The Annals of Probability
2015, Vol. 43, No. 4, 1777–1822
DOI: 10.1214/14-AOP919

ON A FUNCTIONAL CONTRACTION METHOD

BY RALPH NEININGER AND HENNING SULZBACH

Goethe University Frankfurt

Methods for proving functional limit laws are developed for sequences of stochastic processes which allow a recursive distributional decomposition either in time or space. Our approach is an extension of the so-called contraction method to the space $\mathcal{C}[0,1]$ of continuous functions endowed with uniform topology and the space $\mathcal{D}[0,1]$ of càdlàg functions with the Skorokhod topology. The contraction method originated from the probabilistic analysis of algorithms and random trees where characteristics satisfy natural distributional recurrences. It is based on stochastic fixed-point equations, where probability metrics can be used to obtain contraction properties and allow the application of Banach's fixed-point theorem. We develop the use of the Zolotarev metrics on the spaces $\mathcal{C}[0,1]$ and $\mathcal{D}[0,1]$ in this context. Applications are given, in particular, a short proof of Donsker's functional limit theorem is derived and recurrences arising in the probabilistic analysis of algorithms are discussed.

1. Introduction.

The contraction method is an approach for proving convergence in distribution for sequences of random variables which satisfy recurrence relations in distribution. Such recurrence relations for a sequence $(Y_n)_{n\geq 0}$ are often of the form

$$Y_n \stackrel{d}{=} \sum_{r=1}^{K} A_r(n) Y^{(r)}_{I^{(n)}_r} + b(n), \qquad n \geq n_0, \tag{1}$$

where $\stackrel{d}{=}$ denotes that the left-hand side and right-hand side are identically distributed, and $(Y^{(r)}_j)_{j\geq 0}$ have the same distribution as $(Y_n)_{n\geq 0}$ for all $r = 1, \ldots, K$, where $K \geq 1$ and $n_0 \geq 0$ are fixed integers. Moreover, $I^{(n)} = (I^{(n)}_1, \ldots, I^{(n)}_K)$ is a vector of random integers in $\{0, \ldots, n\}$. The basic independence assumption that fixes the distribution of the right-hand side is that $(Y^{(1)}_j)_{j\geq 0}, \ldots, (Y^{(K)}_j)_{j\geq 0}$ and $(A_1(n), \ldots, A_K(n), b(n), I^{(n)})$ are independent. Note, however, that dependencies between the coefficients $A_r(n)$, $b(n)$ and the integers $I^{(n)}_r$ are allowed.

Recurrences of the form (1) come up in diverse fields, for example, in the study of random trees, the probabilistic analysis of recursive algorithms, in branching

Received April 2013; revised February 2014.
MSC2010 subject classifications. Primary 60F17, 68Q25; secondary 60G18, 60C05.
Key words and phrases. Functional limit theorem, contraction method, recursive distributional equation, Zolotarev metric, Donsker's invariance principle.

processes, in the context of random fractals and in models from stochastic geometry where a recursive decomposition can be found, as well as in information and coding theory. For surveys of such occurrences, see [21, 22, 29]. In some applications, one may need K to depend on n or the case $K = \infty$, where generalizations of the results for our case of fixed K can be stated; cf. [22], Section 4.3, for such extensions in the finite-dimensional case.

The sequence $(Y_n)_{n\geq 0}$ satisfying (1) often is a sequence of real random variables with real coefficients $A_r(n)$, $b(n)$. However, the same recurrence appears also for sequences of random vectors $(Y_n)_{n\geq 0}$ in $\mathbb{R}^d$. Then the $A_r(n)$ are random linear maps from $\mathbb{R}^d$ to $\mathbb{R}^d$ and $b(n)$ is a random vector in $\mathbb{R}^d$. We will also review below work that considered random sequences $(Y_n)_{n\geq 0}$ into a separable Hilbert space satisfying (1) where $A_r(n)$ become random linear operators on the space and $b(n)$ a random vector in the Hilbert space. In the present work, we develop a limit theory for such sequences in separable Banach spaces, where our main applications are first to the space $\mathcal{C}[0, 1]$ endowed with the uniform topology. Secondly, although not a Banach space, we will also be able to cover the space $\mathcal{D}[0, 1]$ equipped with the Skorokhod topology. Hence, we consider sequences $(Y_n)_{n\geq 0}$ of stochastic processes with state space $\mathbb{R}$ and time parameter $t \in [0, 1]$ with continuous, respectively, cádlág paths and are interested in conditions that together with (1) allow to deduce functional limit theorems for rescaled versions of $(Y_n)_{n\geq 0}$.

For functions $f \in \mathcal{C}[0, 1]$ or $f \in \mathcal{D}[0, 1]$, we denote the uniform norm by

$$\|f\|_\infty := \sup_{x\in[0,1]} \bigl|f(x)\bigr|.$$

For functions $f, g \in \mathcal{D}[0, 1]$, the Skorokhod distance $d_{\mathrm{sk}}(f, g)$ is used; see Section 2.2.

The rescaling of the process $(Y_n)_{n\geq 0}$ can be done by centering and normalization by the order of the standard deviation in case moments of sufficient order are available. Subsequently, we assume that the scaling has already been done and we denote the scaled process by $(X_n)_{n\geq 0}$. Note that affine scalings of the Y_n implies that the sequence $(X_n)_{n\geq 0}$ also does satisfy a recurrence of type (1), where only the coefficients are changed:

$$X_n \stackrel{d}{=} \sum_{r=1}^{K} A_r^{(n)} X^{(r)}_{I_r^{(n)}} + b^{(n)}, \qquad n \geq n_0 \tag{2}$$

with conditions on identical distributions and independence similar to recurrence (1). The coefficients $A_r^{(n)}$ and $b^{(n)}$ in the modified recurrence (2) are typically directly computable from the original coefficients $A_r(n)$, $b(n)$ and the scaling used; see, for example, for the case of random vectors in $\mathbb{R}^d$, [22], equation (4). Subsequently, we consider equations of type (2) together with assumptions on the moments of X_n which in applications have to be obtained by an appropriate scaling.

For the asymptotic distributional analysis of sequences $(X_n)_{n\geq 0}$ satisfying (2), the so-called contraction method has become a powerful tool. In the seminal paper [26], Rösler introduced this methodology for deriving a limit law for a special instant of this equation that arises in the analysis of the complexity of the Quicksort algorithm. In the framework of the contraction method, first one derives limits of the coefficients $A_r^{(n)}$, $b^{(n)}$,

$$A_r^{(n)} \to A_r, \qquad b^{(n)} \to b \qquad (n \to \infty) \tag{3}$$

in an appropriate sense. If with $n \to \infty$, also the $I_r^{(n)}$ become large and it is plausible that the quantities X_n converge, say to a random variable X; then, by letting formally $n \to \infty$, equation (2) turns into

$$X \stackrel{d}{=} \sum_{r=1}^{K} A_r X^{(r)} + b \tag{4}$$

with $X^{(1)}, \dots, X^{(K)}$ distributed as X and $X^{(1)}, \dots, X^{(K)}$, $(A_1, \dots, A_k, b)$ independent. Hence, one can use the distributional fixed-point equation (4) to characterize the limit distribution $\mathcal{L}(X)$. The idea from Rösler [26] to formalize such an approach and to derive at least weak convergence $X_n \to X$ consists of first using the right-hand side of (4) to define a map as follows: if X_n are B-valued random variables, denote by $\mathcal{M}(B)$ the space of all probability measures on B and

$$T : \mathcal{M}(B) \to \mathcal{M}(B), \tag{5}$$

$$T(\mu) = \mathcal{L}\left(\sum_{r=1}^{K} A_r Z^{(r)} + b\right), \tag{6}$$

where $(A_1, \dots, A_K, b)$, $Z^{(1)}, \dots, Z^{(K)}$ are independent and $Z^{(1)}, \dots, Z^{(K)}$ have distribution μ. Then a random variable X solves (4) if and only if its distribution $\mathcal{L}(X)$ is a fixed point of the map T. To obtain fixed points of T appropriate subspaces of $\mathcal{M}(B)$ are endowed with a complete metric, such that the restriction of T becomes a contraction. Then Banach's fixed-point theorem yields a (in the subspace) unique fixed point of T and one can as well use the metric to also derive convergence of $\mathcal{L}(X_n)$ to $\mathcal{L}(X)$ in this metric. If the metric is also strong enough to imply weak convergence, one has obtained the desired limit law $X_n \to X$.

This approach has been established and applied to a couple of examples in Rösler [26, 27] and Rachev and Rüschendorf [25]. In the latter paper also the flexibility of the approach by using various probability metrics has been demonstrated. Later on general convergence theorems have been derived stating conditions under which convergence of the coefficients of the form (3) together with a contraction property of the map (5) implies convergence in distribution $X_n \to X$. For random variables in $\mathbb{R}$ with the minimal ℓ_2 metric, see Rösler [28], and Neininger [20] for $\mathbb{R}^d$ with the same metric. For a more widely applicable framework for random variables in $\mathbb{R}^d$, see Neininger and Rüschendorf [22], where in particular various problems with normal limit laws could be solved which seem to be beyond the scope of the minimal ℓ_p metric; see also [23]. An extension of these theorems to continuous time, that is, to processes $(X_t)_{t\geq 0}$ satisfying recurrences similar to (2) was given in Janson and Neininger [17].

For the case of random variables in a separable Hilbert space leading to functional limit laws, general limit theorems for recurrences (1) have been developed in Drmota, Janson and Neininger [12]. The main application there was a functional limit law for the profile of random trees which, via a certain encoding of the profile, led to random variables in the Bergman space of square integrable analytic functions on a domain in the complex plane. In Eickmeyer and Rüschendorf [13], general limit theorems for recurrences in $\mathcal{D}[0,1]$ under the L_p-topology were developed. Note that the uniform topology for $\mathcal{C}[0,1]$ and the Skorokhod topology for $\mathcal{D}[0,1]$ considered in the present paper are finer than the L_p-topology. In $\mathcal{C}[0,1]$, the uniform topology provides more continuous functionals such as the supremum $f \mapsto \sup_{t\in[0,1]} f(t)$ or projections $f \mapsto f(s_1,\dots,s_k)$, for fixed $s_1,\dots,s_k \in [0,1]$, to which the continuous mapping theorem can be applied. In $\mathcal{D}[0,1]$, these functionals are also appropriate for the continuous mapping theorem if the limit random variable has continuous sample paths.

Besides the minimal ℓ_p metrics the probability metrics that have proved useful in most of the papers mentioned above is the family of Zolotarev metrics ζ_s being reviewed and further developed here in Section 2. All generalizations from $\mathbb{R}$ via $\mathbb{R}^d$ to separable Hilbert spaces are based on the fact that convergence in ζ_s implies weak convergence; see Section 2. However, for Banach spaces this is not true in general. Counterexamples have been reported in Bentkus and Rachkauskas [4], sketched here in Section 2.1. Also completeness of the ζ_s metrics on appropriate subspaces of $\mathcal{M}(B)$ is only known for the case of separable Hilbert spaces; see [12], Theorem 5.1.

Our study of the spaces $(\mathcal{C}[0,1], \|\cdot\|_\infty)$ and $(\mathcal{D}[0,1], d_{\mathrm{sk}})$ is also based on the Zolotarev metrics ζ_s. Hence, we mainly have to deal with implications that can be drawn from convergence in the ζ_s metrics as well as with the lack of knowledge about completeness of ζ_s. In Section 2.3, implications of convergence in the Zolotarev metric are discussed together with additional conditions that enable to deduce in general weak convergence from convergence in ζ_s. A key ingredient here is a technique developed in Barbour [2] in the context of Stein's method; see also Barbour and Janson [3]. We also obtain criteria for the uniform integrability of $\{\|X_n\|_\infty^s | n \geq 0\}$ for $0 \leq s \leq 3$ in the presence of convergence in the Zolotarev metric. This enables in applications as well to obtain moments convergence of the sup-functional.

In Section 3, we give general convergence theorems in the framework of the contraction method first for a general separable Banach space and then apply and refine this to the space $(\mathcal{C}[0,1], \|\cdot\|_\infty)$ and develop a technique to also apply this to the metric space $(\mathcal{D}[0,1], d_{\mathrm{sk}})$. In particular, based on Janson and Kaijser [16], we give a criterion for the finiteness of the Zolotarev metric on appropriate subspaces that can easily be checked in applications.

To compensate for the lack of knowledge about completeness of the ζ_s metrics, we need to assume that the map T in (5) has a fixed point in an appropriate subspace of $\mathcal{M}(\mathcal{C}[0,1])$ and $\mathcal{M}(\mathcal{D}[0,1])$, respectively. In applications, one may verify this existence of a fixed point either by guessing one successfully: in the application of our framework to Donsker's functional limit theorem in Section 4.1, the Wiener measure can easily be guessed and be seen to be the fixed point of the map T coming up there. Alternatively, in general the existence of a fixed point may arise from infinite iteration of the map T: applied to some probability measure, such an iteration has a series representation for which one may be able to show that it is the desired fixed point. This path is being taken in an application of our framework outlined in Section 4.2.

In Section 4.1, we apply our functional contraction method to derive a short proof of Donsker's functional limit theorem. This does not require the full generality of our setting but illustrates how self-similarities can easily been exploited with this approach. The application in Section 4.2 is on the asymptotic study of fundamental complexities in computer science. Here, the full generality of our approach is needed to obtain a functional limit law. We highlight and discuss the use of our conditions (C1)–(C5) formulated in Section 3 on the recurrence (2) at this example. Details on the verification of the conditions are contained in Broutin, Neininger and Sulzbach [6] where, based on the functional limit law, also various long open standing problems on the complexities in computer science are solved.

2. The Zolotarev metric. Let $(B, \|\cdot\|)$ be a real Banach space and $\mathcal{B}$ its Borel σ-algebra. In Section 2.1, we assume that the norm on B induces a separable topology. We denote by $\mathcal{M}(B)$ the set of all probability measures on $(B, \mathcal{B})$. First, we introduce the Zolotarev metric ζ_s and collect some of its basic properties, mainly covered in [32, 33]. In the second subsection, we define our use of the Zolotarev metrics on the metric space $(\mathcal{D}[0,1], d_{\mathrm{sk}})$. Although not a Banach space, we will be able to declare the Zolotarev metrics ζ_s on $(\mathcal{D}[0,1], d_{\mathrm{sk}})$ using the notion of differentiability of functions $\mathcal{D}[0,1] \to \mathbb{R}$ induced by the supremum norm on $\mathcal{D}[0,1]$. We also comment in Remarks 6 and 7 on delicate measurability issues for the nonseparable Banach space $(B, \|\cdot\|) = (\mathcal{D}[0,1], \|\cdot\|_\infty)$ and the realm of our methodology when working with the coarser (separable) topology on $\mathcal{D}[0,1]$ induced by the Skorokhod metric. In the third subsection, conditions that allow to conclude from convergence in ζ_s to weak convergence are studied for the case $(B, \|\cdot\|) = (\mathcal{C}[0,1], \|\cdot\|_\infty)$ as well as for the case $(\mathcal{D}[0,1], d_{\mathrm{sk}})$. We also discuss further implications from ζ_s-convergence in these two spaces as well as criteria for finiteness of ζ_s. Additional material to the content of this section can be found in the second author's dissertation [31], Chapter 2.

2.1. *Definition and basic properties.* For functions $f : B \to \mathbb{R}$, which are Fréchet differentiable, the derivative of f at a point x is denoted by $Df(x)$. Note that $Df(x)$ is an element of the space $L(B, \mathbb{R})$ of continuous linear forms on B. We also consider higher order derivatives, where $D^m f(x)$ denotes the mth derivative of f at a point x. Thus, $D^m f(x)$ is a continuous m-linear (or multilinear) form on B. The space of continuous multilinear forms $g : B^m \to \mathbb{R}$ is equipped with the norm

$$\|g\| = \sup_{\|h_1\| \le 1, \dots, \|h_m\| \le 1} \left| g(h_1, \dots, h_m) \right|.$$

For a comprehensive account on differentiability in Banach spaces, we refer to Cartan [7]. Subsequently, $s > 0$ is fixed and for $m := \lceil s \rceil - 1$ and $\alpha := s - m$ we define

$$\mathcal{F}_s = \left\{ f : B \to \mathbb{R} : \left\| D^m f(x) - D^m f(y) \right\| \le \|x - y\|^\alpha,\ \forall x, y \in B \right\}. \tag{7}$$

For $\mu, \nu \in \mathcal{M}(B)$, the Zolotarev distance between μ and ν is defined by

$$\zeta_s(\mu, \nu) = \sup_{f \in \mathcal{F}_s} \left| \mathbf{E}\left[f(X) - f(Y) \right] \right|, \tag{8}$$

where X and Y are B-valued random variables with $\mathcal{L}(X) = \mu$ and $\mathcal{L}(Y) = \nu$. Here, $\mathcal{L}(X)$ denotes the distribution of the random variable X. The expression in (8) does not need to be finite or even well defined. However, we have $\zeta_s(\mu, \nu) < \infty$ if

$$\int \|x\|^s \, d\mu(x), \int \|x\|^s \, d\nu(x) < \infty \tag{9}$$

and

$$\int f(x, \dots, x) \, d\mu(x) = \int f(x, \dots, x) \, d\nu(x) \tag{10}$$

for any bounded k-linear form f on B and any $1 \le k \le m$. For random variables X, Y in B, we use the abbreviation $\zeta_s(X, Y) := \zeta_s(\mathcal{L}(X), \mathcal{L}(Y))$. Finiteness of $\zeta_s(X, Y)$ in $\mathbb{R}^d$ fails to hold if X and Y do not have the same mixed moments up to order m. The assumption on the finite absolute moment of order s can be relaxed slightly; see Theorem 4 in [34].

We denote

$$\mathcal{M}_s(B) := \left\{ \mu \in \mathcal{M}(B) \middle| \int \|x\|^s \, d\mu(x) < \infty \right\}$$

and for all $\nu \in \mathcal{M}_s(B)$ denote

$$\mathcal{M}_s(\nu) := \left\{ \mu \in \mathcal{M}_s(B) | \mu \text{ and } \nu \text{ satisfy (10)} \right\}.$$

Then ζ_s is a metric on the space $\mathcal{M}_s(\nu)$ for any $\nu \in \mathcal{M}_s(B)$; see [35], Remark 1, page 198.

A crucial property of ζ_s in the context of recursive decompositions of stochastic processes is the following lemma; see Theorem 3 in [34]. A short proof is given for the reader's convenience.

LEMMA 1. *Let B' be a Banach space and $g : B \to B'$ a linear and continuous operator. Then we have*

$$\zeta_s\big(g(X), g(Y)\big) \le \|g\|^s \zeta_s(X, Y), \qquad \mathcal{L}(X), \mathcal{L}(Y) \in \mathcal{M}_s(\nu).$$

Here, $\|g\|$ *denotes the operator norm of* g, *that is*, $\|g\| = \sup_{x \in B, \|x\| \le 1} \|g(x)\|$.

PROOF. Note that g is also bounded. It suffices to show that

$$\{\|g\|^{-s} f \circ g : f \in \mathcal{F}_s'\} \subseteq \mathcal{F}_s,$$

where $\mathcal{F}_s'$ is defined analogously to $\mathcal{F}_s$ in B'. Let $f \in \mathcal{F}_s$ and $\eta := \|g\|^{-s} f \circ g$. Then η is m-times continuously differentiable and we have $D^m \eta(x) = \|g\|^{-s}(D^m(f(g(x))) \circ g^{\otimes m}$ for $x \in B$. Here, $g^{\otimes m} : B^m \to (B')^m$ denotes the mapping $g^{\otimes m}(h_1, \dots, h_m) = (g(h_1), \dots, g(h_m))$. This implies

$$\begin{aligned}\|D^m \eta(x) - D^m \eta(y)\| &= \|g\|^{-s} \big\|\big(D^m f\big(g(x)\big)\big) \circ g^{\otimes m} - \big(D^m f\big(g(y)\big)\big) \circ g^{\otimes m}\big\| \\ &\le \|g\|^{-\alpha} \big\|g(x) - g(y)\big\|^\alpha \\ &= \|g\|^{-\alpha} \big\|g(x - y)\big\|^\alpha \le \|x - y\|^\alpha.\end{aligned}$$

The assertion follows. $\square$

Another basic property is that ζ_s is $(s, +)$ ideal.

LEMMA 2. *The metric ζ_s is ideal of order s on $\mathcal{M}_s(\nu)$ for any $\nu \in \mathcal{M}_s(B)$, that is*, *we have*

$$\zeta_s(cX, cY) = |c|^s \zeta_s(X, Y),$$
$$\zeta_s(X + Z, Y + Z) \le \zeta_s(X, Y)$$

for any $c \in \mathbb{R} \setminus \{0\}$, $\mathcal{L}(X), \mathcal{L}(Y) \in \mathcal{M}_s(\nu)$ and random variables Z in B, *such that (X, Y) and Z are independent*.

The lemma directly implies

$$\zeta_s(X_1 + X_2, Y_1 + Y_2) \le \zeta_s(X_1, Y_1) + \zeta_s(X_2, Y_2) \tag{11}$$

for $\mathcal{L}(X_1), \mathcal{L}(Y_1) \in \mathcal{M}_s(\nu_1)$ and $\mathcal{L}(X_2), \mathcal{L}(Y_2) \in \mathcal{M}_s(\nu_2)$ with arbitrary $\nu_1, \nu_2 \in \mathcal{M}_s(B)$ such that (X_1, Y_1) and (X_2, Y_2) are independent.

We want to give a result similar to Lemma 1 where the linear operator may also be random itself. We focus on the case that B' either equals B or $\mathbb{R}$ where an extension to $\mathbb{R}^d$ for $d > 1$ is straightforward. Let B^* be the topological dual of B and $\widehat{B}$ be the space of all continuous linear maps from B to B. Endowed with the operator norms

$$\|f\|_{\mathrm{op}} = \sup_{x \in B, \|x\| \le 1} |f(x)|, \qquad \|f\|_{\mathrm{op}} = \sup_{x \in B, \|x\| \le 1} \|f(x)\|,$$

both spaces, B^* and $\widehat{B}$, respectively, are Banach spaces. However, these spaces are typically nonseparable, hence not suitable for our purposes of measurability. Therefore, we will equip them with smaller σ-algebras. Similar to the use of weak-* convergence, let $\mathcal{B}^*$ be the σ-algebra on B^* that is generated by all continuous (with respect to $\|\cdot\|_{\mathrm{op}}$) linear forms φ on B^* (i.e., elements of the bidual B^{**}) of the form $\varphi(a) = a(x)$ for some $x \in B$. Note that the set of these continuous linear forms coincides with the bidual B^{**} if and only if B is reflexive, a property that is not satisfied in our applications. We move on to $\widehat{B}$ and define $\widehat{\mathcal{B}}$ to be the σ-algebra generated by all continuous (with respect to $\|\cdot\|_{\mathrm{op}}$) linear maps ψ from $\widehat{B}$ to B of the form $\psi(a) = a(x)$ for some $x \in B$. By Pettis' theorem, we have $\mathcal{B} = \sigma(\ell \in B^*)$. Hence, if $S \subseteq B^*$ with $\mathcal{B} = \sigma(\ell \in S)$, then $\widehat{\mathcal{B}}$ is also generated by the continuous linear forms ϱ on $\widehat{B}$ that can be written as $\varrho(a) = \ell(a(x))$ for $\ell \in S$ and $x \in B$.

Using the separability of B, it is now easy to see that the norm-functionals $B^* \to \mathbb{R}$, $f \mapsto \|f\|_{\mathrm{op}}$ and $\widehat{B} \to \mathbb{R}$, $f \mapsto \|f\|_{\mathrm{op}}$ are $\mathcal{B}^*$–$\mathcal{B}(\mathbb{R})$ measurable and $\widehat{\mathcal{B}}$–$\mathcal{B}(\mathbb{R})$ measurable, respectively.

DEFINITION 3. By a *random continuous linear form on B*, we denote any random variable with values in $(B^*, \mathcal{B}^*)$. Analogously, *random continuous linear operators on B* are random variables with values in $(\widehat{B}, \widehat{\mathcal{B}})$.

Note that the definition of the σ-algebras $\mathcal{B}^*$ and $\widehat{\mathcal{B}}$ implies in particular that for any $a \in B^*$ or $a \in \widehat{B}$, $x \in B$, random continuous linear form or operator A and random variable X in B, we have that the compositions $a(X)$, $A(x)$ and $A(X)$ are again random variables. The latter property follows from measurability of the map $(a, x) \mapsto a(x)$ with respect to $(\mathcal{B}^* \otimes \mathcal{B})$–$\mathcal{B}(\mathbb{R})$ and $(\widehat{\mathcal{B}} \otimes \mathcal{B})$–$\mathcal{B}$, respectively. In the case of the dual space, this follows as for any $r \in \mathbb{R}$ we have

$$\begin{aligned}&\{(a, x) \in B^* \times B : a(x) < r\}\\&\quad= \bigcup_{k\ge 1} \bigcup_{m\ge 1} \bigcap_{n\ge m} \bigcup_{i\ge 1} \{a \in B^* : a(e_i) < r - 1/k\} \times \{x \in B : \|x - e_i\| < 1/n\},\end{aligned}$$

where $\{e_i | i \ge 1\}$ denotes a countable dense subset of B; the case $\widehat{B}$ being analogous.

The following lemma follows from Lemma 1 by conditioning.

LEMMA 4. *Let $\mathcal{L}(X), \mathcal{L}(Y) \in \mathcal{M}_s(\nu)$ for some $\nu \in \mathcal{M}_s(B)$. Then, for any random linear continuous form or operator A with $\mathbf{E}[\|A\|_{\mathrm{op}}^s] < \infty$ independent of X and Y, we have*

$$\zeta_s\bigl(A(X), A(Y)\bigr) \le \mathbf{E}\bigl[\|A\|_{\mathrm{op}}^s\bigr]\zeta_s(X, Y).$$

Zolotarev gave upper and lower bounds for ζ_s, most of them being valid if more structure on B is assumed. Subsequently, only an upper bound in terms of the minimal ℓ_p metric is needed. For $p > 0$ and $\mu, \nu \in \mathcal{M}_p(B)$, the minimal ℓ_p distance between μ and ν is defined by

$$\ell_p(\mu, \nu) = \inf \mathbf{E}\big[\|X - Y\|^p\big]^{(1/p)\wedge 1},$$

where the infimum is taken over all common distributions $\mathcal{L}(X, Y)$ with marginals $\mathcal{L}(X) = \mu$ and $\mathcal{L}(Y) = \nu$. We abbreviate $\ell_p(X, Y) := \ell_p(\mathcal{L}(X), \mathcal{L}(Y))$.

The next lemma gives an upper bound of ζ_s in terms of ℓ_s where the first statement follows from the Kantorovich–Rubinstein theorem and the second essentially coincides with Lemma 5.7 in [12].

LEMMA 5. *Let $\mathcal{L}(X), \mathcal{L}(Y) \in \mathcal{M}_s(\nu)$ for some $\nu \in \mathcal{M}_s(B)$ with B separable. If $s \leq 1$ then*

$$\zeta_s(X, Y) = \ell_s(X, Y). \tag{12}$$

If $s > 1$ then

$$\zeta_s(X, Y) \leq \big(\mathbf{E}\big[\|X\|^s\big]^{1-1/s} + \mathbf{E}\big[\|Y\|^s\big]^{1-1/s}\big)\ell_s(X, Y).$$

If X_n, X are real-valued random variables, $n \geq 1$, then $\zeta_s(X_n, X) \to 0$ implies convergence of absolute moments of order up to s since there is a constant $C_s > 0$ such that the function $x \mapsto C_s|x|^s$ is an element of $\mathcal{F}_s$, hence $|\mathbf{E}[|X_n|^s - |X|^s]| \leq C_s^{-1}\zeta_s(X_n, X)$.

We proceed with the fundamental question of how convergence in the ζ_s distance relates to weak convergence on B. By the first statement of the previous lemma, or more elementary, by the proof of the Portmanteau lemma [5], Theorem 2.1(ii)–(iii), one obtains that for $0 < s \leq 1$ convergence in the ζ_s metric implies weak convergence; see also [12], page 300.

If B is a separable Hilbert space, then for any $s > 0$ convergence in the ζ_s metric implies weak convergence. This was first proved by Giné and León in [15], see also Theorem 5.1 in [12]. In infinite-dimensional Banach spaces convergence in the ζ_s metric does not need to imply weak convergence: for any probability distribution μ on $B = \mathcal{C}[0, 1]$ with zero mean and $\int \|x\|_\infty^s \, d\mu(s) < \infty$ for some $s > 2$, that is pre-Gaussian, that is, there exists a Gaussian measure ν on $\mathcal{C}[0, 1]$ with zero mean and the same covariance as μ, one has ζ_s-convergence of a rescaled sum of independent random variables with distribution μ toward ν; see inequality (48) in [32]. However, pre-Gaussian probability distributions supported by a bounded subset of $\mathcal{C}[0, 1]$ that do not satisfy the central limit theorem can be found in [30]. For the central limit theorem in Banach spaces, see [18]. Note that convergence with respect to ζ_s implies convergence of the characteristic functions, hence $\zeta_s(X_n, X) \to 0$ implies that $\mathcal{L}(X)$ is the only possible accumulation point of $(\mathcal{L}(X_n))_{n\geq 0}$ in the weak topology.

2.2. *The Zolotarev metric on* $(\mathcal{D}[0,1], d_{\mathrm{sk}})$. In this section, we discuss our use of the Zolotarev metric on the metric space $(\mathcal{D}[0,1], d_{\mathrm{sk}})$ of càdlàg functions on $[0,1]$ endowed with the Skorokhod metric defined by

$$\begin{aligned} &d_{\mathrm{sk}}(f,g) \\ &\quad = \inf\{\varepsilon > 0 \mid \max\{|f(t) - g(\tau(t))|, |\tau(t) - t|\} < \varepsilon \text{ for all } t \in [0,1] \\ &\qquad\qquad \text{for some monotonically increasing and bijective } \tau : [0,1] \to [0,1]\}. \end{aligned}$$

The Borel σ-algebra of the induced topology is denoted by $\mathcal{B}_{\mathrm{sk}}$. For a general introduction to this space, see Billingsley [5], Chapter 3. In particular, $(\mathcal{D}[0,1], d_{\mathrm{sk}})$ is a Polish space, $\mathcal{B}_{\mathrm{sk}}$ coincides with the σ-algebra generated by the finite-dimensional projections, the σ-algebra generated by the open spheres (with respect to the uniform metric) and the σ-algebra generated by all norm-continuous linear forms on $\mathcal{D}[0,1]$; see [24], Theorem 3. Subsequently, norm on $\mathcal{D}[0,1]$ will always refer to the uniform norm $\|\cdot\|_\infty$. Moreover, the norm function $\mathcal{D}[0,1] \to \mathbb{R}$, $f \mapsto \|f\|_\infty$ is $\mathcal{B}_{\mathrm{sk}}$–$\mathcal{B}(\mathbb{R})$ measurable. By Theorem 2, respectively, Theorem 4, in [24], any norm-continuous linear form on $\mathcal{D}[0,1]$ is $\mathcal{B}_{\mathrm{sk}}$–$\mathcal{B}(\mathbb{R})$ measurable and any norm-continuous linear map from $\mathcal{D}[0,1]$ to $\mathcal{D}[0,1]$ is $\mathcal{B}_{\mathrm{sk}}$–$\mathcal{B}_{\mathrm{sk}}$ measurable. Recently, Janson and Kaijser [16], Theorem 15.8, generalized the latter result and proved that any norm-continuous k-linear form on $\mathcal{D}[0,1]$ is $(\mathcal{B}_{\mathrm{sk}})^{\otimes k}$–$\mathcal{B}(\mathbb{R})$ measurable. We do, however, not know whether $\mathcal{F}_s$ defined in (7) based on the uniform norm on $\mathcal{D}[0,1]$ is a subset of the $\mathcal{B}_{\mathrm{sk}}$–$\mathcal{B}(\mathbb{R})$ measurable functions. Hence, we denote the $\mathcal{B}_{\mathrm{sk}}$–$\mathcal{B}(\mathbb{R})$ measurable functions by $\mathcal{E}$ and define the Zolotarev metrics analogously to (8) by

$$\zeta_s(\mu,\nu) = \sup_{f \in \mathcal{F}_s \cap \mathcal{E}} |\mathbf{E}[f(X) - f(Y)]|,$$

where X and Y are $(\mathcal{D}[0,1], d_{\mathrm{sk}})$-valued random variables with $\mathcal{L}(X) = \mu$ and $\mathcal{L}(Y) = \nu$.

We denote by $\mathcal{M}_s(\mathcal{D}[0,1])$ the set of probability distributions μ on $\mathcal{D}[0,1]$ with $\int \|x\|_\infty^s \, d\mu(x) < \infty$ and for $\nu \in \mathcal{M}_s(\mathcal{D}[0,1])$, we define $\mathcal{M}_s(\nu)$ to be the subset of measures μ from $\mathcal{M}_s(\mathcal{D}[0,1])$ satisfying (10). Then ζ_s is a metric on $\mathcal{M}_s(\nu)$ for all $\nu \in \mathcal{M}_s(\mathcal{D}[0,1])$, Lemmas 1 and 2, inequality (11), Lemma 5 where (12) is to be replaced by $\zeta_s(X,Y) \le \ell_s(X,Y)$, and the implication $\zeta_s(X_n, X) \to 0 \Rightarrow X_n \to X$ in distribution if $0 < s \le 1$ remain valid.

The situation becomes more involved concerning random linear forms and operators as defined in Definition 3 in the separable Banach case. Let $\mathcal{D}[0,1]^*$ and $\widehat{\mathcal{D}[0,1]}$ be the dual space, respectively, the space of norm-continuous endomorphisms on $\mathcal{D}[0,1]$ as in the Banach case. For reasons of measurability, we need to restrict to smaller subspaces. Let $\mathcal{D}[0,1]^*_c \subseteq \mathcal{D}[0,1]^*$ be the subset of functions that are additionally continuous with respect to d_{sk}. Analogously, $\widehat{\mathcal{D}[0,1]}_c \subseteq \widehat{\mathcal{D}[0,1]}$ are those endomorphism which are continuous regarded as maps from $(\mathcal{D}[0,1], d_{\mathrm{sk}})$ to $(\mathcal{D}[0,1], d_{\mathrm{sk}})$. We endow $\mathcal{D}[0,1]^*_c$ with the σ-algebra generated by the function $f \mapsto \|f\|_{\mathrm{op}}$ and all elements φ of $\mathcal{D}[0,1]^{**}$ of the form $\varphi(a) = a(x)$ for some $x \in \mathcal{D}[0,1]$. Also the σ-algebra on $\widehat{\mathcal{D}[0,1]}_c$ is generated by the function $f \mapsto \|f\|_{\mathrm{op}}$ and the continuous linear maps $\psi : \widehat{\mathcal{D}[0,1]} \to \mathcal{D}[0,1]$ of the form $\varphi(a) = a(x)$ for some $x \in \mathcal{D}[0,1]$. Under these conditions, we have the same measurability results as in the Banach case and Lemma 4 remains valid.

REMARK 6. Note that we could as well develop the use of the Zolotarev metric together with the contraction method for the Banach space $(\mathcal{D}[0,1], \|\cdot\|_\infty)$. This can be done analogously to the discussion of Sections 2.3 and 3 and in fact would lead to a proof of Donsker's theorem similar to the one given in Section 4.1.1 when replacing the linear interpolation $S^n = (S_t^n)_{t\in[0,1]}$ by a constant (càdlàg) interpolation of the random walk. However, the applicability of such a framework seems to be limited due to measurability problems in the nonseparable space $(\mathcal{D}[0,1], \|\cdot\|_\infty)$: for example, the random function X defined by

$$X_t = \mathbf{1}_{\{t \geq U\}}, \qquad t \in [0,1]$$

with U being uniformly distributed on the unit interval is known to be nonmeasurable with respect to the Borel-σ-algebra on $(\mathcal{D}[0,1], \|\cdot\|_\infty)$. However, we have applications of the functional contraction method developed here in mind on processes with jumps at random times. A typical example in the context of random trees is given in Section 4.2; see also [6]. Hence, in order to even have measurability of the processes considered it requires to work with the coarser Skorokhod topology than the uniform topology and this is our reason for using the Zolotarev metric on $(\mathcal{D}[0,1], d_{\mathrm{sk}})$ instead of $(\mathcal{D}[0,1], \|\cdot\|_\infty)$.

REMARK 7. Although the methodology developed below covers sequences $(X_n)_{n\geq 0}$ of processes with jumps at random times these times will typically need to be the same for all $n \geq n_0$. In particular, sequences of processes with jumps at random times that require a (uniformly small) deformation of the time scale to be aligned cannot be covered by this methodology. The technical reason is that in condition (C1) below (see Section 3) the convergence of the random continuous endomorphisms $\|A_r^{(n)} - A_r\|_s$ is with respect to the operator norm based on the uniform norm which in general does not allow a deformation of the time scale.

2.3. *Weak convergence on* $(\mathcal{C}[0,1], \|\cdot\|_\infty)$ *and* $(\mathcal{D}[0,1], d_{\mathrm{sk}})$. In this subsection, we only consider the spaces $(\mathcal{C}[0,1], \|\cdot\|_\infty)$ and $(\mathcal{D}[0,1], d_{\mathrm{sk}})$.

For random variables $X = (X(t))_{t\in[0,1]}$, $Y = (Y(t))_{t\in[0,1]}$ in $(\mathcal{C}[0,1], \|\cdot\|_\infty)$ with $\zeta_s(X,Y) < \infty$ we have

$$\zeta_s\big((X(t_1),\dots,X(t_k)),(Y(t_1),\dots,Y(t_k))\big) \leq k^{s/2}\zeta_s(X,Y) \tag{13}$$

for all $0 \leq t_1 \leq \cdots \leq t_k \leq 1$. This follows from Lemma 1 using the continuous and linear function $g : \mathcal{C}[0,1] \to \mathbb{R}^k$, $g(f) = (f(t_1),\dots,f(t_k))$ and observing that

$\|g\| = \sqrt{k}$. The bound $\zeta_s((X(t_1), \ldots, X(t_k)), (Y(t_1), \ldots, Y(t_k))) \le \zeta_s(X, Y)$ can be obtained if $\mathbb{R}^k$ is endowed with the max-norm instead of the Euclidean norm. However, no use of this is made here. Hence, we obtain for random variables X_n, X in $(\mathcal{C}[0,1], \|\cdot\|_\infty)$, $n \ge 1$, the implication

$$\zeta_s(X_n, X) \to 0 \quad \Rightarrow \quad X_n \stackrel{\text{f.d.d.}}{\longrightarrow} X.$$

Here, $\stackrel{\text{f.d.d.}}{\longrightarrow}$ denotes weak convergence of all finite-dimensional marginals of the processes. Additionally, if Z is a random variable in $[0,1]$, independent of (X_n) and X, then applying Lemma 4 with the random continuous linear form A defined by $A(f) = f(Z)$ implies

$$\zeta_s\big(X_n(Z), X(Z)\big) \le \mathbf{E}\big[Z^s\big]\zeta_s(X_n, X). \tag{14}$$

In the càdlàg case, that is, $X = (X(t))_{t\in[0,1]}$, $Y = (Y(t))_{t\in[0,1]}$ being random variables in $(\mathcal{D}[0,1], d_{\mathrm{sk}})$ inequality (13) remains true by Lemma 1. (The fact that g is not continuous with respect to the product Skorokhod topology does not cause problems since measurability is sufficient here.) Next, in general, the operator A is no element of $\mathcal{D}[0,1]_c^*$. Hence, we cannot apply Lemma 4 to deduce (14). Nevertheless, by Theorem 2 in [34], the convergence of the characteristic functions of $X_n(t)$ is uniform in t, hence we also have convergence in distribution of $X_n(Z)$ to $X(Z)$. The same argument works for the moments of $X_n(Z)$. We summarize these properties in the following proposition, where $\stackrel{d}{\longrightarrow}$ denotes convergence in distribution.

PROPOSITION 8. *For random variables X_n, X in $(\mathcal{C}[0,1], \|\cdot\|_\infty)$ or $(\mathcal{D}[0,1], d_{\mathrm{sk}})$, $n \ge 1$, with $\zeta_s(X_n, X) \to 0$ for $n \to \infty$ we have*

$$X_n \stackrel{\text{f.d.d.}}{\longrightarrow} X.$$

$\mathcal{L}(X)$ is the only possible accumulation point of $(\mathcal{L}(X_n))_{n\ge1}$ in the weak topology. For all $t \in [0,1]$ we have

$$X_n(t) \stackrel{d}{\longrightarrow} X(t), \qquad \mathbf{E}\big[|X_n(t)|^s\big] \to \mathbf{E}\big[|X(t)|^s\big].$$

For any random variable Z in $[0,1]$ being independent of (X_n) and X, we have

$$\mathbf{E}\big[|X_n(Z)|^s\big] \to \mathbf{E}\big[|X(Z)|^s\big], \qquad X_n(Z) \stackrel{d}{\longrightarrow} X(Z).$$

To conclude from convergence in the ζ_s metric to weak convergence on $(\mathcal{C}[0,1], \|\cdot\|_\infty)$ or $(\mathcal{D}[0,1], d_{\mathrm{sk}})$, further assumptions are needed. Let, for $r > 0$,

$$\begin{aligned}\mathcal{C}_r[0,1] := \big\{ f \in \mathcal{C}[0,1] \,|\, \exists 0 = t_1 < t_2 < \cdots < t_\ell = 1,\ \forall i = 1, \ldots, \ell : \\ |t_i - t_{i-1}| \ge r,\ f|_{[t_{i-1}, t_i]} \text{ is linear}\big\}\end{aligned} \tag{15}$$

denote the set of all continuous functions for which there is a decomposition of $[0, 1]$ into intervals of length at least r such that the function is piecewise linear on those intervals. Analogously, we define

$$\begin{aligned}\mathcal{D}_r[0,1] := \bigl\{f \in \mathcal{D}[0,1] | \exists 0 = t_1 < t_2 < \cdots < t_\ell = 1,\ \forall i = 1,\dots,\ell:\\ |t_i - t_{i-1}| \geq r,\ f|_{[t_{i-1},t_i)} \text{ is constant, continuous in } 1\bigr\}.\end{aligned} \tag{16}$$

THEOREM 9. *Let X_n be random variables in $\mathcal{C}_{r_n}[0, 1]$, $n \geq 0$, and X a random variable in $\mathcal{C}[0, 1]$. Assume that for $0 < s \leq 3$ with $s = m + \alpha$ as in* (7)

$$\zeta_s(X_n, X) = o\biggl(\log^{-m}\biggl(\frac{1}{r_n}\biggr)\biggr). \tag{17}$$

Then $X_n \to X$ in distribution. The assertion remains valid if $\mathcal{C}[0, 1]$, $\mathcal{C}_{r_n}[0, 1]$ are replaced by $\mathcal{D}[0, 1]$, $\mathcal{D}_{r_n}[0, 1]$ endowed with the Skorokhod topology and X has continuous sample paths.

As discussed above, ζ_s convergence does not imply weak convergence in the spaces $\mathcal{C}[0, 1]$ and $\mathcal{D}[0, 1]$ without any further assumption such as (17). In the counterexample from [30], the sequence $S_n/\sqrt{n}$ there converges to a Gaussian limit with respect to ζ_s for $2 < s \leq 3$ where the rate of convergence is upper bounded by the order $n^{1-s/2}$; see [32] or [31]. Moreover, the sequence is piecewise linear but the sequence r_n can only be chosen of the order $(cn)^{-2n}$ for some $c > 0$. Hence, (17) is not satisfied.

In applications such as our proof of Donsker's functional limit law in Section 4.1.1 or the application of the present methodology to a problem from the probabilistic analysis of algorithms in [6], the rate of convergence will typically be of polynomial order which is fairly sufficient.

We postpone the proof of the theorem to the end of this section and state two variants, where the first one, Corollary 10, contains a slight relaxation of the assumptions that is useful in applications such as in the analysis of the complexity of partial match queries in quadtrees; see Section 4.2 or [6]. The second one will be needed in the case $s > 2$; see Section 4.1.

COROLLARY 10. *Let X_n, X be $\mathcal{C}[0, 1]$ valued random variables, $n \geq 0$, and $0 < s \leq 3$ with $s = m + \alpha$ as in* (7). *Suppose $X_n = Y_n + h_n$ with Y_n being $\mathcal{C}[0, 1]$ valued random variables and $h_n \in \mathcal{C}[0, 1]$, $n \geq 0$, such that $\|h_n - h\|_\infty \to 0$ for a $h \in \mathcal{C}[0, 1]$ and*

$$\mathbf{P}\bigl(Y_n \notin \mathcal{C}_{r_n}[0,1]\bigr) \to 0. \tag{18}$$

If

$$\zeta_s(X_n, X) = o\biggl(\log^{-m}\biggl(\frac{1}{r_n}\biggr)\biggr),$$

then

$$X_n \xrightarrow{d} X.$$

The statement remains true if $\mathcal{C}[0,1]$ *and* $\mathcal{C}_{r_n}[0,1]$ *are replaced by* $\mathcal{D}[0,1]$ *and* $\mathcal{D}_{r_n}[0,1]$ *endowed with the Skorokhod topology, respectively,* X *has continuous sample paths and* h *remains continuous.*

COROLLARY 11. *Let* X_n, Y_n, X *be* $\mathcal{C}[0,1]$ *valued random variables,* $n \geq 0$, *and* $0 < s \leq 3$ *with* $s = m + \alpha$ *as in* (7). *Suppose* $X_n \in \mathcal{C}_{r_n}[0,1]$ *for all* n *and* $Y_n \to X$ *in distribution. If*

$$\zeta_s(X_n, Y_n) = o\left(\log^{-m}\left(\frac{1}{r_n}\right)\right),$$

then

$$X_n \xrightarrow{d} X.$$

The statement remains true if $\mathcal{C}[0,1]$ *and* $\mathcal{C}_{r_n}[0,1]$ *are replaced by* $\mathcal{D}[0,1]$ *and* $\mathcal{D}_{r_n}[0,1]$ *endowed with the Skorokhod topology, respectively, and* X *has continuous sample paths.*

In $\mathcal{C}[0,1]$ (or $\mathcal{D}[0,1]$, if the limit X has continuous paths), convergence in distribution implies distributional convergence of the supremum norm $\|X_n\|_\infty$ by the continuous mapping theorem. In applications, one is also interested in convergence of moments of the supremum. For random variables X in $\mathcal{C}[0,1]$ or $\mathcal{D}[0,1]$, we denote by

$$\|X\|_s := \left(\mathbf{E}\left[\|X\|_\infty^s\right]\right)^{(1/s)\wedge 1}$$

the L_s-norm of the supremum norm.

THEOREM 12. *Let* X_n, X *be* $\mathcal{C}[0,1]$ *valued random variables and* $0 < s \leq 3$ *with* $\|X_n\|_s, \|X\|_s < \infty$ *for all* $n \geq 0$. *Suppose one of the following conditions is satisfied:*

(1) $X_n \in \mathcal{C}_{r_n}[0,1]$ *for all* n *and*

$$\zeta_s(X_n, X) = o\left(\log^{-m}\left(\frac{1}{r_n}\right)\right). \tag{19}$$

(2) $X_n = Y_n + h_n$ *with* Y_n *being* $\mathcal{C}[0,1]$ *valued random variables and* $h_n \in \mathcal{C}[0,1]$, $n \geq 0$, *such that* $\|h_n - h\|_\infty \to 0$ *for a* $h \in \mathcal{C}[0,1]$,

$$\mathbf{E}\left[\|X_n\|_\infty^s \mathbf{1}_{\{Y_n \notin \mathcal{C}_{r_n}[0,1]\}}\right] \to 0 \tag{20}$$

and

$$\zeta_s(X_n, X) = o\left(\log^{-m}\left(\frac{1}{r_n}\right)\right).$$

(3) $(Y_n)_{n\geq 0}$ *is a sequence of* $\mathcal{C}[0,1]$ *valued random variables with* $Y_n \leq Z$ *almost surely for a* $\mathcal{C}[0,1]$ *valued random variable* Z *with* $\|Z\|_s < \infty$, $X_n \in \mathcal{C}_{r_n}[0,1]$ *for all* n *and*

$$\zeta_s(X_n, Y_n) = o\left(\log^{-m}\left(\frac{1}{r_n}\right)\right).$$

Then $\{\|X_n\|_\infty^s | n \geq 0\}$ *is uniformly integrable. All statements remain true if* $\mathcal{C}[0,1], \mathcal{C}_{r_n}[0,1]$ *are replaced by* $\mathcal{D}[0,1], \mathcal{D}_{r_n}[0,1]$ *and* h *in item* (2) *remains continuous.*

It is of interest whether the metric space $(\mathcal{M}_s(\nu), \zeta_s)$ is complete. This is true for $0 < s \leq 1$. Also, in the case that B is a separable Hilbert space, this holds true; see Theorem 5.1 in [12]. Nevertheless, the problem remains open in the general case, in particular in the cases $\mathcal{C}[0,1]$ and $\mathcal{D}[0,1]$ with $s > 1$. We can only state the following proposition.

PROPOSITION 13. *Let* $B = (\mathcal{C}[0,1], \|\cdot\|_\infty)$ *or* $B = (\mathcal{D}[0,1], d_{\mathrm{sk}})$, $s > 0$ *and* $\nu \in \mathcal{M}_s(B)$. *Furthermore, let* $(\mu_n)_{n\geq 0}$ *be a sequence of probability measures from* $\mathcal{M}_s(\nu)$ *which is a Cauchy sequence with respect to the* ζ_s *metric. Then there exists a probability measure* μ *on* $\mathbb{R}^{[0,1]}$ *such that, as* $n \to \infty$,

$$\mu_n \xrightarrow{\text{f.d.d.}} \mu. \tag{21}$$

PROOF. Let $\mathcal{L}(X_n) = \mu_n$ for all $n \geq 0$. According to (13), $(X_n(t_1), \dots, X_n(t_k))_{n\geq 0}$ is a Cauchy sequence and hence it exists a random variable $Y_{t_1,\dots,t_k}$ in $\mathbb{R}^k$ with

$$\big(X_n(t_1), \dots, X_n(t_k)\big) \xrightarrow{d} Y_{t_1,\dots,t_k} \qquad (n \to \infty).$$

The set of distributions of $Y_{t_1,\dots,t_k}$ for $0 \leq t_1 < \cdots < t_k \leq 1$ and $k \in \mathbb{N}$ is consistent so there exists a process Y on the product space $\mathbb{R}^{[0,1]}$ whose distribution satisfies (21). □

REMARK 14. If the distribution μ found in Proposition 13 has a version with continuous paths then condition (10) for μ_n and μ is satisfied.

We now present proofs of the theorems and corollaries of the present sections. Theorem 9 essentially follows directly from Theorem 2 in [2]; see also [3]. Nevertheless, we present a version of the proof given there so that we can deduce the variants and implications given in our other statements. A basic tool are Theorems 2.2, 2.3 and 2.4 in Billingsley [5].

LEMMA 15. *Let* $(\mu_n)_{n\geq 0}, \mu$ *be probability measures on a separable metric space* (S,d). *For* $r > 0, x \in S$ *let* $B_r(x) = \{y \in S : d(x,y) < r\}$. *If for any* $x_1,\dots,x_k \in S, \gamma_1,\dots,\gamma_k > 0$ *with* $\mu(\partial B_{\gamma_i}(x_i)) = 0$ *for* $i = 1,\dots,k$ *it holds*

$$\mu_n\Big(\bigcap_{i\in I} B_{\gamma_i}(x_i)\Big) \to \mu\Big(\bigcap_{i\in I} B_{\gamma_i}(x_i)\Big), \tag{22}$$

where $I = \{1,\dots,k\}$, *then* $\mu_n \to \mu$ *weakly.*

Let $(S,d) = (\mathcal{D}[0,1], d_{\mathrm{sk}})$. *Then the assertion remains true when the balls* $B_{\gamma_i}(x_i)$ *are still defined with respect to the* uniform *distance and* $\mu(\mathcal{C}[0,1]) = 1$.

PROOF. The first part of the lemma is a special case of Theorem 2.4 in [5]. To prove the assertion in the càdlàg space, we apply Theorem 2.2 in [5] upon choosing $\mathcal{A}_P$ there to be the set of finite intersection of sets A where A is either a μ-continuous open sphere (in the uniform distance) whose center lies in $\mathcal{C}[0,1]$ or a measurable set with positive uniform distance from $\mathcal{C}[0,1]$. Using (22) and the inclusion-exclusion formula, it is easy to see that $\mu_n(C) \to 0$ for any measurable set C with positive uniform distance from $\mathcal{C}[0,1]$, in particular $\mu_n(A) \to \mu(A)$ for any $A \in \mathcal{A}_P$. Moreover, we can decompose any open set $O \in \mathcal{D}[0,1]$ (in the Skorokhod topology) into O' and $O \setminus O'$ with

$$O' := \bigcup_{x,\delta} B_x^{\|\cdot\|}(\delta),$$

where the union is over all $x \in O \cap \mathcal{C}'$ for a countable set $\mathcal{C}'$ that is dense in $\mathcal{C}[0,1]$ and $\delta \in \mathbb{Q}^+$ such that $B_x^{\|\cdot\|}(\delta) \subseteq O$ and $B_x^{\|\cdot\|}(\delta)$ is μ-continuous. We have $O \cap \mathcal{C}[0,1] \subseteq O'$ since any ball in the metric d_{sk} with center in $\mathcal{C}[0,1]$ contains a concentric ball in the uniform distance. Hence,

$$O \setminus O' = \bigcup_{\delta\in\mathbb{Q}^+} \big\{x \in O \setminus O' : \|y - x\| > \delta \text{ for all } y \in \mathcal{C}[0,1]\big\}.$$

Thus, any open set O is a countable union of sets in $\mathcal{A}_P$ which proves all conditions of Theorem 2.2 in [5] to be satisfied and the claim follows. $\square$

A main difficulty in deducing weak convergence from convergence in ζ_s compared to the Hilbert space case is the nondifferentiability of the norm function $x \mapsto \|x\|_\infty$; see [10], page 147. We will instead use the smoother L_p-norm which approximates the supremum norm in the sense that

$$L_p(x) \to \|x\|_\infty \tag{23}$$

for any fixed $x \in \mathcal{C}[0,1]$ as $p \to \infty$.

For the remaining part of this section, p, for fixed values or tending to infinity, is always to be understood as an even integer with $p \geq 4$. We use the Bachmann–Landau big-O notation.

LEMMA 16. *For $x, y \in \mathcal{C}[0,1]$ let*

$$L_p(x) = \left(\int_0^1 [x(t)]^p\,dt\right)^{1/p}, \qquad \psi_{p,y}(x) = L_p\big((1+[x-y]^2)^{1/2}\big).$$

Then L_p is smooth on $\mathcal{C}[0,1]\setminus\{\mathbf{0}\}$ where $\mathbf{0}$ is the zero-function and $\psi_{p,y}$ is smooth on $\mathcal{C}[0,1]$ for all $y \in \mathcal{C}[0,1]$. Furthermore, for $k \in \{1,2,3\}$, we have

$$\|D^k L_p(x)\| = O\big(p^{k-1} L_p^{1-k}(x)\big),$$

uniformly for p and $x \in \mathcal{C}[0,1]\setminus\{\mathbf{0}\}$. Moreover, again for $k \in \{1,2,3\}$,

$$\|D^k \psi_{p,y}(x)\| = O\big(p^{k-1}\big) \tag{24}$$

uniformly for p and $x, y \in \mathcal{C}[0,1]$. All assertions remain valid when $\mathcal{C}[0,1]$ is replaced by $\mathcal{D}[0,1]$, moreover both functions L_p and $\psi_{p,y}$ are continuous with respect to the Skorokhod metric for all p and $y \in \mathcal{D}[0,1]$.

PROOF. The smoothness properties are obvious. Differentiating L_p by the chain rule yields

$$DL_p(x)[h] = \left(\int_0^1 [x(t)]^p\,dt\right)^{1/p-1} \int_0^1 [x(t)]^{p-1} h(t)\,dt.$$

For $h \in \mathcal{C}[0,1]$ with $\|h\| \le 1$ by Jensen's inequality and $L_p(h) \le \|h\|$, we obtain that the right-hand side of the latter display is uniformly bounded by 1. The bounds on the norms of the higher order derivatives follow along the same lines. Using the same ideas, it is easy to see that

$$\|D^k \psi_{p,y}(x)\| = O\left(\sum_{j=1}^k p^{j-1} L_p^{1-j}\big(\omega_y(x)\big)\right),$$

uniformly in p and $x, y \in \mathcal{C}[0,1]$ where $\omega_y(x) = (1+|x-y|^2)^{1/2}$. This gives (24). $\square$

Note that the convergence in (23) holds pointwise; it is easy to construct a sequence of continuous functions $(x_p)_{p\ge 0}$ such that $L_p(x_p) \to 0$ and $\|x_p\|_\infty \to \infty$ as $p \to \infty$. Additionally to the obvious bound $L_p(x) \le \|x\|_\infty$, we will need the following simple lemma which contains sort of a converse of this inequality.

LEMMA 17. *Let λ denote the Lebesgue measure on the unit interval and let $\gamma > 0$ and $0 < \vartheta < 1$.*

(a) *For all $f \in \mathcal{D}_r[0,1]$, we have*

$$\|f\|_\infty \ge \gamma \quad \Rightarrow \quad \lambda\big(\{t : |f(t)| \ge (1-\vartheta)\gamma\}\big) \ge r.$$

Moreover, for any $g \in \mathcal{C}[0,1]$, there exists a $\delta = \delta(g,\gamma,\vartheta) > 0$ such that

$$\|f-g\|_\infty \ge \gamma \quad \Rightarrow \quad \lambda\big(\{t : |f(t)-g(t)| \ge (1-\vartheta)\gamma\}\big) \ge \min(r,\delta).$$

(b) *For all* $f \in \mathcal{C}_r[0,1]$, *we have*

$$\|f\|_\infty \geq \gamma \quad \Rightarrow \quad \lambda\bigl(\{t : |f(t)| \geq (1-\vartheta)\gamma\}\bigr) \geq \frac{\vartheta}{2} r.$$

Moreover, for $g \in \mathcal{C}[0,1]$, *there exists a* $\delta = \delta(g,\gamma,\vartheta) > 0$ *with*

$$\|f-g\|_\infty \geq \gamma \quad \Rightarrow \quad \lambda\bigl(\{t : |f(t)-g(t)| \geq (1-\vartheta)\gamma\}\bigr) \geq \frac{\vartheta}{4} \min(r,\delta).$$

PROOF. Ad (a): The first assertion is trivial. The second one follows by choosing $\delta > 0$ small enough such that $|g(x)-g(y)| \leq \frac{\vartheta\gamma}{2}$ for all $|x-y| < \delta$.

Ad (b): For the first statement, assume $\|f\|_\infty \geq \gamma$ and let $[e_0, e_1]$ be an interval where f attains its maximum. A geometric argument shows that the quantity $\lambda(\{t \in [e_0,e_1] : |f(t)| \geq (1-\vartheta)\gamma\})$ is minimized when $f(e_0) = \gamma$ and $f(e_1) = -(1-\vartheta)\gamma$. In this case, the quantity equals $\vartheta r/(2(2-\vartheta))$ which implies the assertion since $0 < \vartheta < 1$. Finally, the last statement follows from a combination of the latter argument and by choosing $\delta > 0$ again such that $|g(x)-g(y)| \leq \frac{\vartheta\gamma}{2}$ for all $|x-y| < \delta$. $\square$

We start with the proofs of Theorem 9 and its corollaries in the continuous case.

PROOF OF THEOREM 9. For $r > 0$, $x \in \mathcal{C}[0,1]$ let $B_r(x) = \{y \in \mathcal{C}[0,1] : \|y - x\|_\infty < r\}$. According to Lemma 15, we need to verify that

$$\mathbf{P}\Bigl(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Bigr) \to \mathbf{P}\Bigl(X \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Bigr) \tag{25}$$

for $I = \{1,\dots,k\}$ and $x_1,\dots,x_k \in S, \gamma_1,\dots,\gamma_k > 0$ such that $\mathbf{P}(X \in (\partial B_{\gamma_i}(x_i))) = 0$. The lack of uniformity in (23) leads us to find lower and upper bounds on the desired quantity. We will establish

$$\limsup_{n\to\infty} \mathbf{P}\Bigl(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Bigr) \leq \mathbf{P}\Bigl(X \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Bigr) \tag{26}$$

and

$$\liminf_{n\to\infty} \mathbf{P}\Bigl(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Bigr) \geq \mathbf{P}\Bigl(X \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Bigr) \tag{27}$$

separated from each other. To this end, it is sufficient to construct functions $g_{i,n}, \tilde{g}_{i,n} : \mathcal{C}[0,1] \to [0,1]$ satisfying

$$\tilde{g}_{i,n}(x) \leq \mathbf{1}_{B_{\gamma_i}(x_i)}(x) \leq g_{i,n}(x) \qquad \text{for all } x \in \mathcal{C}_{r_n}[0,1], \tag{28}$$

$$g_{i,n}(x), \tilde{g}_{i,n}(x) \to \mathbf{1}_{B_{\gamma_i}(x_i)}(x) \qquad \text{for all } x \in \mathcal{C}[0,1] \setminus \partial B_{\gamma_i}(x_i) \tag{29}$$

and such that $a_n \prod_{i\in I} g_{i,n}$, $\tilde{a}_n \prod_{i\in I} \tilde{g}_{i,n} \in \mathcal{F}_s$ for appropriate constants $a_n, \tilde{a}_n > 0$ such that $a_n^{-1}\zeta_s(X_n, X) \to 0$ and $\tilde{a}_n^{-1}\zeta_s(X_n, X) \to 0$ as $n \to \infty$. This is sufficient since we then may conclude

$$\begin{aligned} \mathbf{P}\Big(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Big) &\le \mathbf{E}\Big[\prod_{i\in I} g_{i,n}(X_n)\Big] \\ &\le \mathbf{E}\Big[\prod_{i\in I} g_{i,n}(X)\Big] + a_n^{-1}\zeta_s(X_n, X) \end{aligned} \tag{30}$$

and

$$\begin{aligned} \mathbf{P}\Big(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Big) &\ge \mathbf{E}\Big[\prod_{i\in I} \tilde{g}_{i,n}(X_n)\Big] \\ &\ge \mathbf{E}\Big[\prod_{i\in I} \tilde{g}_{i,n}(X)\Big] - \tilde{a}_n^{-1}\zeta_s(X_n, X). \end{aligned} \tag{31}$$

While this is the basic idea subsequently, the construction is slightly more involved.

We first give a motivation of how to construct the functions $g_{i,n}$: according to (29), asymptotically, the functions $g_{i,n}$ have to separate points $x \in \mathcal{C}[0,1]$ which are in $B_{\gamma_i}(x_i)$ from those which are not. This is why we use the L_p norm. Consider ψ_{p,x_i} as introduced in Lemma 16. If $x \in \overline{B_{\gamma_i}(x_i)}$, then $\psi_{p,x_i}(x) \le (1+\gamma_i^2)^{1/2}$ whereas if $x \notin \overline{B_{\gamma_i}(x_i)}$ then $\liminf_{p\to\infty} \psi_{p,x_i}(x) > (1+\gamma_i^2)^{1/2}$.

Let $\varphi : \mathbb{R} \to [0,1]$ be a three times continuously differentiable function with $\varphi(u) = 1$ for $u \le 0$ and $\varphi(u) = 0$ for $u \ge 1$. For $\varrho \in \mathbb{R}$ and $\eta > 0$, we denote $\varphi_{\varrho,\eta} : \mathbb{R}^+ \to [0,1]$ by $\varphi_{\varrho,\eta}(u) = \varphi((u-\varrho)/\eta)$.

Let $g_i(x) = \varphi_{(1+\gamma_i^2)^{1/2},\eta}(\psi_{p,x_i}(x))$. Let $g_{i,n} = g_i$ with $\eta = \eta_n \downarrow 0$ and $p = p_n \uparrow \infty$. Then $g_{i,n}$ has the properties in (28) and (29).

We do not know how to construct functions $\tilde{g}_{i,n}$ with the properties (28) and (29). Instead, we construct functions $\bar{g}_{i,n}$ satisfying related conditions: let $0 < \vartheta < 1$ and $x \in \mathcal{C}_{r_n}[0,1]$. By Lemma 17(b), we can find $\delta = \delta(\vartheta)$ (also depending on $x_1, \dots, x_k, \gamma_1, \dots, \gamma_k$ which are kept fixed) with

$$\begin{aligned} &\{\|x - x_i\|_\infty \ge \gamma_i\} \\ &\quad \subseteq \Big\{\lambda\big(\{t : |x(t) - x_i(t)| \ge \gamma_i(1-\vartheta)\}\big) \ge \frac{\vartheta}{4}\min(r_n, \delta)\Big\} \\ &\quad \subseteq \Big\{\psi_{p,x_i}(x) \ge \big(1 + \gamma_i^2(1-\vartheta)^2\big)^{1/2}\Big(\frac{\vartheta}{4}\min(r_n, \delta)\Big)^{1/p}\Big\} \\ &\quad \subseteq \{\bar{g}_{i,n}(x) = 0\} \end{aligned} \tag{32}$$

with $\bar{g}_{i,n}(x) = \varphi_{(1+\gamma_i^2(1-\vartheta)^2)^{1/2}(\vartheta \min(r_n,\delta)/4)^{1/p} - \eta,\eta}(\psi_{p,x_i}(x))$. This gives (28). $\bar{g}_{i,n}$ does not fulfill (29), but we have

$$\bar{g}_{i,n}(x) \to \mathbf{1}_{B_{\gamma_i(1-\vartheta)}(x_i)}(x)$$

for $x \in \mathcal{C}[0,1] \setminus \partial B_{\gamma_i(1-\vartheta)}(x_i)$ and $p = p_n \uparrow \infty$, $\eta = \eta_n \downarrow 0$ such that $r_n^{1/p_n} \to 1$. This gives for every $0 < \vartheta < 1$ with $\mathbf{P}(X \in \partial B_{\gamma_i(1-\vartheta)}(x_i)) = 0$ for all $i \in I$

$$\lim_{n\to\infty} \mathbf{E}\Big[\prod_{i\in I} \bar{g}_{i,n}(X)\Big] = \mathbf{P}\Big(X \in \bigcap_{i\in I} B_{\gamma_i(1-\vartheta)}(x_i)\Big).$$

Assuming that $\bar{a}_n \prod_{i\in I} \bar{g}_{i,n} \in \mathcal{F}_s$ and letting n tend to infinity (31) rewrites as

$$\begin{aligned} &\liminf_{n\to\infty} \mathbf{P}\Big(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Big) \\ &\qquad \geq \mathbf{P}\Big(X \in \bigcap_{i\in I} B_{\gamma_i(1-\vartheta)}(x_i)\Big) - \limsup_{n\to\infty} \bar{a}_n^{-1} \zeta_s(X_n, X), \end{aligned} \tag{33}$$

where $\bar{a}_n$ may depend on ϑ and δ. Below, we will see that the error term on the right-hand side of (33) vanishes as $n \to \infty$ uniformly in ϑ, δ. So, choosing $\vartheta \downarrow 0$ such that $\mathbf{P}(X \in \partial B_{\gamma_i(1-\vartheta)}(x_i)) = 0$ for all $i \in I$ the assertion

$$\liminf_{n\to\infty} \mathbf{P}\Big(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Big) \geq \mathbf{P}\Big(X \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Big)$$

follows.

It remains to show that the error terms vanish in the limit. By Lemma 16 $g(x) = \varphi_{\varrho,\eta}(\psi_{p,y}(x))$ and using the mean value theorem, we obtain for $m = 0, 1, 2$

$$\|g^{(m)}(x+h) - g^{(m)}(x)\| \leq C_m p^m \eta^{-(m+1)} \|h\|_\infty^\alpha$$

for $p \geq 4$, $\eta < 1$ and some constants $C_m > 0$. It is easy to check that the same is valid for products of functions of form g with different constants, independent of the parameters. It follows that both error terms in (30) and (33) are bounded by $C'_m p_n^m \eta_n^{-(m+1)} \zeta_s(X_n, X)$ for all n, uniformly in ϑ, δ, where C'_m denotes a fixed constant for each $m \in \{0, 1, 2\}$. By (17), we can choose $p_n \uparrow \infty$ and $\eta_n \downarrow 0$ such that both $r_n^{1/p_n} \to 1$ and the error terms vanish in the limit. $\square$

PROOF OF COROLLARY 10. Again, according to Lemma 15, we only have to verify (25), for which we modify the proof of Theorem 9: first note that the assumption of piecewise linearity of X_n and the convergence rate for $\zeta_s(X_n, X)$ are not necessary for the upper bound

$$\limsup_{n\to\infty} \mathbf{P}\Big(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Big) \leq \mathbf{P}\Big(X \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Big).$$

For the lower bound let $\varepsilon > 0$ and note that

$$\mathbf{P}\Big(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Big) \geq \mathbf{P}\Big(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i) \cap \{Y_n \in \mathcal{C}_{r_n}[0,1]\}\Big).$$

We modify the functions $\bar{g}_{i,n}(x)$. Let $0 < \gamma_{K_i} < \gamma_i$ such that

$$\mathbf{P}\Big(X \in \bigcap_{i\in I} B_{\gamma_{K_i}}(x_i)\Big) \geq \mathbf{P}\Big(X \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Big) - \varepsilon$$

and $\mathbf{P}(X \in \partial B_{\gamma_{K_i}}(x_i)) = 0$ for all i. Let $0 < \vartheta < 1$ and n_0 be large enough such that $\varrho_n = \|h_n - h\|_\infty < \min_i(\gamma_{K_i}(1-\vartheta) \wedge \gamma - \gamma_{K_i})$ and $\mathbf{P}(Y_n \notin \mathcal{C}_{r_n}[0,1]) < \varepsilon$ for all $n \geq n_0$. By Lemma 17(b), there exists $\delta = \delta(\vartheta)$ such that for $y \in \mathcal{C}_{r_n}[0,1]$ with $x = y + h_n$ and $n \geq n_0$

$$\begin{aligned}
&\{\|x - x_i\|_\infty \geq \gamma_i\}\\
&\quad\subseteq \{\|y + h - x_i\|_\infty \geq \gamma_{K_i}\}\\
&\quad\subseteq \Big\{\lambda\big(\{t : |y(t) + h(t) - x_i(t)| \geq \gamma_{K_i}(1-\vartheta)\}\big) \geq \frac{\vartheta}{4}\min(r_n, \delta)\Big\}\\
&\quad\subseteq \Big\{\lambda\big(\{t : |x(t) - x_i(t)| \geq \gamma_{K_i}(1-\vartheta) - \varrho_n\}\big) \geq \frac{\vartheta}{4}\min(r_n, \delta)\Big\}\\
&\quad\subseteq \Big\{\psi_{p,x_i}(x) \geq \big(1 + (\gamma_{K_i}(1-\vartheta) - \varrho_n)^2\big)^{1/2}\Big(\frac{\vartheta}{4}\min(r_n,\delta)\Big)^{1/p}\Big\}\\
&\quad\subseteq \{\bar{g}_{i,n}(x) = 0\}
\end{aligned}$$

with $\bar{g}_{i,n}(x) = \varphi_{(1+(\gamma_{K_i}(1-\vartheta)-\varrho_n)^2)^{1/2}(\vartheta \min(r_n,\delta)/4)^{1/p}-\eta,\eta}(\psi_{p,x_i}(x))$. Hence,

$$\mathbf{P}\Big(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Big) \geq \mathbf{E}\Big[\prod_{i\in I} \bar{g}_{i,n}(X_n)\mathbf{1}_{\{Y_n \in \mathcal{C}_{r_n}[0,1]\}}\Big] \geq \mathbf{E}\Big[\prod_{i\in I} \bar{g}_{i,n}(X_n)\Big] - \varepsilon$$

for $n \geq n_0$. The upper bound of the error term $\bar{a}_n^{-1}\zeta_s(X_n, X)$ is a function of p and η so it is uniform in $\varrho_n, \vartheta, \delta$. Following the same lines as in the proof of Theorem 9 gives

$$\begin{aligned}
\liminf_{n\to\infty} \mathbf{P}\Big(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Big) &\geq \mathbf{P}\Big(X \in \bigcap_{i\in I} B_{\gamma_{K_i}}(x_i)\Big) - \varepsilon\\
&\geq \mathbf{P}\Big(X \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Big) - 2\varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, the result follows. □

PROOF OF COROLLARY 11. In the setting of the proof of Theorem 9, (30) rewrites as

$$\begin{aligned}
&\mathbf{P}\Big(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Big)\\
&\quad\leq \mathbf{E}\Big[\prod_{i\in I} g_{i,n}(X_n)\Big] \leq \mathbf{E}\Big[\prod_{i\in I} g_{i,n}(Y_n)\Big] + a_n^{-1}\zeta_s(X_n, Y_n)\\
&\quad= \mathbf{E}\Big[\prod_{i\in I} g_{i,n}(Y_n)\Big] - \mathbf{E}\Big[\prod_{i\in I} g_{i,n}(X)\Big] + \mathbf{E}\Big[\prod_{i\in I} g_{i,n}(X)\Big] + a_n^{-1}\zeta_s(X_n, Y_n).
\end{aligned}$$

We may choose $Y_n \to X$ almost surely. On the event $\{X \in B_{\gamma_i}(x_i)\}$, we have $\lim_n g_{i,n}(Y_n) = \lim_n g_{i,n}(X) = 1$ and on $\{X \notin \overline{B_{\gamma_i}(x_i)}\}$ we have $\lim_n g_{i,n}(Y_n) = \lim_n g_{i,n}(X) = 0$. Since $\mathbf{P}(X \in \partial B_{\gamma_i}(x_i)) = 0$, it follows

$$\prod_{i\in I} g_{i,n}(Y_n) - \prod_{i\in I} g_{i,n}(X) \to 0$$

for $n \to \infty$ almost surely and dominated convergence yields

$$\limsup_{n\to\infty} \mathbf{P}\Big(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Big) \le \mathbf{P}\Big(X \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\Big),$$

just like in the proof of Theorem 9. The lower bound follows similarly. $\square$

We now head over to the case of càdlàg functions. We only discuss the approach in the proof of Theorem 9. Following exactly the same arguments as in the continuous case and using the additional statements of Lemmas 16 and 17(a), it is easy to see that we also obtain (25) if the balls $B_{\gamma_i}(x_i)$ are defined with the uniform metric in $\mathcal{D}[0, 1]$. Remember that we still have $x_i \in \mathcal{C}[0, 1]$. Thus, Lemma 15 yields the assertion.

The proof of Theorem 12 is close to the one of Lemma 5.3 in [12]. The L_p approximation of the supremum norm complicates the argument slightly. We only give all details in the continuous case.

PROOF OF THEOREM 12. Suppose $0 \le s \le 3$ and that the first assumption of Theorem 12 is satisfied. Let $\kappa : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ be a smooth, monotonic function with $\kappa(u) = 0$ for $u \le \frac{1}{2}$ and $\kappa(u) = u^s$ for $u \ge 1$. We may as well assume that the interpolation for $\frac{1}{2} \le u \le 1$ is done smoothly such that we have $\kappa(u) \le u^s$ for $\frac{1}{2} \le u \le 1$, thus $\kappa(u) \le u^s$ for all $u \in \mathbb{R}_0^+$. Let $f, f^{(p)} : \mathcal{C}[0, 1] \to \mathbb{R}$ be given by

$$f(x) = \kappa\big(\|x\|_\infty\big),$$
$$f^{(p)}(x) = \kappa\big(L_p(x)\big).$$

By Lemma 16, the restrictions of L_p and $f^{(p)}$ to $\mathcal{C}[0, 1] \setminus \{\mathbf{0}\}$ are smooth. Furthermore, all derivatives of $f^{(p)}$ vanish for $\|x\|_\infty < 1/2$ which implies that $f^{(p)}$ is smooth on $\mathcal{C}[0, 1]$. Again, by Lemma 16 it is easy to check that for any $k \in \{1, \dots, m+1\}$,

$$\|D^k f^{(p)}(x)\| = O\big(p^{k-1} L_p^{s-k}(x)\big),$$

uniformly in p and $x \in \mathcal{C}[0, 1]$. Let $x, y \in \mathcal{C}[0, 1]$ with $L_p(x), L_p(y) \le 2\|x - y\|_\infty$. Then

$$\begin{aligned}\|D^m f^{(p)}(x) - D^m f^{(p)}(y)\| &\le \|D^m f^{(p)}(x)\| + \|D^m f^{(p)}(y)\| \\ &= O\big(p^{m-1} \|x - y\|_\infty^\alpha\big).\end{aligned}$$

Conversely let $2\|x-y\|_\infty \le L_p(x)$ [the case $2\|x-y\|_\infty \le L_p(y)$ being analogous]. Then, by the mean value theorem, there exists $z \in [x,y] := \{\lambda x + (1-\lambda)y | \lambda \in [0,1]\}$, such that

$$\begin{aligned}\|D^m f^{(p)}(x) - D^m f^{(p)}(y)\| &= \|D^{m+1} f^{(p)}(z)\| \cdot \|x-y\|_\infty \\ &= O\big(p^m L_p^{\alpha-1}(x)\big) \cdot \|x-y\|_\infty \\ &= O\big(p^m \|x-y\|_\infty^\alpha\big).\end{aligned}$$

Hence, there is a constant $c>0$ such that $cp^{-m} f^{(p)} \in \mathcal{F}_s$ for all $p \ge 4$. We define, for $r>0$,

$$\begin{aligned}f_r(x) &:= cr^s f(x/r),\\ f_r^{(p)}(x) &:= cr^s f^{(p)}(x/r).\end{aligned}$$

Then $p^{-m} f_r^{(p)} \in \mathcal{F}_s$. Furthermore, $f_r(x)$ and $f_r^{(p)}(x)$ are bounded by $c\|x\|^s$ for all $x \in \mathcal{C}[0,1]$, uniformly in p. For any fixed x we have $f_r(x) \to 0$ and $\sup_{p\ge 4} f_r^{(p)}(x) \to 0$ as $r \to \infty$. Hence, by $\mathbf{E}[\|X\|^s] < \infty$ and dominated convergence this implies

$$\mathbf{E}\Big[\sup_{p\ge 4} f_r^{(p)}(X)\Big] \to 0 \qquad (r \to \infty). \tag{34}$$

By the definition of ζ_s, we have

$$\mathbf{E}\big[f_r^{(p)}(X_n)\big] \le \mathbf{E}\big[f_r^{(p)}(X)\big] + p^m \zeta_s(X_n, X).$$

By the definition of f_r, for $\|x\| > r$ we have $\|x\|^s = c^{-1} f_r(x)$. Hence,

$$\begin{aligned}&\mathbf{E}\big[\|X_n\|_\infty^s \mathbf{1}_{\{\|X_n\|_\infty \ge 2r\}}\big]\\ &\quad = c^{-1}\mathbf{E}\big[f_r(X_n)\mathbf{1}_{\{\|X_n\|_\infty \ge 2r\}}\big]\\ &\quad \le c^{-1}\mathbf{E}\big[f_r^{(p)}(X_n)\big] + c^{-1}\big(\mathbf{E}\big[\big(f_r(X_n) - f_r^{(p)}(X_n)\big)\mathbf{1}_{\{\|X_n\|_\infty \ge 2r\}}\big]\big) \qquad (35)\\ &\quad \le c^{-1}\mathbf{E}\big[f_r^{(p)}(X)\big] + c^{-1} p^m \zeta_s(X_n, X)\\ &\qquad + c^{-1}\big(\mathbf{E}\big[\big(f_r(X_n) - f_r^{(p)}(X_n)\big)\mathbf{1}_{\{\|X_n\|_\infty \ge 2r\}}\big]\big).\end{aligned}$$

Now, let $\varepsilon > 0$ be arbitrary. By (34), fix $r>0$ such that $\mathbf{E}[f_r^{(p)}(X)] < \varepsilon$ for all $p \ge 4$. Additionally, by the given assumptions there exists a sequence $p_n \uparrow \infty$ such that

$$\frac{\log r_n}{p_n} \to 0, \qquad p_n^m \zeta_s(X_n, X) \to 0 \qquad (n \to \infty).$$

Therefore, let N_0 be large enough such that $p_n^m \zeta_s(X_n, X) < \varepsilon$ for all $n \ge N_0$. It remains to bound the third summand in (35). Using Lemma 17(a), piecewise linearity of X_n implies that for all $0 < \vartheta < 1$,

$$L_p(X_n) \ge \|X_n\|_\infty (1-\vartheta)\left(\frac{\vartheta r_n}{2}\right)^{1/p_n}.$$

In particular, we have $L_p(X_n) \geq \frac{\|X_n\|_\infty}{2}$ for all n sufficiently large. For those n and $\|X_n\| > 2r$ we also have $f_r^{(p)}(X_n) = cL_p^s(X_n)$. This yields

$$\begin{aligned}
&\mathbf{E}\big[\big(f_r(X_n) - f_r^{(p)}(X_n)\big)\mathbf{1}_{\{\|X_n\|_\infty \geq 2r\}}\big] \\
&\quad = c\mathbf{E}\big[\big(\|X_n\|_\infty^s - L_p^s(X_n)\big)\mathbf{1}_{\{\|X_n\|_\infty \geq 2r\}}\big] \qquad (36)\\
&\quad \leq c\big(1 - 2^{-s}\big)\mathbf{E}\big[\|X_n\|_\infty^s \mathbf{1}_{\{\|X_n\|_\infty \geq 2r\}}\big] \qquad (37)
\end{aligned}$$

for all n sufficiently large. Increasing N_0 if necessary, inserting (37) into (35) and rearranging terms implies

$$\mathbf{E}\big[\|X_n\|_\infty^s \mathbf{1}_{\{\|X_n\|_\infty \geq 2r\}}\big] \leq 2^{1+s} c^{-1} \varepsilon$$

for all $n \geq N_0$. Since ε was arbitrary, the assertion follows.

Now, suppose the second assumption is satisfied. Then we have to modify the last part of the proof. In (36), we can decompose

$$L_p^s(X_n) = L_p^s(X_n)\mathbf{1}_{\{Y_n \in \mathcal{C}_{r_n}[0,1]\}} + L_p^s(X_n)\mathbf{1}_{\{Y_n \notin \mathcal{C}_{r_n}[0,1]\}}.$$

Using $L_p^s(X_n) \leq \|X_n\|_\infty^s$, the assumptions guarantee the expectation of the second term to be small in the limit $n \to \infty$. For the first one, using similar arguments as above, given $\{Y_n \in \mathcal{C}_{r_n}[0, 1]\}$, we find

$$L_p(X_n) \geq \frac{\|X_n\|_\infty}{2} - 2\varrho_n$$

with $\varrho_n = \|h_n - h\|_\infty$ for all n sufficiently large. Proceeding as in the first part, we obtain the result. Given the third assumption, it only remains to bound $\mathbf{E}[f_r^{(p)}(Y_n)]$ which appears instead of $\mathbf{E}[f_r^{(p)}(X)]$ by $\mathbf{E}[f_r^{(p)}(Z)]$ in (35). $\square$

3. The contraction method. In this section, the contraction method is developed first for a general separable Banach space B. Then the framework is specialized to the cases $(\mathcal{C}[0, 1], \|\cdot\|_\infty)$ and $(\mathcal{D}[0, 1], d_{\mathrm{sk}})$. For this section, B will always denote a separable Banach space or $(\mathcal{D}[0, 1], d_{\mathrm{sk}})$.

We recall the recursive equation (2). We have

$$X_n \stackrel{d}{=} \sum_{r=1}^{K} A_r^{(n)} X_{I_r^{(n)}}^{(r)} + b^{(n)}, \qquad n \geq n_0, \qquad (38)$$

where $A_1^{(n)}, \dots, A_K^{(n)}$ are random continuous linear operators, $b^{(n)}$ is a B-valued random variable, $(X_n^{(1)})_{n\geq 0}, \dots, (X_n^{(K)})_{n\geq 0}$ are distributed like $(X_n)_{n\geq 0}$, and $I^{(n)} = (I_1^{(n)}, \dots, I_K^{(n)})$ is a vector of random integers in $\{0, \dots, n\}$. Moreover, $(A_1^{(n)}, \dots, A_K^{(n)}, b^{(n)}, I^{(n)}), (X_n^{(1)})_{n\geq 0}, \dots, (X_n^{(K)})_{n\geq 0}$ are independent and $n_0 \in \mathbb{N}$.

Recall that in order to be a random continuous linear operator, A has to take values in the set of continuous endomorphisms on $\mathcal{C}[0, 1]$, respectively, the set of norm-continuous endomorphisms that are continuous with respect to d_{sk} on $\mathcal{D}[0,1]$ such that $A(x)(t)$ is a real-valued random variable for all $x \in \mathcal{C}[0,1]$, respectively, $x \in \mathcal{D}[0,1]$ and $t \in [0,1]$. In $\mathcal{D}[0,1]$, we additionally have to guarantee $\|A\|_{\mathrm{op}}$ to be a real-valued random variable; see Section 2.2.

We make assumptions about the moments and the asymptotic behavior of the coefficients $A_1^{(n)}, \dots, A_K^{(n)}, b^{(n)}$. For a random continuous linear operator A, we write

$$\|A\|_s := \mathbf{E}\big[\|A\|_{\mathrm{op}}^s\big]^{1\wedge(1/s)}.$$

We consider the following conditions with an $s > 0$:

(C1) We have $\|X_0\|_s, \dots, \|X_{n_0-1}\|_s$, $\|A_r^{(n)}\|_s, \|b^{(n)}\|_s < \infty$ for all $r = 1, \dots, K$ and $n \geq 0$ and there exist random continuous linear operators $A_1, \dots, A_K$ on B and a B-valued random variable b such that, as $n \to \infty$,

$$\gamma(n) := \|b^{(n)} - b\|_s + \sum_{r=1}^K \big(\|A_r^{(n)} - A_r\|_s + \|\mathbf{1}_{\{I_r^{(n)} \leq n_0\}} A_r^{(n)}\|_s\big) \to 0 \tag{39}$$

and for all $\ell \in \mathbb{N}$,

$$\mathbf{E}\big[\mathbf{1}_{\{I_r^{(n)} \in \{0,\dots,\ell\} \cup \{n\}\}} \|A_r^{(n)}\|_{\mathrm{op}}^s\big] \to 0. \tag{40}$$

(C2) We have

$$L := \sum_{r=1}^K \mathbf{E}\big[\|A_r\|_{\mathrm{op}}^s\big] < 1.$$

The limits of the coefficients determine the limiting operator T from (5):

$$\begin{aligned} T &: \mathcal{M}(B) \to \mathcal{M}(B), \\ \mu &\mapsto \mathcal{L}\Bigg(\sum_{r=1}^K A_r Z^{(r)} + b\Bigg), \end{aligned} \tag{41}$$

where $(A_1, \dots, A_K, b)$, $Z^{(1)}, \dots, Z^{(K)}$ are independent and $Z^{(1)}, \dots, Z^{(K)}$ have distribution μ.

(C3) The map T has a fixed point $\eta \in \mathcal{M}_s(B)$, such that $\mathcal{L}(X_n) \in \mathcal{M}_s(\eta)$ for all $n \geq n_0$.

The existence of a fixed point is not in general implied by contraction properties of T with respect to a Zolotarev metric due to the lack of knowledge of completeness of the metric on a the space B. However, we can argue that there is at most one fixed point of T in $\mathcal{M}_s(\eta)$:

LEMMA 18. *Assume the sequence* $(X_n)_{n\ge 0}$ *satisfies* (38). *Under conditions* (C1)–(C3), *we have* $T(\mathcal{M}_s(\eta)) \subseteq \mathcal{M}_s(\eta)$ *and*

$$\zeta_s\big(T(\mu), T(\lambda)\big) \le L\zeta_s(\mu,\lambda) \qquad \text{for all } \mu,\lambda \in \mathcal{M}_s(\eta).$$

In particular, the restriction of T *to* $\mathcal{M}_s(\eta)$ *is a contraction and has the unique fixed-point* η.

PROOF. Let $\mu \in \mathcal{M}_s(\eta)$. Recall that we have $s = m+\alpha$ with $m \in \mathbb{N}_0$ and $\alpha \in (0,1]$. We introduce an accompanying sequence

$$Q_n := \sum_{r=1}^{K} A_r^{(n)}\big(\mathbf{1}_{\{I_r^{(n)}<n_0\}} X_{I_r^{(n)}}^{(r)} + \mathbf{1}_{\{I_r^{(n)}\ge n_0\}} Z^{(r)}\big) + b^{(n)}, \qquad n \ge n_0, \tag{42}$$

where $(A_1^{(n)},\dots,A_K^{(n)},b^{(n)})$, $Z^{(1)},\dots,Z^{(K)}$ are independent and $Z^{(1)},\dots,Z^{(K)}$ have distribution μ.

We first show that $\mathcal{L}(Q_n) \in \mathcal{M}_s(\eta)$ for all $n \ge n_0$. Condition (C1), conditioning on the coefficients and Minkowski's inequality, implies $\mathbf{E}[\|Q_n\|_\infty^s] < \infty$ for all n. For $s \le 1$, we already obtain $\mathcal{L}(Q_n) \in \mathcal{M}_s(\eta)$.

For $s > 1$, we choose arbitrary $1 \le k \le m$ and multilinear and bounded $f : B^k \to \mathbb{R}$. We have

$$\begin{aligned}\mathbf{E}[f(Z,\dots,Z)] &= \mathbf{E}[f(X_n,\dots,X_n)]\\ &= \mathbf{E}\Bigg[f\Bigg(\sum_{r=1}^{K} A_r^{(n)} X_{I_r^{(n)}}^{(r)} + b^{(n)},\dots,\sum_{r=1}^{K} A_r^{(n)} X_{I_r^{(n)}}^{(r)} + b^{(n)}\Bigg)\Bigg].\end{aligned}$$

To show $\mathcal{L}(Q_n) \in \mathcal{M}_s(\eta)$, we need to verify that the latter display is equal to $\mathbf{E}[f(Q_n,\dots,Q_n)]$. Since f is multilinear, both terms can be expanded as a sum and it suffices to show that the corresponding summands are equal:

$$\mathbf{E}[f(C_{j_1}^{(n)},\dots,C_{j_k}^{(n)})] = \mathbf{E}[f(D_{j_1}^{(n)},\dots,D_{j_k}^{(n)})], \tag{43}$$

where $j_1,\dots,j_k \in \{1,\dots,K\}$ and for each $i \in \{1,\dots,k\}$ we either have

$$C_{j_i}^{(n)} = A_{j_i}^{(n)} X_{I_{j_i}^{(n)}}^{(j_i)} \quad \text{and} \quad D_{j_i}^{(n)} = A_{j_i}^{(n)}\big(\mathbf{1}_{\{I_{j_i}^{(n)}<n_0\}} X_{I_{j_i}^{(n)}}^{(j_i)} + \mathbf{1}_{\{I_{j_i}^{(n)}\ge n_0\}} Z^{(j_i)}\big) \tag{44}$$

or

$$C_{j_i}^{(n)} = b^{(n)} \quad \text{and} \quad D_{j_i}^{(n)} = b^{(n)}. \tag{45}$$

The equality in (43) is obvious for the case where we have (45) for all $i = 1,\dots,k$. For the other cases, we have (44) for at least $1 \le \ell \le k$ arguments of f, say, for simplicity of presentation, for the first ℓ with $1 \le \ell_1 < \dots < \ell_d = \ell$ such that

$j_s = j_{\ell_i}$ for all $s = \ell_{i-1}+1, \dots, \ell_i, i = 1, \dots, d$ and j_{ℓ_i} pairwise different for $i = 1, \dots, d$ (by convention $\ell_0 := 0$). The claim in (43) reduces to

$$\begin{aligned}&\mathbf{E}\big[f\big(C^{(n)}_{j_{\ell_1}}, \dots, C^{(n)}_{j_{\ell_1}}, C^{(n)}_{j_{\ell_2}}, \dots, C^{(n)}_{j_{\ell_d}}, b^{(n)}, \dots, b^{(n)}\big)\big]\\ &\quad= \mathbf{E}\big[f\big(D^{(n)}_{j_{\ell_1}}, \dots, D^{(n)}_{j_{\ell_1}}, D^{(n)}_{j_{\ell_2}}, \dots, D^{(n)}_{j_{\ell_d}}, b^{(n)}, \dots, b^{(n)}\big)\big].\end{aligned} \tag{46}$$

We will prove that, for each $p \in \{1, \dots, d\}$,

$$\mathbf{E}\big[f\big(C^{(n)}_{j_{\ell_1}}, \dots, C^{(n)}_{j_{\ell_{p-1}}}, C^{(n)}_{j_{\ell_p}}, \dots, C^{(n)}_{j_{\ell_p}}, D^{(n)}_{j_{\ell_{p+1}}}, \dots, D^{(n)}_{j_{\ell_d}}, b^{(n)}, \dots, b^{(n)}\big)\big] \tag{47}$$

$$\begin{aligned}&= \mathbf{E}\big[f\big(C^{(n)}_{j_{\ell_1}}, \dots, C^{(n)}_{j_{\ell_{p-1}}}, D^{(n)}_{j_{\ell_p}}, \dots, D^{(n)}_{j_{\ell_p}},\\ &\qquad D^{(n)}_{j_{\ell_{p+1}}}, \dots, D^{(n)}_{j_{\ell_d}}, b^{(n)}, \dots, b^{(n)}\big)\big],\end{aligned} \tag{48}$$

which in turn implies (46). Abbreviating $Y_i^{(r)} = (\mathbf{1}_{\{i<n_0\}} X_i^{(r)} + \mathbf{1}_{\{i \ge n_0\}} Z^{(r)})$ and denoting by Υ the joint distribution of $(A^{(n)}_{j_{\ell_1}}, \dots, A^{(n)}_{j_{\ell_d}}, I^{(n)}_{j_{\ell_1}}, \dots, I^{(n)}_{j_{\ell_d}}, b^{(n)})$ we have

$$\begin{aligned}&\mathbf{E}\big[f\big(C^{(n)}_{j_{\ell_1}}, \dots, C^{(n)}_{j_{\ell_{i-1}}}, C^{(n)}_{j_{\ell_i}}, \dots, C^{(n)}_{j_{\ell_i}}, D^{(n)}_{j_{\ell_{i+1}}}, \dots, D^{(n)}_{j_{\ell_d}}, b^{(n)}, \dots, b^{(n)}\big)\big]\\ &\quad= \int f(\alpha_1 x_1, \dots, \alpha_{p-1} x_{p-1}, \alpha_p x_p, \dots, \alpha_p x_p,\\ &\qquad\qquad \alpha_{p+1} x_{p+1}, \dots, \alpha_d x_d, b, \dots, b)\\ &\qquad \times d\mathbb{P}_{X_{i_1}}(x_1) \cdots d\mathbb{P}_{X_{i_p}}(x_p)\, d\mathbb{P}_{Y_{i_{p+1}}}(x_{p+1}) \cdots d\mathbb{P}_{Y_{i_d}}(x_d)\\ &\qquad \times d\Upsilon(\alpha_1, \dots, \alpha_d, i_1, \dots, i_d, b)\\ &\quad= \int \mathbf{E}\big[g(X_{i_p}, \dots, X_{i_p})\big]\, d\mathbb{P}_{X_{i_1}} \cdots d\mathbb{P}_{X_{i_{p-1}}}\, d\mathbb{P}_{Y_{i_{p+1}}} \cdots d\mathbb{P}_{Y_{i_d}}\, d\Upsilon,\end{aligned}$$

where, for all fixed $\alpha_1, \dots, \alpha_d, i_1, \dots, i_d, b, x_1, \dots, x_{p-1}, x_{p+1}, \dots, x_d$, we use the bounded and multilinear function $g : B^{\ell_p - \ell_{p-1}} \to \mathbb{R}$,

$$\begin{aligned}&g(y_1, \dots, y_{\ell_p - \ell_{p-1}})\\ &\quad:= f(\alpha_1 x_1, \dots, \alpha_{p-1} x_{p-1}, \alpha_p y_1, \dots, \alpha_p y_{\ell_p - \ell_{p-1}},\\ &\qquad\qquad \alpha_{p+1} x_{p+1}, \dots, \alpha_d x_d, b, \dots, b).\end{aligned}$$

Since $\mathcal{L}(X_m), \mathcal{L}(Z) \in \mathcal{M}_s(\eta)$ for all $m \ge n_0$ we can replace X_{i_p} by Y_{i_p}. This shows the equality (47), hence (43). Altogether, we obtain $\mathcal{L}(Q_n) \in \mathcal{M}_s(\eta)$ for all $n \ge n_0$.

Now, we show $T(\mu) \in \mathcal{M}_s(\eta)$. Let W be a random variable with distribution $T(\mu)$. By (C2), in particular, $\|A_r\|_s < \infty$ for $r = 1, \dots, K$, by (C1) we have $\|b\|_s < \infty$. Thus, as for Q_n, from Minkowski's inequality we obtain $\mathbf{E}[\|W\|_\infty^s] < \infty$, hence $T(\mu) \in \mathcal{M}_s(\eta)$ for $s \le 1$. For the case $s > 1$, we consider again arbitrary $1 \le k \le m$ and multilinear and bounded $f : B^k \to \mathbb{R}$. It suffices to show $\mathbf{E}[f(Q_n,\dots,Q_n)]=\mathbf{E}[f(W,\dots,W)]$ for some $n\ge n_0$. In fact, we will show that $\lim_{n\to\infty}\mathbf{E}[f(Q_n,\dots,Q_n)]=\mathbf{E}[f(W,\dots,W)]$. For this, we expand

$$\mathbf{E}\big[f(W,\dots,W)\big]=\mathbf{E}\left[f\left(\sum_{r=1}^K A_r Z^{(r)}+b,\dots,\sum_{r=1}^K A_r Z^{(r)}+b\right)\right]$$

into summands corresponding to (43) and have to show that

$$\lim_{n\to\infty}\mathbf{E}\big[f\big(D_{j_1}^{(n)},\dots,D_{j_k}^{(n)}\big)\big]=\mathbf{E}\big[f(E_{j_1},\dots,E_{j_k})\big], \tag{49}$$

where $j_1,\dots,j_k\in\{1,\dots,K\}$. For each $i\in\{1,\dots,k\}$, we have in case (44) that $E_{j_i}=A_{j_i}Z^{(j_i)}$, in case (45) that $E_{j_i}=b$. We obtain, introducing a telescoping sum and using Hölder's inequality,

$$\begin{aligned}
&\big|\mathbf{E}\big[f\big(D_{j_1}^{(n)},\dots,D_{j_k}^{(n)}\big)\big]-\mathbf{E}\big[f(E_{j_1},\dots,E_{j_k})\big]\big|\\
&\quad=\Bigg|\sum_{q=1}^k \mathbf{E}\big[f\big(E_{j_1},\dots,E_{j_{q-1}},D_{j_q}^{(n)},\dots,D_{j_k}^{(n)}\big)\\
&\qquad\qquad -f\big(E_{j_1},\dots,E_{j_q},D_{j_{q+1}}^{(n)},\dots,D_{j_k}^{(n)}\big)\big]\Bigg|\\
&\quad\le\sum_{q=1}^k\big|\mathbf{E}\big[f\big(E_{j_1},\dots,E_{j_{q-1}},D_{j_q}^{(n)}-E_{j_q},D_{j_{q+1}}^{(n)},\dots,D_{j_k}^{(n)}\big)\big]\big|\\
&\quad\le\sum_{q=1}^k\|f\|\big\|D_{j_q}^{(n)}-E_{j_q}\big\|_k\prod_{v=1}^{q-1}\|E_{j_v}\|_k\prod_{v=q+1}^{k}\big\|D_{j_v}^{(n)}\big\|_k.
\end{aligned}$$

Note that the $\|E_{j_v}\|_k$ and $\|D_{j_v}^{(n)}\|_k$ are all uniformly bounded by independence, (C1), and $\|X_0\|_s,\dots,\|X_{n_0-1}\|_s$, $\|Z\|_s<\infty$. Hence, it suffices to show that $\|D_{j_v}^{(n)}-E_{j_v}\|_k\to 0$ for all j_v. In case (45), this is $\|b^{(n)}-b\|_k\to 0$ by condition (C1). In case (45), we have, abbreviating $r=j_i$,

$$\begin{aligned}
&\big\|A_r^{(n)}\big(\mathbf{1}_{\{I_r^{(n)}<n_0\}}X_{I_r^{(n)}}^{(r)}+\mathbf{1}_{\{I_r^{(n)}\ge n_0\}}Z^{(r)}\big)-A_rZ^{(r)}\big\|_k\\
&\quad\le\big\|\big(A_r^{(n)}-A_r\big)Z^{(r)}\big\|_k+\big\|A_r^{(n)}\big(\mathbf{1}_{\{I_r^{(n)}<n_0\}}\big(X_{I_r^{(n)}}^{(r)}-Z^{(r)}\big)\big)\big\|_k.
\end{aligned}$$

The first summand of the latter display tends to zero by independence, $\|Z\|_s<\infty$ and condition (C1). The second summand tends to zero applying Hölder's inequality, condition (C1), which implies that $\|A_r^{(n)}\|_s$ in uniformly bounded, $\|X_0\|_s,\dots,\|X_{n_0-1}\|_s,\|Z\|_s<\infty$ and conditions (C1) and (C3). Altogether we obtain $T(\mu)\in\mathcal{M}_s(\eta)$.

Let $\mu, \lambda \in \mathcal{M}_s(\eta)$. Conditioning on the coefficients, using Lemma 1 and (11), it follows that

$$\zeta_s\big(T(\mu), T(\lambda)\big) \le \left(\sum_{r=1}^{K} \mathbf{E}\big[\|A_r\|_{\mathrm{op}}^s\big]\right)\zeta_s(\mu, \lambda).$$

Thus, by condition (C2), the restriction of T to $\mathcal{M}_s(\eta)$ is a contraction with respect to ζ_s.

Assume, μ was a fixed point of T as well. Then the contraction property implies

$$\zeta_s(\mu, \eta) = \zeta_s\big(T(\mu), T(\eta)\big) \le L\zeta_s(\mu, \eta),$$

hence $\zeta_s(\mu, \eta) = 0$. Since the ζ_s-distance is a metric on $\mathcal{M}_s(\eta)$ it follows $\mu = \eta$. □

We now turn to the problem of convergence of the sequence $(X_n)_{n\ge 0}$ to the fixed-point η.

Aiming to proof $X_n \to X$ condition (C1) is natural in the context of contraction method. Condition (C2) is necessary if working with ζ_s metrics. We will discuss this in detail for the cases $\mathcal{C}[0,1]$ and $\mathcal{D}[0,1]$ below. The existence of a solution of the fixed-point equation in condition (C3) is required since we miss knowledge about completeness of the ζ_s metrics. If $\mu \in \mathcal{M}_s(B)$, then $(T^n(\mu))_{n\ge 0}$ is a Cauchy sequence with respect to ζ_s, the proof being similar to the one of the previous lemma. Then, for $B = \mathcal{C}[0,1]$ or $B = \mathcal{D}[0,1]$, by Proposition 13, all finite-dimensional marginals of $T^n(\mu)$ converge to the corresponding marginals of some measure ν on $\mathbb{R}^{[0,1]}$, the natural candidate for a fixed-point of (41). In the application discussed in Section 4.2, the solution of the fixed-point equation (69) is constructed via a sequence $(Z_n)_{n\ge 0}$ of random variables that satisfy $\mathcal{L}(Z_n) = T^n(\mu)$ and converge uniformly almost surely (cf. [6] for details). The starting point is the Dirac measure $\mu = \delta_f$ with a specific function $f \in \mathcal{C}[0,1]$.

The following proposition uses the ideas developed so far to infer convergence of X_n to X in the ζ_s distance. The proof extends a similar proof for the case $B = \mathbb{R}^d$; see [22], Theorem 4.1. We draw further implications from this proof; see Corollary 21.

PROPOSITION 19. *Let $(X_n)_{n\ge 0}$ satisfy recurrence* (38) *with conditions* (C1)–(C3). *Then for the fixed-point $\eta = \mathcal{L}(X)$ of T in* (41) *we have, as $n \to \infty$,*

$$\zeta_s(X_n, X) \to 0.$$

PROOF. We use the accompanying sequence defined in (42). Throughout the proof, let $n \ge n_0$. Again since the ζ_s-distance is a metric, we have

$$\zeta_s(X_n, X) \le \zeta_s(X_n, Q_n) + \zeta_s(Q_n, X). \tag{50}$$

First, we consider the second term. By (C1) and Minkowski's inequality, absolute moments of order s of the sequence $(Q_n)_{n\ge n_0}$ are bounded, hence using Lemma 5 it suffices to show

$$\ell_s(Q_n, X) \to 0.$$

Using the same set of independent random variables $X^{(1)}, \dots, X^{(K)}$ for Q_n and in the recurrence of X, we obtain

$$\begin{aligned}
\ell_s(Q_n, X) &\le \Bigg\| \sum_{r=1}^K \big(A_r - \mathbf{1}_{\{I_r^{(n)} \ge n_0\}} A_r^{(n)}\big) X^{(r)} \Bigg\|_s + \Bigg\| \sum_{r=1}^K \mathbf{1}_{\{I_r^{(n)} < n_0\}} A_r^{(n)} X^{(r)}_{I_r^{(n)}} \Bigg\|_s \\
&\quad + \big\| b^{(n)} - b \big\|_s \\
&\le \sum_{r=1}^K \big(\| A_r - A_r^{(n)} \|_s + \| \mathbf{1}_{\{I_r^{(n)} < n_0\}} \| A_r^{(n)} \|_{\mathrm{op}} \|_s \big) \|X\|_s + \| b^{(n)} - b \|_s \\
&\quad + \Bigg\| \sum_{r=1}^K \mathbf{1}_{\{I_r^{(n)} < n_0\}} A_r^{(n)} X^{(r)}_{I_r^{(n)}} \Bigg\|_s .
\end{aligned}$$

By (C1) the first two summands tend to zero. Also, the third one converges to zero using (C1) and

$$\big\| \mathbf{1}_{\{I_r^{(n)} < n_0\}} \| A_r^{(n)} \|_{\mathrm{op}} X^{(r)}_{I_r^{(n)}} \big\|_s \le \big\| \mathbf{1}_{\{I_r^{(n)} < n_0\}} \| A_r^{(n)} \|_{\mathrm{op}} \big\|_s \Big\| \sup_{j < n_0} \|X_j\| \Big\|_s .$$

Furthermore, conditioning on the coefficients and using that ζ_s is $(s, +)$ ideal and Lemma 1, it is easy to see that

$$\zeta_s(Q_n, X_n) \le p_n \zeta_s(X_n, X) + \mathbf{E}\Bigg[\sum_{r=1}^K \mathbf{1}_{\{n_0 \le I_r^{(n)} \le n-1\}} \| A_r^{(n)} \|_{\mathrm{op}}^s \zeta_s(X_{I_r^{(n)}}, X) \Bigg] \tag{51}$$

$$\le p_n \zeta_s(X_n, X) + \Bigg(\sum_{r=1}^K \mathbf{E}\big[\| A_r^{(n)} \|_{\mathrm{op}}^s \big] \Bigg) \sup_{n_0 \le i \le n-1} \zeta_s(X_i, X), \tag{52}$$

where

$$p_n = \mathbf{E}\Bigg[\sum_{r=1}^K \mathbf{1}_{\{I_r^{(n)} = n\}} \| A_r^{(n)} \|_{\mathrm{op}}^s \Bigg] \to 0, \qquad n \to \infty.$$

Combining (50) and (52) implies

$$\zeta_s(X_n, X) \le \frac{1}{1 - p_n} \Bigg[\sum_{r=1}^K \mathbf{E}\big[\| A_r^{(n)} \|_{\mathrm{op}}^s \big] \sup_{n_0 \le i \le n-1} \zeta_s(X_i, X) + o(1) \Bigg].$$

From this, it follows that $\zeta_s(X_n, X)$ is bounded. Let

$$\bar{\eta} := \sup_{n \ge n_0} \zeta_s(X_n, X), \qquad \eta := \limsup_{n \to \infty} \zeta_s(X_n, X)$$

and $\varepsilon > 0$ arbitrary. Then there exists $\ell > 0$ with $\zeta_s(X_n, X) \le \eta + \varepsilon$ for all $n \ge \ell$. Using (50), (51) and splitting $\{n_0 \le I_r^{(n)} \le n-1\}$ into $\{n_0 \le I_r^{(n)} \le \ell\}$ and $\{\ell < I_r^{(n)} \le n-1\}$, we obtain

$$\zeta_s(X_n, X) \le \frac{\bar{\eta}}{1-p_n}\mathbf{E}\left[\sum_{r=1}^K \mathbf{1}_{\{n_0 \le I_r^{(n)} \le \ell\}} \|A_r^{(n)}\|_{\mathrm{op}}^s\right] + \frac{\eta+\varepsilon}{1-p_n}\mathbf{E}\left[\sum_{r=1}^K \|A_r^{(n)}\|_{\mathrm{op}}^s\right] + o(1),$$

which, by (C1), finally implies

$$\eta \le \mathbf{E}\left[\sum_{r=1}^K \|A_r\|_{\mathrm{op}}^s\right](\eta+\varepsilon).$$

Since $\varepsilon > 0$ is arbitrary and by condition (C2), we obtain $\eta = 0$. $\square$

REMARK 20. As pointed out in [13] for a related convergence result, the statements of Lemma 18 and Proposition 19 remain true if condition (C1) is weakened by replacing

$$\sum_{r=1}^K \|A_r^{(n)} - A_r\|_s \to 0$$

by

$$\sum_{r=1}^K \|(A_r^{(n)} - A_r)f\|_s \to 0, \qquad \|A_r^{(n)}\|_s \to \|A_r\|_s$$

for all $f \in \mathcal{C}[0,1]$ and uniform boundedness of $\|A_r^{(n)}\|_s$ for all $n \ge 0$ and all $r = 1, \dots, K$. This follows from the given independence structure and the dominated convergence theorem.

To be able to apply the results of the previous section to deduce weak convergence from convergence in ζ_s for the special cases $\mathcal{C}[0,1]$ and $\mathcal{D}[0,1]$, rates of convergence for ζ_s are required. We impose a further assumption on the convergence rate of the coefficients to establish a rate of convergence for the process that strengthens condition (C2). We use the Bachmann–Landau big-O notation for sequences of numbers.

(C4) The sequence $(\gamma(n))_{n \ge n_0}$ from condition (C1) satisfies $\gamma(n) = O(R(n))$ as $n \to \infty$ for some positive sequence $R(n) \downarrow 0$ such that

$$L^* = \limsup_{n\to\infty} \mathbf{E}\left[\sum_{r=1}^K \|A_r^{(n)}\|_{\mathrm{op}}^s \frac{R(I_r^{(n)})}{R(n)}\right] < 1.$$

COROLLARY 21. *Let $(X_n)_{n\geq 0}$ satisfy recurrence* (38) *with conditions* (C1), (C3) *and* (C4). *Then for the fixed-point $\eta = \mathcal{L}(X)$ of T in* (41) *we have, as $n \to \infty$,*

$$\zeta_s(X_n, X) = O\big(R(n)\big).$$

PROOF. We consider the quantities introduced in the proof of Proposition 19 again. By condition (C4), we have $\zeta_s(Q_n, X) \leq CR(n)$ for some $C > 0$ and all n. Furthermore, we can choose $\gamma > 0$ and $n_1 > 0$ such that

$$\mathbf{E}\Bigg[\sum_{r=1}^{K} \|A_r^{(n)}\|_{\mathrm{op}}^s \frac{R(I_r^{(n)})}{R(n)}\Bigg] \leq 1 - \gamma, \qquad p_n \leq \frac{\gamma}{2}$$

for $n \geq n_1$. Obviously, for any $n_2 \geq n_1$, we can choose $K \geq 2C/\gamma$ such that $d(n) := \zeta_s(X_n, X) \leq KR(n)$ for all $n < n_2$. Using (51), this implies

$$d(n_2) \leq p_{n_2} d(n_2) + \mathbf{E}\Bigg[\sum_{r=1}^{K} \mathbf{1}_{\{I_r^{(n_2)} \leq n_2 - 1\}} \|A_r^{(n_2)}\|_{\mathrm{op}}^s d\big(I_r^{(n_2)}\big)\Bigg] + CR(n_2)$$

hence

$$\begin{aligned}
d(n_2) &\leq \frac{1}{1 - p_{n_2}} \Bigg(\mathbf{E}\Bigg[\sum_{r=1}^{K} \|A_r^{(n_2)}\|_{\mathrm{op}}^s KR\big(I_r^{(n_2)}\big)\Bigg] + CR(n_2)\Bigg) \\
&= \frac{1}{1 - p_{n_2}} \Bigg(KR(n_2)\mathbf{E}\Bigg[\sum_{r=1}^{K} \|A_r^{(n_2)}\|_{\mathrm{op}}^s \frac{R(I_r^{(n_2)})}{R(n_2)}\Bigg] + CR(n_2)\Bigg) \\
&\leq \frac{1}{1 - p_{n_2}} \big((1 - \gamma)K + C\big)R(n_2) \leq KR(n_2).
\end{aligned}$$

Inductively, $d(n) \leq KR(n)$ for all n. $\square$

We now consider the special cases $\mathcal{C}[0,1]$ and $\mathcal{D}[0,1]$. Related to Corollary 10, we consider the following additional assumption, where the notation $\mathcal{C}_r[0,1]$ defined in (15) is used.

(C5) *Case* ($\mathcal{C}[0,1], \|\cdot\|_\infty$): we have $X_n = Y_n + h_n$ for all $n \geq 0$, where $\|h_n - h\|_\infty \to 0$ with $h_n, h \in \mathcal{C}[0,1]$, and there exists a positive sequence $(r_n)_{n\geq 0}$ such that

$$\mathbf{P}\big(Y_n \notin \mathcal{C}_{r_n}[0,1]\big) \to 0.$$

Case ($\mathcal{D}[0,1], d_{\mathrm{sk}}$): we have $X_n = Y_n + h_n$ for all $n \geq 0$, where $\|h_n - h\|_\infty \to 0$ with $h_n \in \mathcal{D}[0,1]$, $h \in \mathcal{C}[0,1]$, and there exists a positive sequence $(r_n)_{n\geq 0}$ such that

$$\mathbf{P}\big(Y_n \notin \mathcal{D}_{r_n}[0,1]\big) \to 0.$$

We now state the main theorem of this section. It follows immediately from Proposition 8, Corollary 10, Proposition 19 and Corollary 21.

THEOREM 22. *Let $(X_n)_{n\geq 0}$ be a sequence of random variables in* $(\mathcal{C}[0,1], \|\cdot\|_\infty)$ *or* $(\mathcal{D}[0,1], d_{\mathrm{sk}})$ *satisfying recurrence* (38) *with conditions* (C1), (C2), (C3) *being satisfied. Then, for $\mathcal{L}(X)=\eta$, we have for all $t\in[0,1]$*

$$X_n(t) \xrightarrow{d} X(t), \qquad \mathbf{E}\bigl[\bigl|X_n(t)\bigr|^s\bigr] \to \mathbf{E}\bigl[\bigl|X(t)\bigr|^s\bigr]. \tag{53}$$

If Z is distributed on $[0,1]$ *and independent of (X_n) and X then*

$$X_n(Z) \xrightarrow{d} X(Z), \qquad \mathbf{E}\bigl[\bigl|X_n(Z)\bigr|^s\bigr] \to \mathbf{E}\bigl[\bigl|X(Z)\bigr|^s\bigr]. \tag{54}$$

If moreover conditions (C4) *and* (C5) *are satisfied, where $R(n)$ in* (C4) *and r_n in* (C5) *can be chosen with*

$$R(n) = o\left(\frac{1}{\log^m(1/r_n)}\right), \qquad n\to\infty, \tag{55}$$

then we have convergence in distribution:

$$X_n \xrightarrow{d} X.$$

Finally, we give sufficient criteria to verify condition (C3) for the cases $\mathcal{C}[0,1]$ and $\mathcal{D}[0,1]$. First, consider the general case where $\mathcal{L}(Y)=\nu$ is a probability distribution on a separable Banach space $(B,\|\cdot\|)$ with $\mathbf{E}[\|Y\|^s]<\infty$. If B is a Hilbert space, it is easy to see (and already indicated in [32] for $m=2$) that for a probability measure $\mathcal{L}(X)=\mu$ on B to be in $\mathcal{M}_s(\nu)$ the defining properties (9) and (10) are equivalent to $\mathbf{E}[\|X\|^s]<\infty$ and

$$\mathbf{E}\bigl[\varphi_1(X)\cdots\varphi_k(X)\bigr] = \mathbf{E}\bigl[\varphi_1(Y)\cdots\varphi_k(Y)\bigr]$$

for all $0<k\leq m$ and continuous linear forms $\varphi_1,\dots,\varphi_n$ on B. A generalization of this equivalence to Banach spaces does not hold in general, a counterexample is constructed in Janson and Kaijser [16]. However, with deeper arguments from functional analysis, Janson and Kaijser [16] proved that this equivalence does hold for separable Banach spaces having the approximation property, such as $\mathcal{C}[0,1]$. The case $\mathcal{D}[0,1]$ is also treated in [16]. Combining (9), (10) and Theorems 1.3 and 16.13 in [16] implies the following lemma.

LEMMA 23. *Let $\mathcal{L}(Y)=\mathcal{L}((Y_t)_{t\in[0,1]})=\nu$ and $\mathcal{L}(X)=\mathcal{L}((X_t)_{t\in[0,1]})=\mu$ be probability measures on $\mathcal{C}[0,1]$. For $0<s\leq 1$ we have $\mu\in\mathcal{M}_s(\nu)$ if*

$$\mathbf{E}\bigl[\|X\|_\infty^s\bigr], \mathbf{E}\bigl[\|Y\|_\infty^s\bigr] < \infty. \tag{56}$$

For $1<s\leq 2$ we obtain $\mu\in\mathcal{M}_s(\nu)$ if we have condition (56) *and*

$$\mathbf{E}[X_t] = \mathbf{E}[Y_t] \qquad \text{for all } 0\leq t\leq 1. \tag{57}$$

For $2 < s \leq 3$ *we obtain* $\mu \in \mathcal{M}_s(\nu)$ *if we have conditions* (56), (57) *and*

$$\operatorname{Cov}(X_t, X_u) = \operatorname{Cov}(Y_t, Y_u) \qquad \textit{for all } 0 \leq t, u \leq 1. \tag{58}$$

The assertions remain true if $\mathcal{C}[0, 1]$ *is replaced by* $\mathcal{D}[0, 1]$.

REMARK 24. Interpreting $\mathbf{E}[X]$ as a Bochner integral in the continuous case, condition (57) is equivalent to $\mathbf{E}[X] = \mathbf{E}[Y]$. This is due to the fact that $\mathbf{E}[X]$ is a continuous function with $\mathbf{E}[X](t) = \mathbf{E}[X(t)]$ and $\varphi(\mathbf{E}[X]) = \mathbf{E}[\varphi(X)]$ for all continuous linear forms φ on $\mathcal{C}[0, 1]$. Also the higher moments can be interpreted similarly as expectations of corresponding tensor products; see [12] or, for an elaborate account [16].

REMARK 25. Note that condition (58) typically cannot be achieved for a sequence $(X_n)_{n\geq 0}$ that arises as in (2) by an affine scaling from a sequence $(Y_n)_{n\geq 0}$ as in (1). This fundamental problem for developing a functional contraction method on the basis of the Zolotarev metrics ζ_s with $2 < s \leq 3$ was already mentioned in [12], Remark 6.2. We describe a way to circumvent this problem in our application to Donsker's invariance principle by a perturbation argument; see Section 4.1.

4. Applications.

As applications, we first give as a toy example a short proof of Donsker's invariance principle in Section 4.1. In Section 4.2, we discuss further examples from the probabilistic analysis of algorithms on partial match queries which requires the full generality of our abstract setting. This allows to settle various long standing open questions about asymptotics of the complexity of such queries.

4.1. *Donsker's invariance principle.* Let $(V_n)_{n\in\mathbb{N}}$ be a sequence of independent, identically distributed real valued random variables with $\mathbf{E}[V_1] = 0$, $\operatorname{Var}(V_1) = 1$ (for simplicity) and $\mathbf{E}[|V_1|^{2+\varepsilon}] < \infty$ for some $\varepsilon > 0$. We consider the properly scaled and linearized random walk $S^n = (S^n_t)_{t\in[0,1]}$, $n \geq 1$, defined by

$$S^n_t = \frac{1}{\sqrt{n}}\left(\sum_{k=1}^{\lfloor nt\rfloor} V_k + \bigl(nt - \lfloor nt\rfloor\bigr)V_{\lfloor nt\rfloor+1}\right), \qquad t \in [0, 1].$$

With $W = (W_t)_{t\in[0,1]}$, a standard Brownian motion Donsker's function limit law states the following.

THEOREM 26 (Donsker [11]). *We have* $S^n \xrightarrow{d} W$ *as* $n \to \infty$ *in* $(\mathcal{C}[0, 1], \|\cdot\|_\infty)$.

4.1.1. *A contraction proof.* In this section, we apply the general methodology of Sections 2 and 3 to give a short proof of Theorem 26. For a recursive decomposition of S^n and W, we define operators for $\beta > 1$,

$$\varphi_\beta : \mathcal{C}[0,1] \to \mathcal{C}[0,1], \qquad \varphi_\beta(f)(t) = \mathbf{1}_{\{t \le 1/\beta\}} f(\beta t) + \mathbf{1}_{\{t > 1/\beta\}} f(1),$$
$$\psi_\beta : \mathcal{C}[0,1] \to \mathcal{C}[0,1], \qquad \psi_\beta(f)(t) = \mathbf{1}_{\{t \le 1/\beta\}} f(0) + \mathbf{1}_{\{t > 1/\beta\}} f\left(\frac{\beta t - 1}{\beta - 1}\right).$$

Note that both φ_β and ψ_β are linear, continuous and $\|\varphi_\beta(f)\|_\infty = \|\psi_\beta(f)\|_\infty = \|f\|_\infty$ for all $f \in \mathcal{C}[0,1]$, hence we have $\|\varphi_\beta\|_{\mathrm{op}} = \|\psi_\beta\|_{\mathrm{op}} = 1$. By construction, we have

$$S^n \stackrel{d}{=} \sqrt{\frac{\lceil n/2 \rceil}{n}} \varphi_{n/\lceil n/2 \rceil}(S^{\lceil n/2 \rceil}) + \sqrt{\frac{\lfloor n/2 \rfloor}{n}} \psi_{n/\lceil n/2 \rceil}(\widehat{S}^{\lfloor n/2 \rfloor}), \qquad n \ge 2, \tag{59}$$

where $(S^1, \dots, S^n)$ and $(\widehat{S}^1, \dots, \widehat{S}^n)$ are independent and S^j and $\widehat{S}^j$ are identically distributed for all $j \ge 1$. Therefore, $(S^n)_{n \ge 1}$ satisfies recurrence (38) choosing

$$K = 2, \qquad I_1^{(n)} = \lceil n/2 \rceil, \qquad I_2^{(n)} = \lfloor n/2 \rfloor, \qquad n_0 = 2,$$
$$A_1^{(n)} = \sqrt{\frac{\lceil n/2 \rceil}{n}} \varphi_{n/\lceil n/2 \rceil}, \qquad A_2^{(n)} = \sqrt{\frac{\lfloor n/2 \rfloor}{n}} \psi_{n/\lceil n/2 \rceil}, \qquad b^{(n)} = 0.$$

In the following, let $\widehat{W} = (\widehat{W}_t)_{t \in [0,1]}$ be a standard Brownian motion, independent of W. Properties of Brownian motion imply

$$W \stackrel{d}{=} \sqrt{\frac{1}{\beta}} \varphi_\beta(W) + \sqrt{\frac{\beta - 1}{\beta}} \psi_\beta(\widehat{W}) \tag{60}$$

for any $\beta > 1$. Hence, the Wiener measure $\mathcal{L}(W)$ is a fixed point of the operator T in (41) with

$$K = 2, \qquad A_1 = \sqrt{\frac{1}{\beta}} \varphi_\beta, \qquad A_2 = \sqrt{\frac{\beta - 1}{\beta}} \psi_\beta, \qquad b = 0. \tag{61}$$

For $\beta = 2$, the coefficients in (59) converge to the ones in (60), that is, as $n \to \infty$,

$$\sqrt{\frac{\lceil n/2 \rceil}{n}} \to \frac{1}{\sqrt{2}}, \qquad \sqrt{\frac{\lfloor n/2 \rfloor}{n}} \to \frac{1}{\sqrt{2}},$$

but the coefficients $A_1^{(n)}, A_2^{(n)}$ only converge to A_1, A_2 in the operator norm for n even. Nevertheless, from the point of view of the contraction method, this suggests weak convergence of S^n to W.

Note that the operator T associated with the fixed-point equation (60), that is, with the coefficients in (61), satisfies condition (C2) only with $s > 2$. In view of condition (C3) and Lemma 23, we need to match the mean and covariance structure. We have $\mathbf{E}[S_t^n] = 0$ for all $0 \le t \le 1$ and a direct computation yields

$$\operatorname{Cov}(S_s^n, S_t^n) = \begin{cases} s, & \text{for } \lfloor ns \rfloor < \lfloor nt \rfloor, \\ \dfrac{1}{n}\big(\lfloor ns \rfloor + (ns - \lfloor ns \rfloor)(nt - \lfloor nt \rfloor)\big), & \text{for } \lfloor ns \rfloor = \lfloor nt \rfloor. \end{cases} \tag{62}$$

Hence, we do not have finite $\zeta_{2+\varepsilon}$-distance between S^n and W since they do not share their covariance functions. To surmount this problem, we consider a linearized version of the Brownian motion W. For fixed $n \in \mathbb{N}$, we divide the unit interval into pieces of length $1/n$ and interpolate W linearly between the points $0, 1/n, 2/n, \dots, (n-1)/n, 1$. The interpolated process $W^n = (W_t^n)_{t \in [0,1]}$ is given by

$$W_t^n := W_{\lfloor nt \rfloor / n} + (nt - \lfloor nt \rfloor)(W_{(\lfloor nt \rfloor + 1)/n} - W_{\lfloor nt \rfloor / n}), \qquad t \in [0, 1].$$

We have $\mathbf{E}[W_t^n] = 0$ and W^n and S^n have the same covariance function (62) for all $n \in \mathbb{N}$. Furthermore, W^n has the same distributional recursive decomposition (59) as S^n.

Note that the linearized Brownian motion does not differ much from the original one:

LEMMA 27. *We have $\|W^n - W\|_\infty \to 0$ as $n \to \infty$ almost surely.*

PROOF. This directly follows from the uniform continuity of W. For $\varepsilon > 0$, there exists a random $\delta > 0$ such that $|W(t) - W(s)| < \varepsilon$ for any $s, t \in [0, 1]$ with $|t - s| < \delta$. The triangle inequality implies $\|W^n - W\|_\infty < 2\varepsilon$ for any $n > 1/\delta$. $\square$

In view of Corollary 11, it suffices to prove that S^n and W^n are close with respect to $\zeta_{2+\varepsilon}$. The proof of this runs along the same lines as the one for Proposition 19, respectively, Corollary 21; in fact, it is much shorter due to the simple form of the recurrence:

PROPOSITION 28. *For any $\delta < \varepsilon/2$ we have $\zeta_{2+\varepsilon}(S^n, W^n) = O(n^{-\delta})$ as $n \to \infty$.*

PROOF. We have

$$\begin{aligned} \zeta_{2+\varepsilon}(S^n, W^n) = \zeta_{2+\varepsilon}\Bigg(&\sqrt{\frac{\lceil n/2 \rceil}{n}}\varphi_{n/\lceil n/2 \rceil}(S^{\lceil n/2 \rceil}) + \sqrt{\frac{\lfloor n/2 \rfloor}{n}}\psi_{n/\lceil n/2 \rceil}(\overline{S}^{\lfloor n/2 \rfloor}), \\ &\sqrt{\frac{\lceil n/2 \rceil}{n}}\varphi_{n/\lceil n/2 \rceil}(W^{\lceil n/2 \rceil}) + \sqrt{\frac{\lfloor n/2 \rfloor}{n}}\psi_{n/\lceil n/2 \rceil}(\overline{W}^{\lfloor n/2 \rfloor})\Bigg) \end{aligned}$$

$$\leq \left(\frac{\lceil n/2\rceil}{n}\right)^{1+\varepsilon/2} \zeta_{2+\varepsilon}\big(S^{\lceil n/2\rceil}, W^{\lceil n/2\rceil}\big) + \left(\frac{\lfloor n/2\rfloor}{n}\right)^{1+\varepsilon/2} \zeta_{2+\varepsilon}\big(S^{\lfloor n/2\rfloor}, W^{\lfloor n/2\rfloor}\big).$$

We abbreviate

$$d_n := \zeta_{2+\varepsilon}\big(S^n, W^n\big), \qquad a_n := \left(\frac{\lceil n/2\rceil}{n}\right)^{1+\varepsilon/2}, \qquad b_n := \left(\frac{\lfloor n/2\rfloor}{n}\right)^{1+\varepsilon/2}$$

and note that we have $a_n + b_n \leq 2^{-\varepsilon/2} + C'/n$ for some constant $C' > 0$ and all $n \in \mathbb{N}$. For arbitrary $\delta < \varepsilon/2$, we prove the assertion by induction: fix $\delta < \delta' < \varepsilon/2$ and choose $m_0 \in \mathbb{N}$ such that $\lfloor n/2\rfloor^{-\delta} \leq (n/2)^{-\delta} 2^{\varepsilon/2-\delta'}$ and $1 + 2^{\varepsilon/2} C'/n \leq 2^{\delta'-\delta}$ for all $n \geq m_0$. Furthermore, let $C > 0$ be large enough such that $d_n \leq C n^{-\delta}$ for all $1 \leq n \leq m_0$. Then, for $n > m_0$, assuming the claim to be verified for all smaller indices,

$$\begin{aligned} d_n &\leq a_n d_{\lceil n/2\rceil} + b_n d_{\lfloor n/2\rfloor} \\ &\leq C\big(a_n (n/2)^{-\delta} + b_n (n/2)^{-\delta} 2^{\varepsilon/2-\delta'}\big) \\ &\leq C n^{-\delta} 2^{\delta} 2^{\varepsilon/2-\delta'} (a_n + b_n) \\ &\leq C n^{-\delta}. \end{aligned}$$

The assertion follows. $\square$

Now Donsker's theorem (Theorem 26) follows from Proposition 28, Lemma 27 and Corollary 11.

Note that our approach requires the assumption $\mathbf{E}[|V_1|^{2+\varepsilon}] < \infty$ for some $\varepsilon > 0$, which in Donsker's theorem can be weakened to $\mathbf{E}[V_1^2] < \infty$.

By Theorem 12, we directly obtain convergence of moments of the supremum.

COROLLARY 29. *Suppose $\mathbf{E}[|V_1|^{2+\alpha}] < \infty$ with $0 < \alpha \leq 1$. Then $\|S^n\|_\infty^{2+\alpha}$ is uniformly integrable. Thus, $\mathbf{E}[\|S^n\|_\infty^\kappa]$ converges to $\mathbf{E}[\|W\|_\infty^\kappa]$ for any $0 < \kappa \leq 2 + \alpha$.*

REMARK 30. Based on the recursion (59), it is easy to show that $\mathbf{E}[\|S^n\|_\infty^k]$ is bounded uniformly in n for integer valued $k \geq 3$ if the increment V_1 has finite absolute moment of order k. In this case, we have $\mathbf{E}[\|S^n\|_\infty^\kappa] \to \mathbf{E}[\|W\|_\infty^\kappa]$ for any real $0 < \kappa < k$.

4.1.2. *Characterizing the Wiener measure by a fixed-point property.* We reconsider the map T corresponding to the fixed-point equation (60) for the case $\beta = 2$:

$$T : \mathcal{M}\big(\mathcal{C}[0,1]\big) \to \mathcal{M}\big(\mathcal{C}[0,1]\big), \qquad T(\mu) = \mathcal{L}\left(\frac{1}{\sqrt{2}}\varphi_2(Z) + \frac{1}{\sqrt{2}}\psi_2(\overline{Z})\right), \tag{63}$$

where $Z, \overline{Z}$ are independent with distribution $\mathcal{L}(Z) = \mathcal{L}(\overline{Z}) = \mu$. Our discussion above implies that the Wiener measure $\mathcal{L}(W)$ is the unique fixed point of T restricted to $\mathcal{M}_{2+\varepsilon}(\mathcal{L}(W))$ for any $\varepsilon > 0$. Note that $\mathcal{M}_{2+\varepsilon}(\mathcal{L}(W))$ is the space of the distributions of all continuous stochastic processes $V = (V_t)_{t\in[0,1]}$ with $\mathbf{E}[\|V\|_\infty^{2+\varepsilon}] < \infty$, $\mathbf{E}[V_t] = 0$ and $\mathrm{Cov}(V_t, V_u) = t \wedge u$ for all $0 \le t, u \le 1$. Note that one easily verifies that $T(\mathcal{M}_{2+\varepsilon}(\mathcal{L}(W))) \subset \mathcal{M}_{2+\varepsilon}(\mathcal{L}(W))$ and the last part of the proof of Lemma 18 implies that T restricted to $\mathcal{M}_{2+\varepsilon}(\mathcal{L}(W))$ is Lipschitz-continuous with Lipschitz constant at most $L = 2^{-\varepsilon/2} < 1$, hence $\mathcal{L}(W)$ is the unique fixed point of T in $\mathcal{M}_{2+\varepsilon}(\mathcal{L}(W))$.

We now show that a more general statement is true, the Wiener measure is also, up to multiplicative scaling, the unique fixed point of T in the larger space of probability measures $\mathcal{L}(V) \in \mathcal{M}(\mathcal{C}[0,1])$ with $V_0 = 0$. For a related statement, see also Aldous [1], page 528. The subsequent proof is based on the fact that the centered normal distributions are the only solutions of the fixed-point equation

$$X \stackrel{d}{=} \frac{X + \overline{X}}{\sqrt{2}}, \tag{64}$$

where $X, \overline{X}$ are independent, identically distributed real-valued random variables; see Theorem 7.2.1 in [19].

THEOREM 31. *Let $X = (X_t)_{t\in[0,1]}$ be a continuous process with $X_0 = 0$. Then $\mathcal{L}(X)$ is a fixed-point of (63) if and only if either $X = \mathbf{0}$ a.s. or there exists a constant $\sigma > 0$, such that $(\sigma^{-1}X_t)_{t\in[0,1]}$ is a standard Brownian motion.*

PROOF. Let $\mathcal{L}(X)$ be a fixed point of (63) and $\overline{X} = (\overline{X}_t)_{t\in[0,1]}$ be independent of X with the same distribution. The fixed point property implies

$$X_1 \stackrel{d}{=} \frac{X_1 + \overline{X}_1}{\sqrt{2}},$$

hence $\mathcal{L}(X_1) = \mathcal{N}(0, \sigma^2)$ for some $\sigma^2 \ge 0$, where $\mathcal{N}(0, \sigma^2)$ denotes the centered normal distribution with variance σ^2. This implies

$$X_{1/2} \stackrel{d}{=} \frac{X_1}{\sqrt{2}},$$

hence $\mathcal{L}(X_{1/2}) = \mathcal{N}(0, \sigma^2/2)$. Let $\mathscr{D} = \{m2^{-n} : m, n \in \mathbb{N}_0, m \le 2^n\}$ by the set of dyadic numbers in $[0,1]$. By induction, we obtain $\mathcal{L}(X_t) = \mathcal{N}(0, \sigma^2 t)$ for all $t \in \mathscr{D}$. For the distribution of the increments, we first obtain

$$X_1 - X_{1/2} \stackrel{d}{=} \frac{X_1}{\sqrt{2}},$$

hence $\mathcal{L}(X_1 - X_{1/2}) = \mathcal{N}(0, \sigma^2/2)$. Again inductively, we obtain $\mathcal{L}(X_1 - X_t) = \mathcal{N}(0, (1-t)\sigma^2)$ for all $t \in \mathscr{D}$. Also by induction, it follows $\mathcal{L}(X_t - X_s) =$

$\mathcal{N}(0, (t-s)\sigma^2)$ for all $s, t \in \mathscr{D}$ with $s < t$. Finally, continuity of X implies the same property for all $s, t \in [0, 1]$. It remains to prove independence of increments. Denoting by $X^{(1)}, X^{(2)}, \ldots$ independent distributional copies of X, we obtain from iterating the fixed-point property

$$(X_t)_{t\in[0,1]} \stackrel{d}{=} \left(2^{-n/2} \sum_{m=1}^{2^n} \mathbf{1}_{\{(m-1)2^{-n} < t \le m2^{-n}\}} X^{(m)}_{2^n t - m + 1} + \mathbf{1}_{\{m2^{-n} < t\}} X^{(m)}_1\right)_{t\in[0,1]}$$

for all $n \in \mathbb{N}$. Hence, for any dyadic points $0 \le t_1 < t_2 < \cdots < t_k \le 1$, choosing n large enough, each $X_{t_{i+1}} - X_{t_i}$ can be expressed as a function of a subset of $X^{(1)}, \ldots, X^{(2^n)}$ these subsets being pairwise disjoint for $i = 0, \ldots, n-1$. Since, $\mathscr{D}$ is dense in $[0, 1]$, this shows that X has independent increments. For $\sigma = 0$, we have $X = \mathbf{0}$ a.s., otherwise $\sigma^{-1} X$ is a standard Brownian motion.

The converse direction of the theorem is trivial. $\square$

REMARK 32. Note that we cannot cancel the assumption on continuity of X without replacement, for example, the process

$$Y_t = \begin{cases} W_t, & t \notin \mathscr{D}, \\ 0, & t \in \mathscr{D} \end{cases}$$

also solves (60) and is not a multiple of Brownian motion. However, it would be sufficient to require càdlàg paths, so $\mathcal{C}[0, 1]$ could be replaced by $\mathcal{D}[0, 1]$ in our statement.

REMARK 33. Our decomposition of Brownian motion in (60) is in time. However, equation (64) suggests to also investigate a decomposition in space

$$(X_t)_{t\in[0,1]} \stackrel{d}{=} \left(\frac{X_t + \overline{X}_t}{\sqrt{2}}\right)_{t\in[0,1]}, \tag{65}$$

where $(X_t)_{t\in[0,1]}$ and $(\overline{X}_t)_{t\in[0,1]}$ are independent and identically distributed. Again, equation (65) induces a map on $\mathcal{M}(\mathcal{C}[0, 1])$ that is a contraction in $\zeta_{2+\varepsilon}$ on the subspace $\mathcal{M}_{2+\varepsilon}(\mathcal{L}(W))$, so the Wiener measure is the only solution in $\mathcal{M}_{2+\varepsilon}(\mathcal{L}(W))$. In this case, we cannot remove the moment assumption as in Theorem 31 since any centered, continuous Gaussian process solves equation (65). Using (64), it is not hard to see that there are no further solutions of (65).

4.2. *Partial match queries in quad trees*. In this section, we outline recurrences coming up in the probabilistic analysis of the performance of data structures and discuss in detail the use and verification of our conditions (C1)–(C5) and Theorem 22. In this example, the full generality of our setup is needed.

For preprocessing and supporting search queries in multidimensional data various types of search trees are in use, most prominently quad trees and k-d trees. Among various other fundamental search operations in multivariate data so-called partial match queries are of particular importance. For a partial match query, one specifies some of the components of the data and asks to report all data in the given set that match the specified components and are arbitrary in the remaining components. We will subsequently not need to introduce these data structures and the partial match queries since there is a geometric reformulation that is discussed and used below. For details about the computer science background and precise definition of the structures and queries, see [6].

Consider a sequence $(U_i, V_i)_{i\geq 1}$ of independent and identically distributed random vectors all with the uniform distribution on the unit square $[0, 1]^2$. We iteratively construct a decomposition of $[0, 1]^2$ as follows. The first point (U_1, V_1) decomposes the square into four rectangles by drawing the two lines through (U_1, V_1) in $[0, 1]^2$ that are perpendicular to its sides. We call these line segments the horizontal and vertical lines. The second point (U_2, V_2) almost surely falls into the interior of one of the four rectangles. We recursively draw the horizontal and vertical lines through (U_2, V_2) within the rectangle. Hence, we then have a decomposition of the original square $[0, 1]^2$ into seven rectangles. Now we iterate this process. After $n - 1$ steps, we have $3(n - 1) + 1$ rectangles and the nth point is used to decompose the rectangle it falls in into four new rectangle by the horizontal and vertical lines through it; see Figure 1. We identify this decomposition of the unit square with all the line segments drawn and call it the *decomposition after n steps*.

Now fix $t \in [0, 1]$ and denote the number of horizontal lines in the decomposition after n steps that are cut by the vertical line $x_1 = t$ by $C_n(t)$; see Figures 1 and 2.

In the computer science setting, this is the measure for the complexity of a partial match query in a random (point) quad tree where the first component is specified as t, the second component is arbitrary and n data are inserted in the uniform model; see [6]. We have $C_0(t) = 0$ and $C_1(t) = 1$ for all $t \in [0, 1]$. We consider the process $(C_n(t))_{t\in[0,1]}$ as a process in $(\mathcal{D}[0, 1], d_{\mathrm{sk}})$.

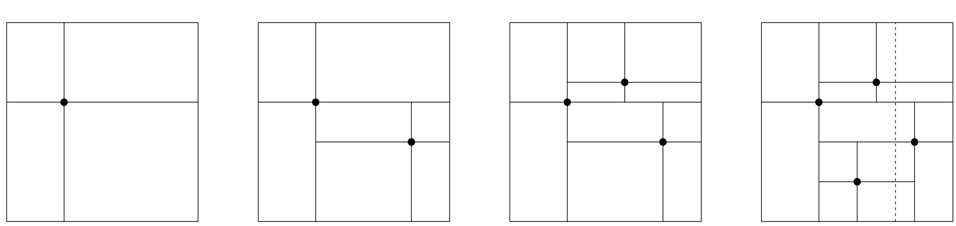

FIG. 1. *The construction of a quad tree at times* $n = 1, 2, 3, 4$. *The dashed line in the right most square indicates the query line* $x_1 = t$.

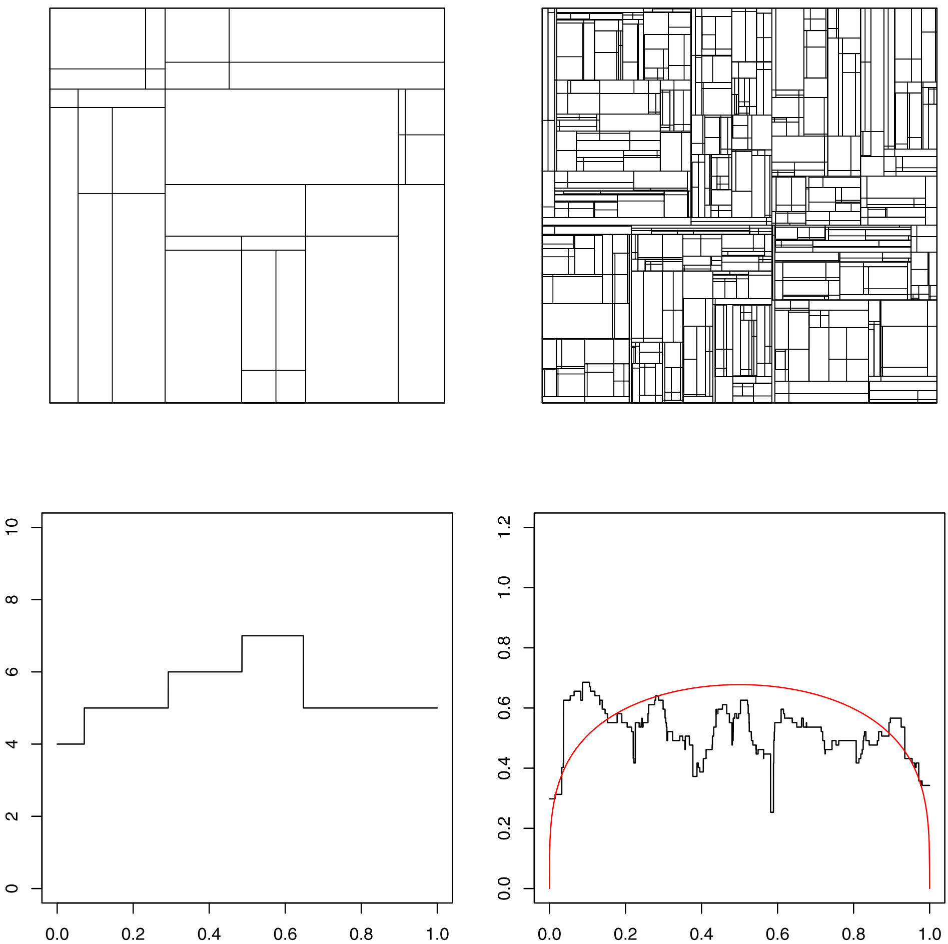

FIG. 2. *Left column: A realization of the decomposition of the unit square for a quad tree of size* $n = 10$ *and its process* $(C_n(t))_{t\in[0,1]}$. *Right column: A realization of the decomposition of the unit square for a quad tree of size* $n = 300$ *and the process* $(X_n(t))_{t\in[0,1]}$. *The smooth curve indicates the function* $t \mapsto (t(1-t))^{\beta/2}$.

For a recursive decomposition of this process, we denote the numbers of points among the first n points which fall into each of the four rectangles generated by the first point $(U_1, V_1) =: (U, V)$ by $I^{(n)} = (I_1^{(n)}, I_2^{(n)}, I_3^{(n)}, I_4^{(n)})$. Hence, conditionally on (U, V), the vector $I^{(n)}$ has the multinomial distribution $M(n-1; UV, U(1-V), (1-U)V, (1-U)(1-V))$, where a numbering of the four quadrants is used. Moreover, conditionally on (U, V) and $I^{(n)}$ we have that each point set within a rectangle is a set of independent and identically distributed points each with the uniform distribution on the particular rectangle and that the four point sets are also independent. Hence, for processes $(C_j^{(r)}(t))_{t\in[0,1]}$ which are independent and independent of $(U, V, I^{(n)})$, and $(C_j^{(r)}(t))_{t\in[0,1]}$ distributed as $(C_j(t))_{t\in[0,1]}$ for $r = 1, \dots, 4$ and $j \in \mathbb{N}_0$ we obtain the recurrence

$$\begin{aligned}(C_n(t))_{t\in[0,1]} \stackrel{d}{=} \biggl(1 + \mathbf{1}_{\{t<U\}}\biggl[C^{(1)}_{I_1^{(n)}}\Bigl(\frac{t}{U}\Bigr) + C^{(2)}_{I_2^{(n)}}\Bigl(\frac{t}{U}\Bigr)\biggr] \\ + \mathbf{1}_{\{t\ge U\}}\biggl[C^{(3)}_{I_3^{(n)}}\Bigl(\frac{t-U}{1-U}\Bigr) + C^{(4)}_{I_4^{(n)}}\Bigl(\frac{t-U}{1-U}\Bigr)\biggr]\biggr)_{t\in[0,1]}.\end{aligned} \tag{66}$$

The arguments t/U and $(1-t)/(1-U)$ adjust that a vertical line $x_1 = t$ within the whole square $[0, 1]^2$, after scaling, corresponds to the line $x_1 = t/U$ in the left rectangles (if $t < U$) and to the line $x_1 = (1-t)/(1-U)$ in the right rectangles (if $t \ge U$). Note that equation (66) has exactly the form (1), where the indicators and rescalings in time in (66) give the random linear maps $A_r(n)$ for $r = 1, \dots, 4$, and we have $b(n) = 1$.

The first asymptotic analysis of this process was done by Flajolet et al. [14], where the one-dimensional averaged complexity $C_n(\xi)$ was considered with ξ uniformly distributed on [0, 1] and independent of the sequence $(U_i, V_i)_{i\in\mathbb{N}}$. In [14], is shown that, as $n \to \infty$,

$$\mathbb{E}[C_n(\xi)] \sim \kappa n^\beta \qquad \text{with } \kappa = \frac{\Gamma(2\beta+2)}{2(\Gamma(\beta+1))^3}, \beta = \frac{\sqrt{17}-3}{2},$$

where Γ denotes the gamma function; see also Chern and Hwang [8] for more refined analysis of this expectation. Recently, Curien and Joseph [9] showed

$$\mathbb{E}[C_n(t)] \sim \chi\bigl(t(1-t)\bigr)^{\beta/2} n^\beta \qquad \text{with } \chi = \frac{\kappa}{B((\beta/2)+1, (\beta/2)+1)}, \tag{67}$$

where $B(\cdot,\cdot)$ denotes the beta function (Euler integral). The analysis beyond expectations, in particular of variances and limit laws either for the process $(C_n(t))_{t\in[0,1]}$ itself or its marginals or the averaged complexity $C_n(\xi)$ or the worst case complexity $\sup_{t\in[0,1]} C_n(t)$ remained open.

We now discuss how our general framework from Section 3 can be applied to a proper normalization of $(C_n(t))_{t\in[0,1]}$ and highlight the use and verification of conditions (C1)–(C5), which can be shown to hold with the choice $s = 2$. The details are worked out in [6]. The resulting functional limit law allows to settle the open questions raised in the previous paragraph.

Let us first use the normalization $X_0(t) := 0$ and

$$X_n(t) := \frac{C_n(t)}{\chi n^\beta}, \qquad n \ge 1, t \in [0, 1] \tag{68}$$

and write $X_n := (X_n(t))_{t\in[0,1]}$. See Figure 2 for a simulation of X_n. For X_n, we obtain the recurrence

$$\begin{aligned}X_n \stackrel{d}{=} \biggl(\frac{1}{\chi n^\beta} + \mathbf{1}_{\{t<U\}}\biggl[\Bigl(\frac{I_1^{(n)}}{n}\Bigr)^\beta X^{(1)}_{I_1^{(n)}}\Bigl(\frac{t}{U}\Bigr) + \Bigl(\frac{I_2^{(n)}}{n}\Bigr)^\beta X^{(2)}_{I_2^{(n)}}\Bigl(\frac{t}{U}\Bigr)\biggr] \\ + \mathbf{1}_{\{t\ge U\}}\biggl[\Bigl(\frac{I_3^{(n)}}{n}\Bigr)^\beta X^{(3)}_{I_3^{(n)}}\Bigl(\frac{t-U}{1-U}\Bigr) + \Bigl(\frac{I_4^{(n)}}{n}\Bigr)^\beta X^{(4)}_{I_4^{(n)}}\Bigl(\frac{t-U}{1-U}\Bigr)\biggr]\biggr)_{t\in[0,1]}\end{aligned}$$

with assumptions on independence and identical distributions as in (66). This suggests that a limit process $X = (X(t))_{t\in[0,1]}$ satisfies

$$
\begin{aligned}
X \stackrel{d}{=} \Bigg(&\mathbf{1}_{\{t<U\}}\bigg[(UV)^{\beta}X^{(1)}\Big(\frac{t}{U}\Big) + \big(U(1-V)\big)^{\beta}X^{(2)}\Big(\frac{t}{U}\Big)\bigg] \\
&+ \mathbf{1}_{\{t\geq U\}}\bigg[\big((1-U)V\big)^{\beta}X^{(3)}\Big(\frac{t-U}{1-U}\Big) \\
&\qquad + \big((1-U)(1-V)\big)^{\beta}X^{(4)}\Big(\frac{t-U}{1-U}\Big)\bigg]\Bigg)_{t\in[0,1]},
\end{aligned}
\tag{69}
$$

where U and V are independent $[0,1]$-uniform random variables and $(X^{(r)}(t))_{t\in[0,1]}$, for $r = 1,\ldots,4$, are independent copies of the process X, also independent of (U, V). Note that (69) is a fixed-point equation of type (4).

This heuristic derivation of equation (69) can be turned into a rigorous approach as follows. First, note that the operators $A_1^{(n)}$ and A_1 on $\mathcal{D}[0,1]$ are given as follows: for $f \in \mathcal{D}[0,1]$, the random functions $A_1^{(n)}(f)$ and $A_1(f)$ are

$$
t \mapsto \mathbf{1}_{\{t<U\}}\bigg(\frac{I_1^{(n)}}{n}\bigg)^{\beta} f\Big(\frac{t}{U}\Big) \quad \text{and} \quad t \mapsto \mathbf{1}_{\{t<U\}}(UV)^{\beta} f\Big(\frac{t}{U}\Big)
\tag{70}
$$

and direct integration shows that condition (C2) is satisfied for the choice $s = 2$.

For condition (C3) first an appropriate process $X = (X(t))_{t\in[0,1]}$ which solves (69) has to be constructed. Since we do not know the completeness of ζ_2 on an appropriate subspace of $\mathcal{M}_2(\mathcal{D}[0,1])$ and also are not able to guess X as a well-known process (as in the example in Section 4.1.1) such a process X has to be constructed individually. In view of (67), the normalization (68), the choice $s = 2$ and Lemma 23 we additionally need to have $\mathbf{E}[\|X\|_\infty^2] < \infty$ and $\mathbf{E}[X(t)] = (t(1-t))^{\beta/2}$ for $t \in [0,1]$. In [6], a sequence of random continuous functions is constructed from a discrete recurrence approximating (69) which converges uniformly. The construction uses concentration inequalities and tail bounds for the saturation level of random quad trees. Its limit X is the stochastic process as needed. Moreover, it can also be shown that it has continuous paths almost surely.

Our normalization does not imply that $\mathcal{L}(X_n) \in \mathcal{M}_2(\mathcal{L}(X))$ for all $n \geq 1$, since the normalization in (68) does violate condition (57). Thus, the processes X_n cannot be compared with X using the ζ_2 distance. To overcome this technical issue, one can instead consider the normalization

$$
\frac{C_n(t) - \mathbf{E}[C_n(t)]}{\chi n^{\beta}}, \qquad t \in [0,1], n \geq 1
\tag{71}
$$

and the shifted limit $(X(t) - (t(1-t))^{\beta/2})_{t\in[0,1]}$. Then condition (C3) is satisfied. This also shows the necessity to allow the perturbation h_n in Corollary 10 and condition (C5) in our general setup. The centering of the sequence X_n and the solution X only affects the additive term $b^{(n)}$ and the toll term b. In particular, condition (C2) remains valid in the centered setting and we have $\|A_1^{(n)} - A_1\|_s \to 0$ for any $s > 0$. Similarly, $\|A_r^{(n)} - A_r\|_s \to 0$ for $r = 2, 3, 4$. Convergence of the additive term $b^{(n)}$ is equivalent to uniformity of the expansion in (67). This is shown in [6]. It is also easily seen that (40) holds, hence condition (C1) is true.

For condition (C4), an appropriated rate of convergence of the coefficients in (39) is needed. Note that such a rate can only be derived if a rate in the asymptotic expansion of the means in (67) is available. Hence, as a technical step in [6] a polynomial additive error term of the order $O(n^{\beta-\varepsilon})$ for some $\varepsilon > 0$ is shown to hold valid uniformly in $t \in [0, 1]$. This implies that the convergence rates $\gamma(n)$ in (39) satisfy $\gamma(n) = O(n^{-\varepsilon})$ as $n \to \infty$. Hence, for the sequence $(R(n))_{n\geq 1}$ in condition (C4) we can choose $R(n) = n^{-\varepsilon'}$ with $0 < \varepsilon' \leq \varepsilon$ sufficiently small such that we obtain $L^* < 1$ in (C4).

Finally, note that the jumps of your piecewise constant processes X_n occur at the random times $U_1, \dots, U_n$ so that interval lengths between consecutive jumps may become arbitrarily small. Condition (C5) allows to cover such instances of processes if the probability for close jumps can be controlled. In our example, it is easy to see that the smallest interval between jumps is of length at least n^{-3} with probability of order $O(1/n)$. Hence, condition (C5) is satisfied with the choice $r_n = n^{-3}$ there. Moreover, the sequences $(r_n)_{n\in\mathbb{N}}$ and $(R(n))_{n\in\mathbb{N}}$ are chosen such that condition (55) is fulfilled. Hence, our main result Theorem 22 applies and we first obtain distributional convergence of the centered normalized sequence in (71) which also implies

$$X_n \xrightarrow{d} X$$

in $(\mathcal{D}[0, 1], d_{\mathrm{sk}})$. Here, we may also apply Theorem 12 to infer convergence of moments of $\|X_n\|$ toward the moments of $\|X\|$.

The use of some other search trees to support partial match queries leads to distributional recurrences related to (66), for example, the 2-d-trees. For the application of our framework in this case, see [6].

Acknowledgements. We thank Svante Janson for his invaluable support during various stages of our research on a functional version of the contraction method, in particular, for bringing the paper [2] to our attention, for his help clarifying the conditions in Lemma 23 and many detailed comments on a first version of this paper. We also thank Alfredas Račkauskas for commenting on the counterexamples discussed in [4] and three anonymous referees for constructive comments.

REFERENCES

[1] ALDOUS, D. (1994). Recursive self-similarity for random trees, random triangulations and Brownian excursion. *Ann. Probab.* **22** 527–545. MR1288122

[2] BARBOUR, A. D. (1990). Stein's method for diffusion approximations. *Probab. Theory Related Fields* **84** 297–322. MR1035659

[3] BARBOUR, A. D. and JANSON, S. (2009). A functional combinatorial central limit theorem. *Electron. J. Probab.* **14** 2352–2370. MR2556014

[4] BENTKUS, V. YU. and RACHKAUSKAS, A. (1984). Estimates for the distance between sums of independent random elements in Banach spaces. *Teor. Veroyatn. Primen.* **29** 49–64. MR0739500

[5] BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd ed. Wiley, New York. MR1700749

[6] BROUTIN, N., NEININGER, R. and SULZBACH, H. (2013). A limit process for partial match queries in random quadtrees and 2-d trees. *Ann. Appl. Probab.* **23** 2560–2603. MR3127945

[7] CARTAN, H. (1971). *Differential Calculus*. Hermann, Paris. MR0344032

[8] CHERN, H.-H. and HWANG, H.-K. (2003). Partial match queries in random quadtrees. *SIAM J. Comput.* **32** 904–915 (electronic). MR2001889

[9] CURIEN, N. and JOSEPH, A. (2011). Partial match queries in two-dimensional quadtrees: A probabilistic approach. *Adv. in Appl. Probab.* **43** 178–194. MR2761153

[10] DIEUDONNÉ, J. (1960). *Foundations of Modern Analysis*. *Pure and Applied Mathematics* **10**. Academic Press, New York. MR0120319

[11] DONSKER, M. D. (1951). An invariance principle for certain probability limit theorems. *Mem. Amer. Math. Soc.* **6** 12. MR0040613

[12] DRMOTA, M., JANSON, S. and NEININGER, R. (2008). A functional limit theorem for the profile of search trees. *Ann. Appl. Probab.* **18** 288–333. MR2380900

[13] EICKMEYER, K. and RÜSCHENDORF, L. (2007). A limit theorem for recursively defined processes in L^p. *Statist. Decisions* **25** 217–235. MR2412071

[14] FLAJOLET, P., GONNET, G., PUECH, C. and ROBSON, J. M. (1993). Analytic variations on quadtrees. *Algorithmica* **10** 473–500. MR1244619

[15] GINÉ, E. and LEÓN, J. R. (1980). On the central limit theorem in Hilbert space. *Stochastica* **4** 43–71. MR0573725

[16] JANSON, S. and KAIJSER, S. (2014). Higher moments of Banach space valued random variables. *Mem. Amer. Math. Soc.* To appear.

[17] JANSON, S. and NEININGER, R. (2008). The size of random fragmentation trees. *Probab. Theory Related Fields* **142** 399–442. MR2438697

[18] LEDOUX, M. and TALAGRAND, M. (1991). *Probability in Banach Spaces*: *Isoperimetry and Processes*. *Ergebnisse der Mathematik und Ihrer Grenzgebiete* (3) [*Results in Mathematics and Related Areas* (3)] **23**. Springer, Berlin. MR1102015

[19] LUKACS, E. (1975). *Stochastic Convergence*, 2nd ed. *Probability and Mathematical Statistics* **30**. Academic Press, New York. MR0375405

[20] NEININGER, R. (2001). On a multivariate contraction method for random recursive structures with applications to Quicksort. *Random Structures Algorithms* **19** 498–524. Analysis of algorithms (Krynica Morska, 2000). MR1871564

[21] NEININGER, R. (2004). *Stochastische Analyse von Algorithmen*, *Fixpunktgleichungen und Ideale Metriken*. Univ. Frankfurt, Habilitation. http://www.math.uni-frankfurt.de/~neiningr/habil.pdf.

[22] NEININGER, R. and RÜSCHENDORF, L. (2004). A general limit theorem for recursive algorithms and combinatorial structures. *Ann. Appl. Probab.* **14** 378–418. MR2023025

[23] NEININGER, R. and RÜSCHENDORF, L. (2004). On the contraction method with degenerate limit equation. *Ann. Probab.* **32** 2838–2856. MR2078559

[24] PESTMAN, W. R. (1995). Measurability of linear operators in the Skorokhod topology. *Bull. Belg. Math. Soc. Simon Stevin* **2** 381–388. MR1355827

[25] RACHEV, S. T. and RÜSCHENDORF, L. (1995). Probability metrics and recursive algorithms. *Adv. in Appl. Probab.* **27** 770–799. MR1341885

[26] RÖSLER, U. (1991). A limit theorem for "Quicksort". *RAIRO Inform. Théor. Appl.* **25** 85–100. MR1104413

[27] RÖSLER, U. (1992). A fixed point theorem for distributions. *Stochastic Process. Appl.* **42** 195–214. MR1176497

[28] RÖSLER, U. (2001). On the analysis of stochastic divide and conquer algorithms. *Algorithmica* **29** 238–261. Average-case analysis of algorithms (Princeton, NJ, 1998). MR1887306

[29] RÖSLER, U. and RÜSCHENDORF, L. (2001). The contraction method for recursive algorithms. *Algorithmica* **29** 3–33. MR1887296

[30] STRASSEN, V. and DUDLEY, R. M. (1969). The central limit theorem and ε-entropy. In *Probability and Information Theory* (*Proc. Internat. Sympos., McMaster Univ., Hamilton, Ont.*, 1968) 224–231. Springer, Berlin. MR0279872

[31] SULZBACH, H. (2012). On a functional contraction method. Dissertation, Univ. Frankfurt. Available at http://publikationen.ub.uni-frankfurt.de/frontdoor/index/index/docId/24858.

[32] ZOLOTAREV, V. M. (1976). Approximation of the distributions of sums of independent random variables with values in infinite-dimensional spaces. *Teor. Veroyatn. Primen.* **21** 741–758. MR0517338

[33] ZOLOTAREV, V. M. (1976). Metric distances in spaces of random variables and of their distributions. *Mat. Sb.* **101(143)** 416–454, 456. MR0467869

[34] ZOLOTAREV, V. M. (1977). Ideal metrics in the problem of approximating the distributions of sums of independent random variables. *Teor. Veroyatn. Primen.* **22** 449–465. MR0455066

[35] ZOLOTAREV, V. M. (1979). Ideal metrics in the problems of probability theory and mathematical statistics. *Austral. J. Statist.* **21** 193–208. MR0561947

INSTITUTE FOR MATHEMATICS
GOETHE UNIVERSITY FRANKFURT
60054 FRANKFURT AM MAIN
GERMANY
E-MAIL: neiningr@math.uni-frankfurt.de
henning.sulzbach@gmail.com