}_{B \in \mathcal{B}}$, on E' to measure the graph distance between two parametric representations:

Definition 2.5 (A family of pseudometrics). *Fix $B \in \mathcal{B}$. Let x and y be elements of $D_{E'}$. For parametric representations $\lambda_x = (z_x, t_x) \in \Lambda_x$ and $\lambda_y = (z_y, t_y) \in \Lambda_x$ define*

$$g_B(\lambda_x, \lambda_y) := \sup_{s \in [0,1]} \max\{p_B(z_x(s) - z_y(s)), |t_x(s) - t_y(s)|\}.$$

A pseudometric on $D_{E'}$ is given by

$$d_{B, \mathrm{M1}}(x, y) := \inf_{\lambda_x \in \Lambda_x, \lambda_y \in \Lambda_y} g_B(\lambda_x, \lambda_y), \qquad \text{for } x, y \in D_{E'}.$$

The reader can verify that Definition 2.5 gives a pseudometric using [34, Thm. 12.3.1]. The family $\{d_{B, \mathrm{M1}}\}_{B \in \mathcal{B}}$ generates a topology which we will call the M1 topology on $D_{E'}$:

Definition 2.6 (M1 topology). *The M1 topology on $D_{E'}$ is defined to be the projective limit topology of $\{d_{B, \mathrm{M1}}\}_{B \in \mathcal{B}}$. That is, the topology with neighbourhoods*

$$\{x \in D_{E'} : d_{B_i, \mathrm{M1}}(x, x_0) < \varepsilon_i, \quad \text{for } 1 \leq i \leq n\},$$

where $x_0 \in D_{E'}$, $B_i \in \mathcal{B}$, $\varepsilon_i > 0$ and $n \in \mathbb{N}$.

The following is a collection of facts that are either straightforward or have very similar proofs to the J1 topology case, and so we omit the full details.

Proposition 2.7 (Some properties of $(D_{E'}, \mathrm{M1})$). *We have the following:*

(i) For every $\phi \in E$, the canonical projection

$$\pi^\phi : (D_{E'}, \mathrm{M1}) \to (D_{\mathbb{R}}, \mathrm{M1}), \qquad \pi^\phi(x) = (x_t(\phi))_{t \in [0,T]} =: x(\phi)$$

is continuous,

(ii) For every $n \geq 0$, the canonical inclusion $\iota_n : (D_{E_{-n}}, \mathrm{M1}) \to (D_{E'}, \mathrm{M1})$ is continuous,

(iii) $(D_{E'}, \mathrm{M1})$ is completely regular,

(iv) Let $\mathfrak{B}(D_{E'})$ denote the Borel σ-algebra on $(D_{E'}, \mathrm{M1})$ and $\mathcal{K}$ the Kolmogorov σ-algebra generated by the projections

$$\pi_{t_1, t_2, \ldots, t_n}^{\phi_1, \phi_2, \ldots, \phi_n} : D_{E'} \to \mathbb{R}^n, \quad \pi_{t_1, t_2, \ldots, t_n}^{\phi_1, \phi_2, \ldots, \phi_n}(x) = (x_{t_1}(\phi_1), x_{t_2}(\phi_2), \ldots, x_{t_n}(\phi_n)).$$

Then $\mathfrak{B}(D_{E'}) = \mathcal{K}$.

Proof. (i) Use that $\{\phi\} \in \mathcal{B}$ for every $\phi \in E$. (ii) Straightforward.

(iii) From [17, Thm. 2.1.1] it suffices to show that for every $x^1 \neq x^2 \in D_{E'}$ there exists a bounded set, B, such that $d_{B, \mathrm{M1}}(x^1, x^2) > 0$. So take $B = \{\phi\}$ for any $\phi \in E$ such that $x^1(\phi) \neq x^2(\phi)$.

(iv) The inclusion $\mathcal{K} \subseteq \mathfrak{B}(D_{E'})$ is straightforward since $\pi_t^\phi = \pi_t \circ \pi^\phi$ with the canonical projections

$$\pi^\phi : (D_{E'}, \mathrm{M1}) \to (D_{\mathbb{R}}, \mathrm{M1}) \quad \text{and} \quad \pi_t : (D_{E'}, \mathbb{R}) \to \mathbb{R}$$

and π^ϕ is continuous by (i) and π_t is measurable by [34, Thm 11.5.2]. To show $\mathfrak{B}(D_{E'}) \subseteq \mathcal{K}$, let $\iota_n : D_{E_{-n}} \hookrightarrow D_{E'}$ be the canonical inclusion. From [17, Thm. 1.3.1], $D_{E'} = \bigcup_{n \geq 0} D_{E_{-n}}$, thus for any open subset, U, of $(D_{E'}, \mathrm{M1})$

$$U = \bigcup_{n=0}^{\infty} \iota_n^{-1}(U). \tag{2.1}$$

By (ii) ι_n is continuous, so

$$\iota_n^{-1}(U) \in \mathfrak{B}(D_{E_{-n}}, \mathrm{M1}) = \mathfrak{B}(D_{E_{-n}}, \mathrm{J1}) = \sigma(\pi_t : D_{E_{-n}} \to E_{-n}) \subseteq \sigma(\pi_t : D_{E'} \to E'),$$

where the first equality follows from [34, Thm. 11.5.2] (since E_{-n} is a Banach space) and the second equality follows from the reasoning in [11, Prop. 3.7.1]. From [17, Thm. 3.1.1], $\mathfrak{B}(E') = \sigma(\pi^\phi : E' \to \mathbb{R})$, so it follows that $\mathcal{K} = \sigma(\pi_t : D_{E'} \to E')$. Therefore $\iota_n^{-1}(U) \in \mathcal{K}$, and the result is complete by (2.1). $\square$

3 Compactness and tightness characterisations

The following three theorems characterise compactness, tightness and weak convergence in $(D_{E'}, \mathrm{M1})$. All notation is as defined in Section 2.

Theorem 3.1 (Compactness characterisation)**.** *Let $A \subseteq D_{E'}$. The following are equivalent:*

(i) A is compact in $(D_{E'}, \mathrm{M1})$,

(ii) For every $\phi \in E$, $\pi^\phi(A) = \{x(\phi) : x \in A\}$ is compact in $(D_{\mathbb{R}}, \mathrm{M1})$,

(iii) There exists $p \in \mathbb{N}$ such that $A \subseteq D_{E_{-p}}$ and A is compact in $(D_{E_{-p}}, \mathrm{M1})$.

Theorem 3.2 (Tightness characterisation)**.** *Let $(\mu_n)_{n \geq 1}$ be a sequence of probability measures on $D_{E'}$. Then $(\mu_n)_{n \geq 1}$ is tight on $(D_{E'}, \mathrm{M1})$ if and only if $\left(\mu_n \circ (\pi^\phi)^{-1}\right)_{n \geq 1}$ is tight on $(D_{\mathbb{R}}, \mathrm{M1})$ for every $\phi \in E$.*

Furthermore, if $(\mu_n)_{n \geq 1}$ is tight on $(D_{E'}, \mathrm{M1})$, then it is relatively compact on $(D_{E'}, \mathrm{M1})$.

Theorem 3.3 (Weak convergence)**.** *If $(\mu_n)_{n \geq 1}$ is tight on $(D_{E'}, \mathrm{M1})$ and for every $k \geq 1$, $\phi_1, \phi_2, \ldots, \phi_k \in E$ and $t_1, t_2, \ldots, t_k \in \mathrm{cont}(\mu)$*

$$\mu_n \circ (\pi_{t_1, t_2, \ldots, t_k}^{\phi_1, \phi_2, \ldots, \phi_k})^{-1} \to \mu \circ (\pi_{t_1, t_2, \ldots, t_k}^{\phi_1, \phi_2, \ldots, \phi_k})^{-1}, \quad \text{weakly on } \mathbb{R}^k,$$

as $n \to \infty$, where $\mathrm{cont}(\mu) = \{t \in [0, T] : \mu\{x \in D_{E'} : x_{t-} = x_t\} = 1\}$, then $(\mu_n)_{n \geq 1}$ converges weakly to μ on $(D_{E'}, \mathrm{M1})$.

Before proving these results we need the following technical result, which exploits the nuclear space structure of E to enable a generic subset $A \subseteq D_{E'}$ to be controlled in terms of its projection onto a finite number of basis vectors. For a normed linear space X, we will denote the *M1 modulus of continuity on D_X* by

$$w_{X, \mathrm{M1}}(x; \delta) := \sup_{t \in [0, T]} \sup_{\mathrm{trip}(t; \delta)} \inf_{\lambda \in [0, 1]} \|x_{t_2} - (1 - \lambda) x_{t_1} - \lambda x_{t_3}\|_X, \qquad x \in D_X, \delta > 0, \quad (3.1)$$

where $\mathrm{trip}(t; \delta) = \{(t_1, t_2, t_3) : \max(0, t - \delta) \leq t_1 < t_2 < t_3 < \min(t + \delta, T)\}$.

Lemma 3.4 (Controlling the modulus of continuity)**.** *Let $p > n$ be such that the inclusion $E_p \hookrightarrow E_n$ is Hilbert–Schmidt and let $A \subseteq D_{E'}$ be such that*

$$c := \sup_{x \in A} \sup_{t \in [0, T]} \|x_t\|_{-n} < \infty.$$

Then, for every $\varepsilon > 0$, there exists $k \geq 1$ and $\phi_1, \phi_2, \ldots, \phi_k \in E$ such that

$$\sup_{x \in A} w_{E_{-p}, \mathrm{M1}}(x; \delta) \leq 2c\varepsilon^{-1} \max_{i = 1, \ldots, k} \sup_{x \in A} w_{\mathbb{R}, \mathrm{M1}}(x(\phi_i); \delta) + 3c\varepsilon, \qquad \text{for every } \delta > 0.$$

The proof is a technical computation for which the next lemma will be helpful. This argument is just a repackaging of the Heine–Borel theorem and is adapted from [13, Lem. A.28].

Lemma 3.5 (A geometric argument)**.** *Fix $M \in \mathbb{N}$ and $\varepsilon > 0$, and let B be the closed unit ball in $\mathbb{R}^M$. Then there exists a finite set $\Theta = \{\theta_1, \ldots, \theta_k\} \subseteq \mathbb{R}^M$ such that $\|\theta_i\|_{\mathbb{R}^M} = 1$, and for all $v_1, v_2 \in B$ satisfying*

$$\min_{\lambda \in [0, 1]} \|\lambda v_1 + (1 - \lambda) v_2\|_{\mathbb{R}^M} \geq \varepsilon,$$

there is an $i \in \{1, 2, \ldots, k\}$ for which

$$\|\lambda v_1 + (1 - \lambda) v_2\|_{\mathbb{R}^M} \leq 2\varepsilon^{-1} |(\lambda v_1 + (1 - \lambda) v_2) \cdot \theta_i|,$$

for every $\lambda \in [0, 1]$. Here, $\|\cdot\|_{\mathbb{R}^M}$ is the Euclidean norm and k depends on M and ε.

Proof. For $v_1, v_2 \in B$, let $\lambda_0 = \lambda_0(v_1, v_2)$ be a minimiser of

$$\lambda \mapsto \|\lambda v_1 + (1 - \lambda) v_2\|_{\mathbb{R}^M}$$

on $[0, 1]$. Define

$$S = \{(v_1, v_2) \in B \times B : \|\lambda_0 v_1 + (1 - \lambda_0) v_2\|_{\mathbb{R}^M} \geq \varepsilon\},$$

then for $(v_1, v_2) \in S$ set $\theta(v_1, v_2)$ to be the unit vector in the direction of the pair's minimiser:

$$\theta(v_1, v_2) = \frac{\lambda_0 v_1 + (1 - \lambda_0) v_2}{\|\lambda_0 v_1 + (1 - \lambda_0) v_2\|_{\mathbb{R}^M}},$$

so that for every $\lambda \in [0, 1]$

$$(\lambda v_1 + (1 - \lambda) v_2) \cdot \theta(v_1, v_2) \geq \varepsilon \geq \varepsilon \|\lambda v_1 + (1 - \lambda) v_2\|_{\mathbb{R}^M} > \frac{\varepsilon}{4} \|\lambda v_1 + (1 - \lambda) v_2\|_{\mathbb{R}^M}.$$

Hence there is an open neighbourhood, $N(v_1, v_2)$, in $\mathbb{R}^M \times \mathbb{R}^M$ about (v_1, v_2) such that for all $(w_1, w_2) \in N(v_1, v_2)$

$$\frac{\varepsilon}{2} \left\| \lambda w_1 + (1 - \lambda) w_2 \right\|_{\mathbb{R}^M} \leq (\lambda w_1 + (1 - \lambda) w_2) \cdot \theta(v_1, v_2).$$

Now S is a compact set with open cover $\{N(v_1, v_2)\}_{(v_1, v_2) \in S}$, therefore it is possible to take a finite sub-cover, $\{N(v_1^i, v_2^i)\}_{i=1,\ldots,k}$. To complete the proof let $\Theta = \{\theta(v_1^i, v_2^i)\}_{i=1,\ldots,k}$.
□

Proof of Lemma 3.4. For readability, introduce the notation

$$y^\lambda = y(\lambda; x; t_1, t_2, t_3) := x_{t_2} - \lambda x_{t_1} - (1 - \lambda) x_{t_3} \in E',$$

where $\lambda \in [0, 1]$, $x \in D_{E'}$ and $t_1, t_2, t_3 \in [0, T]$. Then, recalling the notation of (3.1),

$$w_{E_{-p}, \mathrm{M1}}(x; \delta) = \sup_{t, \mathrm{trip}(t; \delta)} \inf_{\lambda \in [0,1]} \left\| y(\lambda; x; t_1, t_2, t_3) \right\|_{-p},$$

and the triangle inequality gives that $\|y^\lambda\|_{-n} \leq c$, whenever $x \in A$.

To construct the finite family of vectors, first notice that $E_{-n} \subseteq E_{-p}$, so for an orthonormal basis, $\{e_i^p\}_{i \geq 1}$, of E_p

$$\left| y^\lambda(e_i^p) \right| \leq \left\| y^\lambda \right\|_{-n} \left\| e_i^p \right\|_n \leq c \left\| e_i^p \right\|_n,$$

for $x \in A$. Therefore if M is chosen large enough so that $\sum_{i=M+1}^{\infty} \left\| e_i^p \right\|_n^2 \leq \varepsilon$, then

$$\left\| y^\lambda \right\|_{-p}^2 \leq \sum_{i=1}^{M} \left| y^\lambda(e_i^p) \right|^2 + c\varepsilon, \tag{3.2}$$

so it now suffices to work with $e_1^p, e_2^p \ldots, e_M^p$. By introducing the $\mathbb{R}^M$ vectors

$$v_1 := \begin{pmatrix} (x_{t_2} - x_{t_1})(e_1^p) \\ (x_{t_2} - x_{t_1})(e_2^p) \\ \vdots \\ (x_{t_2} - x_{t_1})(e_M^p) \end{pmatrix} \quad \text{and} \quad v_2 := \begin{pmatrix} (x_{t_2} - x_{t_3})(e_1^p) \\ (x_{t_2} - x_{t_3})(e_2^p) \\ \vdots \\ (x_{t_2} - x_{t_3})(e_M^p) \end{pmatrix},$$

the bound in (3.2) gives

$$w_{E_{-p}, \mathrm{M1}}(x; \delta) \leq \sup_{t, \mathrm{trip}(t; \delta)} \inf_{\lambda \in [0,1]} \left\| \lambda v_1 + (1 - \lambda) v_2 \right\|_{\mathbb{R}^M} + c\varepsilon.$$

With this fixed choice of ε and M, take $\Theta = \{\theta_1, \theta_2, \ldots, \theta_k\}$ from Lemma 3.5. If it is the case that

$$\inf_{\lambda \in [0,1]} \left\| \lambda v_1 + (1 - \lambda) v_2 \right\|_{\mathbb{R}^M} \geq 2c\varepsilon \tag{3.3}$$

then, by normalising v_1 and v_2 by $2c$, it follows that there is some $i \in \{1, 2, \ldots, k\}$ such that

$$\inf_{\lambda \in [0,1]} \left\| \lambda v_1 + (1 - \lambda) v_2 \right\|_{\mathbb{R}^M} \leq 2c\varepsilon^{-1} \inf_{\lambda \in [0,1]} \left| \theta_i \cdot (\lambda v_1 + (1 - \lambda) v_2) \right|.$$

If (3.3) fails to be true, then clearly the upper bound above can be replaced by $2c\varepsilon$. Hence, for every $x \in A$

$$w_{E_{-p}, \mathrm{M1}}(x; \delta) \leq 2c\varepsilon^{-1} \sup_{t, \mathrm{trip}(t; \delta)} \max_{i=1,\ldots,k} \inf_{\lambda \in [0,1]} \left| \theta_i \cdot (\lambda v_1 + (1 - \lambda) v_2) \right| + 3c\varepsilon. \tag{3.4}$$

For each $i \in \{1, 2, \dots, k\}$, ϕ_i can now be constructed by defining

$$\phi_i := \sum_{j=1}^{M} \theta_i^{(j)} e_j^p \in E_p,$$

where $\theta_i^{(j)} \in \mathbb{R}$ is the jth coordinate of θ_i. Inequality (3.4) therefore reduces to

$$w_{E_{-p}, \text{M1}}(x; \delta) \le 2c\varepsilon^{-1} \sup_{t, \text{trip}(t; \delta)} \max_{i=1, \dots, k} \inf_{\lambda \in [0,1]} \left| y^\lambda(\phi_i) \right| + 3c\varepsilon \le 2c\varepsilon^{-1} \max_{i=1, \dots, k} w_{\mathbb{R}, \text{M1}}(x(\phi_i; \delta)) + 3c\varepsilon$$

where the second inequality is due to switching the maximum and supremum. Taking a supremum over $x \in A$ and repeating the switch once more yields the result. $\qquad\square$

By taking $\lambda = 1$ in the above proof, the corresponding result for the increments of x is given:

Corollary 3.6 (Controlling increments)**.** *With x, A, c, ε and $\phi_1, \phi_2, \dots, \phi_k$ as in the statement of Lemma 3.4,*

$$\sup_{x \in A} \sup_{t \in (s-\delta, s+\delta) \cap [0,T]} \|x_t - x_s\|_{E_{-p}} \le 2c\varepsilon^{-1} \max_{i=1, \dots, k} \sup_{x \in A} \sup_{t \in (s-\delta, s+\delta) \cap [0,T]} |x_t(\phi_i) - x_s(\phi_i)| + 3c\varepsilon,$$

for every $\delta > 0$ and $s \in [0, T]$.

Proof of Theorem 3.1. (i) $\Rightarrow$ (ii) and (iii) $\Rightarrow$ (i) follow from Proposition 2.7 (i) and (ii). We now prove (ii) $\Rightarrow$ (iii).

The first half of the proof of [17, Thm. 2.4.4] does not depend on the choice of temporal topology, hence we have $p > n$ for which $A \subseteq D_{E_{-n}} \subseteq D_{E_{-p}}$, $E_p \hookrightarrow E_n$ is Hilbert–Schmidt and

$$c := \sup_{x \in A} \sup_{t \in [0,T]} \|x_t\|_{-n} < \infty.$$

Therefore A will be compact in $(D_{E_{-p}}, \text{M1})$ if we can verify the second condition in [34, Thm. 12.12.2] (with $D_{E_{-p}}$ as the range space).

Let $\phi_1, \phi_2, \dots, \phi_k \in E$ be as in the conclusion of Lemma 3.4. Since each $\pi^{\phi_i}(A)$ is compact (by hypothesis), [34, Thm. 12.12.2] implies (recall the notation of (3.1))

$$\max_{i=1,2,\dots,k} \sup_{x \in A} w_{\mathbb{R}, \text{M1}}(x(\phi_i); \delta) \to 0, \qquad \text{as } \delta \to 0.$$

Therefore by Lemma 3.4, $\limsup_{\delta \to 0} \sup_{x \in A} w_{E_{-p}, \text{M1}}(x; \delta) \le 3c\varepsilon$, and likewise for the terms in Corollary 3.6. Since $\varepsilon > 0$ is arbitrary, we are done. $\qquad\square$

Proof of Theorem 3.2. To prove the second statement, assume that (μ_n) is tight on $(D_{E'}, \text{M1})$. For every $p \ge 0$, $D_{E_{-p}}$ is a Polish space, so [29, Sec. 3, Def. 2, Ex. 1] implies that $D_{E_{-p}}$ is a topological Radon space. From [29, Sec. 3, Ex. 4] $(D_{E'}, \text{M1})$ is a topological Radon space, hence every probability measure on $(D_{E'}, \text{M1})$ is a Radon measure. By Theorem 3.1 (iii), every compact subset of $(D_{E'}, \text{M1})$ is metrizable, therefore [29, Sec. 5, Thm. 2] completes the proof of the second result, since $(D_{E'}, \text{M1})$ is completely regular (Proposition 2.7 (iii)).

The first part of the theorem follows from the work in the proof of Theorem 3.1 and [17, Thm. 2.5.1]. $\qquad\square$

Proof of Theorem 3.3. Since the Borel and Kolmogorov σ-algebras on $(D_{E'}, \text{M1})$ coincide (Proposition 2.7 (iv)), the result follows by [24, Prop. 5.1]. $\qquad\square$

4 Application to empirical processes with Dirichlet boundary conditions

In the remainder of the paper we will show how our machinery can be applied to the problem of approximating stochastic evolution equations through empirical averages of microscopic particles. We will be very concrete and consider a problem from mathematical finance, specifically large portfolio credit modelling. Our analysis will show that $(D_{E'}, \mathrm{M1})$ can be a convenient space on which to prove tightness when studying systems with Dirichlet boundary conditions.

Define a collection of correlated Brownian motions, $\{X^{i,N}\}_{i=1,\dots,N}$, started on the half-line and evolving with the dynamics

$$X_t^{i,N} = X_0^i + \int_0^t \rho(L_s^N)dW_s + \int_0^t \sqrt{1-\rho(L_s^N)^2}dW_s^i, \qquad L_t^N := \frac{1}{N}\sum_{i=1}^N \mathbb{1}_{\tau^{i,N}\le t}, \quad (4.1)$$

where $\tau^{i,N} := \inf\{t > 0 : X_t^{i,N} \le 0\}$. Here, $W, W^1, W^2, \dots$ are independent Brownian motions, $\{X^i\}_{i\ge 1}$ are i.i.d. with some density $f : (0,\infty) \to [0,\infty)$ and $\rho : [0,1] \to [0,1]$ is a measurable function. We will not impose any further regularity constraints on ρ in the following tightness calculations. This is a system in which the proportion of particles that have hit the origin determines the correlation in the system.

(To see that such $\{X^{i,N}\}_{i=1,\dots,N}$ exist, notice that $t \mapsto L_t^N$ is piecewise constant. Therefore, to construct the discrete system, take N Brownian motions with initial correlation $\rho(0)$, stop the system at the first hitting of zero, restart the system with correlation $\rho(1/N)$ and repeat.)

Our set-up extends the constant correlation model introduced in [8]. The motivation for this particular form is to address the correlation skew seen in [8, Sec. 5]. To analyse the model, the quantity of interest is the empirical measure of the population:

$$\nu_t^N = \frac{1}{N}\sum_{i=1}^N \mathbb{1}_{t<\tau^{i,N}}\delta_{X_t^{i,N}}, \qquad (\text{hence } L_t^N = 1 - \nu_t^N(0,\infty)), \quad (4.2)$$

where δ_x is the usual Dirac delta measure at the point $x \in \mathbb{R}$.

We would like to establish the weak convergence (at the process level) of $(\nu^N)_{N\ge 1}$ to some limit ν, which should be the solution of the non-linear evolution equation

$$d\nu_t(\phi) = \frac{1}{2}\nu_t(\phi'')dt + \rho(L_t)\nu_t(\phi')dW_t, \qquad L_t = 1 - \nu_t(0,\infty). \quad (4.3)$$

with test functions $\phi \in \mathcal{S}$ that satisfy $\phi(0) = 0$. This is an example of a (stochastic) McKean–Vlasov equation [30]. Proving existence and uniqueness of solutions to this equation would require further regularity constraints on ρ. For now, we will only demonstrate that $(\nu^N)_{N\ge 1}$ is tight on the space $(D_{\mathcal{S}'}, \mathrm{M1})$, where $\mathcal{S}'$ is the space of tempered distributions.

Notice that, for every t, ν_t^N is a sub-probability measure, so is an element of $\mathcal{S}'$, and for every ϕ, $\nu_t^N(\phi)$ is a real-valued càdlàg function. Therefore ν^N has a version that is càdlàg, by [23], so $D_{\mathcal{S}'}$ is an appropriate space to work with. By Theorem 3.2, it suffices to show $\nu^N(\phi)$ is tight in $(D_{\mathbb{R}}, \mathrm{M1})$ for every $\phi \in \mathcal{S}$, and for that it is sufficient to verify [34, Thm. 12.12.3], the first condition of which is trivial since $|\nu_t^N(\phi)| \le \|\phi\|_\infty$. For demonstrating that the second condition of [34, Thm. 12.12.3] holds, we can employ the helpful result [1, Thm. 1], so to summarise what is now required:

Proposition 4.1. *The sequence $(\nu^N)_{N\ge 1}$ is tight in $(D_{\mathcal{S}'}, \mathrm{M1})$ if, for every fixed $\phi \in \mathcal{S}$, there exists $a, b, c > 0$ such that*

$$\mathbf{P}(H_{\mathbb{R}}(\nu_{t_1}^N(\phi), \nu_{t_2}^N(\phi), \nu_{t_3}^N(\phi)) \ge \eta) \le c\eta^{-a}|t_3 - t_1|^{1+b}$$

for all $N \ge 1$, $\eta > 0$, and $0 \le t_1 < t_2 < t_3 \le T$, and

$$\lim_{N\to\infty} \mathbf{P}\big(\sup_{t\in(0,\delta)} |\nu_t^N(\phi) - \nu_0^N(\phi)| + \sup_{t\in(T-\delta,T)} |\nu_T^N(\phi) - \nu_t^N(\phi)| \geq \eta \big) = 0, \qquad \textit{for every } \eta > 0.$$

Here, $H_{\mathbb{R}}(v_1, v_2, v_3) := \inf_{\lambda\in[0,1]} |v_2 - (1-\lambda)v_1 + \lambda v_3|$.

The challenge in working with $(\nu^N)_{N\geq 1}$ is the discontinuity presented by the absorbing boundary at the origin. For the constant ρ case, the authors of [8] use explicit estimates based on 2d Brownian motion in a wedge [15, 22] to control for the boundary effects. With the more complicated interactions in our present model such methods seem intractable, however, the M1 topology provides an alternative approach.

Introduce the process $\bar{\nu}^N \in D_{\mathcal{S}'}$ defined by

$$\bar{\nu}_t^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{X_{t\wedge\tau^{i,N}}^{i,N}}.$$

This has the advantage of being continuous (in time), so its increments are easy to control, and it can be related to ν^N through the simple fact

$$\bar{\nu}_t^N(\phi) = \nu_t^N(\phi) + \phi(0)L_t^N, \qquad \text{for every } \phi \in \mathcal{S}. \tag{4.4}$$

Thus, ν^N is a linear combination of a process with well-behaved increments and a process that has zero M1 modulus of continuity (L^N is monotone). Substituting (4.4) into $H_{\mathbb{R}}$ from Proposition 4.1 gives:

Proposition 4.2 (Decomposition trick). *For every $\phi \in \mathcal{S}$, $N \geq 1$ and $0 \leq t_1 < t_2 < t_3 \leq T$*

$$H_{\mathbb{R}}(\nu_{t_1}^N(\phi), \nu_{t_2}^N(\phi), \nu_{t_3}^N(\phi)) \leq |\bar{\nu}_{t_1}^N(\phi) - \bar{\nu}_{t_2}^N(\phi)| + |\bar{\nu}_{t_2}^N(\phi) - \bar{\nu}_{t_3}^N(\phi)|.$$

Proof. Carry out the aforementioned substitution and apply the triangle inequality to get

$$\text{l.h.s.} \leq \text{r.h.s.} + |\phi(0)| \inf_{\lambda\in[0,1]} |L_{t_2}^N - (1-\lambda)L_{t_1}^N - \lambda L_{t_3}^N|.$$

Since L^N is monotone increasing, the final term is zero for $t_1 < t_2 < t_3$, and this completes the proof. $\qquad\square$

Our trick makes the remainder of the tightness proof routine:

Theorem 4.3. *The sequence $(\nu^N)_{N\geq 1}$ is tight on $(D_{\mathcal{S}'}, \text{M1})$.*

Proof. For $\phi \in E$, let $\|\phi\|_{\text{lip}}$ denote its Lipschitz constant. Since the $X^{i,N}$ are just 1d Brownian motions, Hölder's inequality gives

$$\mathbf{E}[|\bar{\nu}_t^N(\phi) - \bar{\nu}_s^N(\phi)|^4] \leq \frac{1}{N}\sum_{i=1}^{N} \mathbf{E}[|\phi(X_{t\wedge\tau^{i,N}}^{i,N}) - \phi(X_{s\wedge\tau^{i,N}}^{i,N})|^4] \leq \|\phi\|_{\text{lip}}^4 \mathbf{E}[|X_{t\wedge\tau^{1,N}}^{1,N} - X_{s\wedge\tau^{1,N}}^{1,N}|^4],$$

and the final expression is $O(|t-s|^2)$, uniformly in N. Therefore Markov's inequality and Proposition 4.2 give the first statement in Proposition 4.1.

For the second statement in Proposition 4.1, we can first apply the decomposition in (4.4) to get

$$\sup_{t\in(0,\delta)} |\nu_t^N(\phi) - \nu_0^N(\phi)| \leq \sup_{t\in(0,\delta)} |\bar{\nu}_t^N(\phi) - \bar{\nu}_0^N(\phi)| + |\phi(0)|L_\delta^N,$$

and likewise for the term at $T-\delta$. Using Doob's maximal inequality and repeating the Hölder calculation above gives

$$\mathbf{E} \sup_{t\in(0,\delta)} |\nu_t^N(\phi) - \nu_0^N(\phi)| \leq o(1) + |\phi(0)|\mathbf{E}L_\delta^N = o(1) + |\phi(0)|\mathbf{P}(0 < \tau^{1,N} \leq \delta).$$

Since $X^{1,N}$ is a Brownian motion, the final term is $o(1)$ uniformly in N, as $\delta \to 0$. So applying Markov's inequality gives the second statement in Proposition 4.1. $\qquad\square$

Remark 4.4 (Full convergence). To prove full weak convergence of $(\nu^N)_{N\geq 1}$, one approach would be to show that all limit points are supported on solutions of the evolution

equation (4.3) and that this equation has a unique solution. The latter task is dependent on the level of regularity of ρ, however the former could be proved using martingale methods. To this end, an application of Itô's formula gives

$$d\nu_t^N(\phi) = \frac{1}{2}\nu_t^N(\phi'')dt + \rho(L_t^N)\nu_t^N(\phi')dW_t + dI_t^N(\phi), \quad \text{where} \quad \mathbf{E}|I_t^N(\phi)|^2 = O(N^{-1}),$$

for every $\phi \in \mathcal{S}$ such that $\phi(0) = 0$. By Theorem 4.3, we have subsequential weak limits for every term in this equation, and the M1 topology (on $\mathbb{R}$) is well-behaved with respect to integrals [34, Thm. 11.5.1].

Remark 4.5 (Why not work with $\bar{\nu}^N$?)**.** It might seem easier to work with the process $\bar{\nu}^N$ from the start. Notice, however, that the above evolution equation would only hold for $\bar{\nu}^N$ if $\phi'(0) = 0 = \phi''(0)$. This has the effect of moving the boundary problems we had in calculating the J1 modulus of continuity of $(\nu^N)_{N \geq 1}$ to the analysis of the limiting evolution equation.

Remark 4.6 (Why $\mathcal{S}'$?)**.** $\mathcal{S}'$ seems an excessively large range space for the above example, however, it is easy to recover that any limiting process must be measure-valued by the Riesz–Markov–Kakutani theorem. Developing a general theory for $\mathcal{S}'$ also allows us to approach fluctuation problems where the limiting processes would no longer take values in the finite measures.

References

[1] F. Avram and M.S. Taqqu. Probability bounds for M-Skorohod oscillations. *Stochastic Processes and their Applications*, 33(1):63–72, 1989. MR-1027108

[2] F. Avram and M.S. Taqqu. Weak convergence of sums of moving averages in the α-stable domain of attraction. *The Annals of Probability*, 20(1):483–503, 1992. MR-1143432

[3] B. Basrak and D. Krizmanić. A Multivariate Functional Limit Theorem in Weak M_1 Topology. *Journal of Theoretical Probability*, 28:119–136, 2015. MR-3320962

[4] J.J. Becnel and A.N. Sengupta. Nuclear space facts, strange and plain. `http://www.math.lsu.ed/~sengupta/papers/BecSenNuclear.pdf`, 2009.

[5] G. Ben Arous and J. Černý. Scaling limit for trap models on $\mathbb{Z}^d$. *The Annals of Probability*, 35(6):2356–2384, 2007. MR-2353391

[6] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication. MR-1700749

[7] D. Blount and M.A. Kouritzin. On convergence determining and separating classes of functions. *Stochastic Processes and their Applications*, 120:1898–1907, 2010. MR-2673979

[8] N. Bush, B.M. Hambly, H. Haworth, L. Jin, and C. Reisinger. Stochastic Evolution Equations in Portfolio Credit Modelling. *SIAM Journal of Financial Mathematics*, 2(1):627–664, 2011. MR-2836495

[9] F. Delarue, J. Inglis, S. Rubenthaler, and E. Tanré. Particle systems with singular mean-field self-excitation. Application to neuronal networks. *Stochastic Processes and their Applications*, 125(6):2451–2492, 2015. MR-3322871

[10] K. R. Duffy, C. Macci, and G.L. Torrisi. Sample path large deviations for order statistics. *Journal of Applied Probability*, 48(1):238–257, 2011. MR-2809898

[11] S. Ethier and T. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley, 1986. MR-0838085

[12] J.M. Harrison and R.J. William. A multiclasss closed queueing network with unconventional heavy traffic behavior. *The Annals of Applied Probability*, 6(1):1–47, 1996. MR-1389830

[13] R. Holley and D.W. Stroock. Central limit phenomena of various interacting systems. *The Annals of Mathematics*, 110(2):333–393, 1979. MR-0549491

[14] K. Itô. Distribution-valued processes arising from independent Brownian motions. *Mathematische Zeitschrift*, 182(1):17–33, 1983. MR-0686883

[15] S. Iyengar. Hitting lines with two-dimensional Brownian motion. *SIAM Journal of Applied Mathematics*, 45(6):983–989, 1985. MR-0813460

[16] A. Jakubowski. On the Skorokhod topology. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 22(3):263–285, 1986. MR-0871083

[17] G. Kallianpur and J. Xiong. *Stochastic Differential Equations in Infinite Dimensional Spaces*. Institute of Mathematical Statistics: Lecture notes, monograph series. Institute of Mathematical Statistics, 1995. MR-1465436

[18] O. Kella and W. Whitt. Diffusion Approximations for Queues of Server Vacations. *Advances in Applied Probability*, 22(3):706–729, 1990. MR-1066971

[19] M.A. Kouritzin. On tightness of probability measures on Skorokhod spaces. `http://hdl.handle.net/10402/era.37299`, 2014.

[20] D. Krizmanić. Weak convergence of partial maxima processes in the M1 topology. *Extremes*, 17(3):447–465, 2014. MR-3252821

[21] A. Mandelbaum and W.A. Massey. Strong approximations for time-dependent queues. *Mathematics of Operations Research*, 20(1):33–64, 1995. MR-1320446

[22] A. Metzler. On the first passage problem for correlated Brownian motion. *Statistics & Probability Letters*, 80(5–6):277–284, 2010. MR-2593563

[23] I. Mitoma. On the sample continuity of $\mathcal{S}'$-processes. *Journal of the Mathematical Society of Japan*, 35(4):629–636, 1983. MR-0714465

[24] I. Mitoma. Tightness of probabilities on $C\left([0,1]\,;\mathcal{S}'\right)$ and $D\left([0,1]\,;\mathcal{S}'\right)$. *The Annals of Probability*, 11(4):989–999, 1983. MR-0714961

[25] G. Pang and W. Whitt. Continuity of a queueing integral representation in the M_1 topology. *The Annals of Applied Probability*, 20(1):214–237, 2010. MR-2582647

[26] I. Pavlyukevich and M. Riedle. Non-standard skorokhod convergence of Lévy-driven convolution integrals in Hilbert spaces. *Stochastic Analysis and Applications*, 33:271–305, 2015. MR-3305470

[27] S. Resnick and E. van der Berg. Weak convergence of high-speed network traffic models. *Journal of Applied Probability*, 37:575–597, 2000. MR-1781014

[28] A.V. Skorokhod. Limit Theorems for Stochastic Processes. *Theory of Probability and its Applications*, 1(3):261–290, 1956. MR-0084897

[29] O.G. Smolyanov and S.V. Fomin. Measures on linear topological spaces. *Russian Mathematical Surveys*, 31(4):1–53, 1976.

[30] A.-S. Sznitman. Topics in propagation of chaos. In Paul-Louis Hennequin, editor, *Ecole d'Eté de Probabilités de Saint-Flour XIX – 1989*, volume 1464 of *Lecture Notes in Mathematics*, chapter 3, pages 165–251. Springer Berlin Heidelberg, 1991. MR-1108185

[31] W. Whitt. Weak Convergence of First Passage Time Processes. *Journal of Applied Probability*, 8(2):417–422, 1971. MR-0307335

[32] W. Whitt. Limits for cumulative input processes to queues. *Probability in Engineering and Informational Sciences*, 14:123–150, 2000. MR-1757175

[33] W. Whitt. The Reflection Map with Discontinuities. *Mathematics of Operations Research*, 26(3):447–484, 2001. MR-1849880

[34] W. Whitt. *Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues*. Springer Series in Operations Research and Financial Engineering. Springer, 2002. MR-1876437

[35] M.J. Wichura. Functional laws of the iterated logarithm for the partial sums of i.i.d. random variables in the domain of attraction of a completely asymmetric stable law. *The Annals of Probability*, 2(6):1108–1138, 1974. MR-0358950

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