Probab. Theory Relat. Fields (2008) 142:399–442
DOI 10.1007/s00440-007-0110-1

# The size of random fragmentation trees

**Svante Janson · Ralph Neininger**

Received: 13 September 2006 / Revised: 16 October 2007 / Published online: 22 November 2007


**Abstract** We consider the random fragmentation process introduced by Kolmogorov, where a particle having some mass is broken into pieces and the mass is distributed among the pieces at random in such a way that the proportions of the mass shared among different daughters are specified by some given probability distribution (the dislocation law); this is repeated recursively for all pieces. More precisely, we consider a version where the fragmentation stops when the mass of a fragment is below some given threshold, and we study the associated random tree. Dean and Majumdar found a phase transition for this process: the number of fragmentations is asymptotically normal for some dislocation laws but not for others, depending on the position of roots of a certain characteristic equation. This parallels the behavior of discrete analogues with various random trees that have been studied in computer science. We give rigorous proofs of this phase transition, and add further details. The proof uses the contraction method. We extend some previous results for recursive sequences of random variables to families of random variables with a continuous parameter; we believe that this extension has independent interest.

**Mathematics Subject Classification (2000)** Primary: 60F05; Secondary: 60J80 · 60C05 · 68P05

---

Research supported by an Emmy Noether fellowship of the DFG.

---

S. Janson
Department of Mathematics, Uppsala University, PO Box 480, 751 06 Uppsala, Sweden
e-mail: svante.janson@math.uu.se
URL: http://www.math.uu.se/∼svante/

R. Neininger (✉)
Department of Mathematics and Computer Science, J.W. Goethe University,
60054 Frankfurt a.M., Germany
e-mail: neiningr@math.uni-frankfurt.de
URL: http://www.math.uni-frankfurt.de/∼neiningr/

## 1 The problem and result

Consider the following fragmentation process, introduced by Kolmogorov [28], see also Bertoin [3, Chapter 1] and the references in [3, Section 1.6]. Fix $b \geq 2$ and the law for a random vector $\mathbf{V} = (V_1, \dots, V_b)$; this is commonly called the *dislocation law*. We assume throughout the paper that $0 \leq V_j \leq 1$, $j = 1, \dots, b$, and

$$\sum_{j=1}^{b} V_j = 1, \tag{1.1}$$

i.e., that $(V_1, \dots, V_b)$ belongs to the standard simplex. For simplicity we also assume that each $V_j < 1$ a.s. We allow $V_j = 0$, but note that, a.s., $0 < V_j < 1$ for at least one $j$. (The case (1.1) is called *conservative*. The non-conservative case, not treated here, is known to be quite different.)

Starting with an object of mass $x \geq 1$, we break it into $b$ pieces with masses $V_1 x, \dots, V_b x$. Continue recursively with each piece of mass $\geq 1$, using new (independent) copies of the random vector $(V_1, \dots, V_b)$ each time. The process terminates a.s. after a finite number of steps, leaving a finite set of fragments of masses $<1$.

As said above, this model has been studied by many authors, with or without our stopping rule and often without assuming (1.1). The model can be embedded in continuous time (this is immaterial for our purpose), see Bertoin [3, Chapter 1]; in particular, [3, Section 1.4.4] uses the same stopping rule as we do (in a more general situation than ours). Different stopping rules are treated by Gnedin and Yakubovich [19] and Krapivsky et al. [29,30].

We let $N(x)$ be the random number of fragmentation events, i.e., the number of pieces of mass $\geq 1$ that appear during the process; further, let $N_e(x)$ be the final number of fragments, i.e., the number of pieces of mass $<1$ that appear. Dean and Majumdar [12] found (without giving a rigorous proof) that the asymptotic behavior of $N(x)$ as $x \to \infty$ depends on the position of the roots of a certain characteristic equation; the main purpose of this paper is to give a precise version of this in Theorem 1.3 below. Some special cases have earlier been studied by other authors, see Sect. 7.

It is natural to consider the fragmentation process as a tree, with the root representing the original object, its children representing the pieces of the first fragmentation, and so on. It is then convenient to let the fragmentation go on for ever, although we ignore what happens to pieces smaller than 1. Let us mark each node with the mass of the corresponding object.

We thus consider the infinite rooted $b$-ary tree $T_b$, whose nodes are labeled with the strings $J = j_1 \cdots j_k$ with $j_i \in \{1, \dots, b\}$ and $k \geq 0$. Let $\mathcal{B}^*$ denote the set of all such strings, and let $(V_1^{(J)}, \dots, V_b^{(J)})$, $J \in \mathcal{B}^*$, be independent copies of $\mathbf{V}$. Then node $J = j_1 \cdots j_k$ gets the mass $x \prod_{i=1}^{k} V_{j_i}^{(j_1 \cdots j_{i-1})}$. Thus $N(x)$ is the number of nodes with mass $\geq 1$, i.e.,

$$N(x) = \sum_{J \in \mathcal{B}^*} \mathbf{1}_{\left\{x \prod_{i=1}^{k} V_{j_i}^{(j_1 \cdots j_{i-1})} \geq 1\right\}}. \tag{1.2}$$

By the recursive construction of the fragmentation process, we have $N(x) = 0$ for $0 \le x < 1$ and

$$N(x) \overset{\mathrm{d}}{=} 1 + \sum_{j=1}^{b} N^{(j)}(V_j x), \quad x \ge 1, \tag{1.3}$$

where $N^{(j)}(\cdot)$ are copies of the process $N(\cdot)$, independent of each other and of $(V_1, \dots, V_b)$.

We define the *fragmentation tree* $\mathcal{T}(x)$ to be the subtree of $\mathcal{T}(\infty) = T_b$ consisting of all nodes with mass $\ge 1$. Thus $N(x) = |\mathcal{T}(x)|$, the number of nodes in $\mathcal{T}(x)$. More precisely, using standard terminology for trees, we call these nodes *internal nodes* of $\mathcal{T}(x)$, and we say that a node in $\mathcal{T}(\infty)$ is an *external node* of $\mathcal{T}(x)$ if it has label $<1$ but its parent is an internal node of $\mathcal{T}(x)$.

Thus $N(x)$ is the number of internal nodes, and $N_e(x)$ is the number of external nodes. Since each internal node has $b$ internal or external children, we have, for $x \ge 1$, $N(x) + N_e(x) = 1 + bN(x)$, or $N_e(x) = (b-1)N(x) + 1$. Hence the results for $N(x)$ immediately yield similar results for $N_e(x)$ and the total number of external and internal nodes $N(x) + N_e(x)$ too.

In this paper we thus study the size of the fragmentation tree $\mathcal{T}(x)$. Of course, it is interesting to study other properties too, such as height, path length, profile, …

*Remark 1.1* It is obviously equivalent to instead start with mass 1, so node $J = j_1 \cdots j_k$ gets the mass $\prod_{i=1}^{k} V_{j_i}^{(j_1 \cdots j_{i-1})}$, and then keep all nodes with mass $\ge \varepsilon = 1/x$, now considering asymptotics as $\varepsilon \to 0$. This formulation (used for example by Bertoin [3, Section 1.4.4]) is sometimes more convenient, for example, it allows us to define $\mathcal{T}(x)$ for all $x \ge 0$ simultaneously, using the same $V_j^{(J)}$; this defines $(\mathcal{T}(x))_{x \ge 0}$ as an increasing stochastic process of trees. Nevertheless, for our purposes we prefer the formulation above, mainly because of the connection with the discrete models discussed in Remark 9.3.

*Remark 1.2* We assume for convenience that each object is split into the same number $b$ of parts. Our method applies also to some case of a random number of parts. Indeed, if the number of parts is bounded, we can use the results below with $b$ large enough, setting the non-existing $V_j := 0$. It seems possible to extend the proofs below with minor modifications to the cases when the number of parts $b = \infty$ or $b$ is random and unbounded (under suitable assumptions), but we have not pursued this and we leave this extension to the reader.

Note that $N(x)$ makes sense also for $b = \infty$, while $N_e(x) = \infty$ in this case.

Our main result is Theorem 1.3 below on the asymptotic distribution of $N(x)$, together with the corresponding estimates for mean and variance given in Theorem 3.1.

We define (with $0^z := 0$), at least for $\operatorname{Re} z \ge 0$,

$$\phi(z) := \sum_{j=1}^{b} \mathbb{E}\, V_j^z, \tag{1.4}$$

and note that $\phi(z)$ is bounded and analytic in the open right half-plane $\{z : \operatorname{Re} z > 0\}$. More precisely, there exists $a \in [-\infty, 0)$ such that (1.4) converges for real $z > a$ but not for $z < a$; then $\phi(z)$ is analytic in $\{z : \operatorname{Re} z > a\}$. In some cases, $\phi$ may be extended to a meromorphic function in a larger domain. (For several examples of this, see Sect. 7.)

Since we assume (1.1), clearly $\phi(1) = 1$. Since further $0 \leq V_j < 1$ a.s., the function $\phi(z)$ is decreasing for real $z > 0$; hence $\phi(z) > 1$ when $0 < z < 1$ and $\phi(z) < 1$ for $1 < z < \infty$. Further, $|\phi(z)| \leq \sum_j \mathbb{E}\,|V_j^z| = \phi(\operatorname{Re} z)$, so $|\phi(z)| < 1$ when $\operatorname{Re} z > 1$.

A crucial role is played by the solutions to the *characteristic equation*

$$\phi(\lambda) = 1. \tag{1.5}$$

By the comments above, $\lambda = 1$ is one root, and $\operatorname{Re}\lambda \leq 1$ for every root $\lambda$; furthermore, there is no real root in $(0, 1)$. Let, for any $\delta \in [-\infty, \infty)$ such that $\phi$ is analytic, or at least meromorphic, in $\{z : \operatorname{Re} z > \delta\}$, $M(\delta)$ be the number of roots $\lambda$ of (1.5) with $\operatorname{Re}\lambda > \delta$.

We further define

$$\alpha := -\phi'(1) = \sum_{j=1}^{b} \mathbb{E}(-V_j \ln V_j), \tag{1.6}$$

the expected entropy of $(V_1, \dots, V_b)$.

We need a (weak) regularity condition on the distribution of $(V_1, \dots, V_b)$. We find the following convenient, although it can be weakened to Condition B($\delta$) in Sect. 2 for suitable $\delta$. For examples where this regularity and Theorem 1.3 fail, see Example 8.1.

**Condition A** Each $V_j$ has a distribution that is absolutely continuous on $(0, 1)$, although a point mass at 0 is allowed.

Note that there is no condition on the joint distribution. In one case, however, we need also a condition including the joint distribution. (Note that both conditions are satisfied if $\mathbf{V}$ has a density on the standard simplex, i.e., if $(V_1, \dots, V_{b-1})$ has a density.)

**Condition A′** The support of the distribution of $\mathbf{V}$ on the standard simplex has an interior point.

If Condition A holds, then, by Lemmas 2.2 and 2.1 below, the number $M(\delta)$ of roots of $\phi(\lambda) = 1$ in $\{\lambda : \operatorname{Re}\lambda > \delta\}$ is finite for every $\delta > 0$. We may thus order the roots with $\operatorname{Re}\lambda > 0$ as $\lambda_1, \lambda_2, \dots, \lambda_{M(0)}$ with decreasing real parts: $\lambda_1 = 1 > \operatorname{Re}\lambda_2 \geq \operatorname{Re}\lambda_3 \geq \cdots$; we will assume this in the sequel. If $\lambda_1 = 1$ is the only root with $\operatorname{Re}\lambda > 0$, we set $\lambda_2 = -\infty$ for convenience.

We let $\mathcal{M}^{\mathbb{C}}$ denote the space of probability measures on $\mathbb{C}$, and let

$$\mathcal{M}_2^{\mathbb{C}}(\gamma) := \left\{\eta \in \mathcal{M}^{\mathbb{C}} : \int |z|^2 \, d\eta(z) < \infty, \quad \text{and} \quad \int z \, d\eta(z) = \gamma\right\}, \quad \gamma \in \mathbb{C}.$$

We let $T$ denote the map (assuming $\lambda_2 \neq -\infty$)

$$T : \mathcal{M}^{\mathbb{C}} \to \mathcal{M}^{\mathbb{C}}, \quad \eta \mapsto \mathcal{L}\left(\sum_{r=1}^{b} V_r^{\lambda_2} Z^{(r)}\right), \tag{1.7}$$

where $(V_1, \ldots, V_b), Z^{(1)}, \ldots, Z^{(b)}$ are independent and $\mathcal{L}(Z^{(r)}) = \eta$ for $r = 1, \ldots, b$. Note that $T$ maps $\mathcal{M}_2^{\mathbb{C}}(\gamma)$ into itself for each $\gamma$, since $\lambda_2$ satisfies $\phi(\lambda_2) = 1$.

We state our main result. The constant $\alpha > 0$ is defined in (1.6) above and $\beta$ is given explicitly in Theorem 3.1. The $\ell_2$ distance between distributions is defined in Sect. 4.

**Theorem 1.3** *Suppose that Condition* A *holds. Then we have*:

(i) *If* $\operatorname{Re}\lambda_2 < 1/2$ *then* $\mathbb{E}\,N(x) = \alpha^{-1}x + o(\sqrt{x})$, $\operatorname{Var} N(x) \sim \beta x$ *with* $\beta > 0$ *and*

$$\frac{N(x) - \alpha^{-1}x}{\sqrt{x}} \xrightarrow{\mathrm{d}} \mathcal{N}(0, \beta).$$

(ii) *If* $\operatorname{Re}\lambda_2 = 1/2$ *and each root* $\lambda_i$ *with* $\operatorname{Re}\lambda_i = 1/2$ *is a simple root of* $\phi(\lambda) = 1$, *and further Condition* A′ *too holds, then* $\mathbb{E}\,N(x) = \alpha^{-1}x + O(\sqrt{x})$, $\operatorname{Var}(N(x)) \sim \beta x \ln x$ *with* $\beta > 0$ *and*

$$\frac{N(x) - \alpha^{-1}x}{\sqrt{x \ln x}} \xrightarrow{\mathrm{d}} \mathcal{N}(0, \beta).$$

(iii) *If* $\operatorname{Re}\lambda_2 > 1/2$, *and* $\lambda_2$ *and* $\lambda_3 = \overline{\lambda_2}$ *are the only roots of* (1.5) *with this real part, and these roots are simple, then* $\mathbb{E}\,N(x) = \alpha^{-1}x + \operatorname{Re}(\gamma x^{\lambda_2}) + O(x^{\kappa})$, *for some* $\gamma$ *and* $\kappa$ *with* $\gamma \in \mathbb{C} \setminus \{0\}$ *and* $1/2 < \kappa < \operatorname{Re}\lambda_2$, *and*

$$\ell_2\left(\frac{N(x) - \alpha^{-1}x}{x^{\operatorname{Re}\lambda_2}}, \operatorname{Re}\left(\Xi e^{\mathrm{i}\operatorname{Im}\lambda_2 \ln x}\right)\right) = O\left(x^{\kappa - \operatorname{Re}\lambda_2}\right),$$

*for some complex random variable* $\Xi$. *Furthermore,* $\mathcal{L}(\Xi)$ *is the unique fixed point of* $T$ *in* $\mathcal{M}_2^{\mathbb{C}}(\gamma)$.

*Remark 1.4* In case (iii), the normalized $N(x)$ thus does not converge in distribution; instead we have an asymptotic periodicity in $\log x$ of the distribution. This type of asymptotic periodicity is common for properties of some types of random trees, see for example Chern and Hwang [10], Chauvin and Pouyanne [8], Fill and Kapur [16], Janson [22, Example 7.8, Remark 3.20] and Janson [24, Examples 4.4, 4.5].

The trichotomy in the theorem is very similar to the situation for multi-type branching processes and generalized Pólya urns, see [22], in particular Theorems 3.22–3.24 there; in that case, the $\lambda_i$ are the eigenvalues of a certain matrix.

*Remark 1.5* We can regard our process as a general (age-dependent) branching process [21, Chapter 6], provided we make a logarithmic change of time as in Sect. 3. (This approach has been used in related problems by for example Gnedin and Yakubovich [19].) Indeed, there are two versions. For internal nodes, the individuals in the branching process live for ever, and give birth at times $-\ln V_1, \dots, -\ln V_b$. For external nodes, we have a splitting process where each individual when it dies gives birth to new particles with life lengths $-\ln V_1, \dots, -\ln V_b$. For both versions, we obtain a supercritical branching process with Malthusian parameter 1, but the identity (1.1) causes the asymptotics for moments and distributions to be quite different from typical supercritical branching processes; the reason is that the intrinsic martingale [3, Section 1.2.2] degenerates to a constant, unlike in the non-conservative case (such as, e.g., in [2,4,19]).

*Remark 1.6* Distributions that are fixed points of (1.7) can sometimes be found explicitly. For example, if $\lambda_2$ in (1.7) is real, then the stable distributions of index $1/\lambda_2$ are examples of fixed points of $T$. Note, however, that in our case, $\lambda_2$ is never real. Moreover, the fixed points we are interested in have finite variance, and are thus quite different from stable distributions. Other examples with explicit solutions are given in, e.g., Gnedin and Yakubovich [19] (in this case, generalized Mittag–Leffler distributions).

For the related Quicksort fixed point equation, Fill and Janson [15] found a complete characterization of the set of fixed points; in that case, all fixed points are formed by combining certain stable distributions with the unique fixed point with mean 0 and finite variance.

*Remark 1.7* Condition A$'$ is needed only in part (ii), and is needed only to exclude the possibility that for each root $\lambda_i$ of (1.5) with $\operatorname{Re}\lambda_i = 1/2$,

$$\sum_{j=1}^{b} V_j^{\lambda_i} = 1 \quad \text{a.s.} \tag{1.8}$$

This is easily seen to be impossible if Condition A$'$ holds, and even otherwise it seems highly unlikely for any particular example, but it seems possible to construct examples satisfying Condition A where $\mathbf{V}$ is concentrated on a curve, say, such that (1.8) holds.

We will prove the statements on mean and variance, with further refinements, in Sect. 3. To prove convergence in distribution, we will use a continuous time version of the contraction method. We develop a general theorem, that we find to be of independent interest, in Sect. 5. This theorem is applied to our problem in Sect. 6. Some examples are given in Sects. 7 and 8.

*Remark 1.8* As an alternative to using the random vector $\mathbf{V}$ to describe the fragmentation process, one can use the point process $\sum_j \delta_{V_j}$ on $[0, 1]$. Let $\eta$ be the intensity of this process; thus $\eta$ is a measure on $[0, 1]$. In this formulation, $\phi$ is the Mellin transform of the measure $\eta$; further $\mu$ and $\nu$ in Sect. 3 equal the measures $\eta$ and its size biased version $s\eta(ds)$ after the change of variable $s = e^{-x}$.

## 2 Further preliminaries

We define (again with $0^z := 0$)

$$\psi(z, w) := \operatorname{Cov}\left(\sum_{j=1}^{b} V_j^z, \sum_{j=1}^{b} V_j^w\right) = \mathbb{E}\left(\sum_{j=1}^{b} V_j^z \sum_{k=1}^{b} V_k^w\right) - \phi(z)\phi(w). \quad (2.1)$$

In particular, $\psi(z, \bar{z}) = \mathbb{E}\left|\sum_{j=1}^{b} V_j^z - \phi(z)\right|^2 \geq 0$, with equality only if $\sum_{j=1}^{b} V_j^z = \phi(z)$ a.s.

For $\operatorname{Re} z, \operatorname{Re} w \geq 0$, we have $|V_j^z|, |V_j^w| \leq 1$ and thus $|\psi(z, w)| \leq 2b^2$.

We say that $\mathbf{V}$ is *lattice* if there exists a number $R > 1$ such that every $V_j \in \{R^{-n}\}_{n\geq 0} \cup \{0\}$ a.s.; otherwise $\mathbf{V}$ is *non-lattice*. Basic Fourier analysis applied to the probability measure $\nu$ defined in (3.16) shows that $\mathbf{V}$ is non-lattice if and only if $\lambda = 1$ is the only root of (1.5) with $\operatorname{Re} \lambda = 1$. (Otherwise, there is an infinite number of roots with $\operatorname{Re} \lambda = 1$.) We will assume this, and more, below.

We introduce a family of regularity conditions that are weaker than Condition A.

**Condition B($\delta$)** (Here $\delta$ is a real number with $\delta \geq 0$.)

$$\limsup_{t\to\infty} |\phi(\delta + \mathrm{i}t)| < 1.$$

**Lemma 2.1** *If Condition* B($\delta$) *holds for some* $\delta \geq 0$, *then Condition* B($\delta'$) *holds for every* $\delta' > \delta$ *as well*; *moreover*

$$\limsup_{\substack{\operatorname{Re} z\geq\delta \\ \operatorname{Im} z\to\infty}} |\phi(z)| < 1.$$

*Proof* Choose first $\varepsilon > 0$ such that $\limsup_{t\to\infty} |\phi(\delta + \mathrm{i}t)| < 1 - 2\varepsilon$, and then $A$ such that $|\phi(\delta + \mathrm{i}t)| \leq 1 - 2\varepsilon$ if $t \geq A$, and thus also if $t \leq -A$. Recall further that $|\phi(\delta + \mathrm{i}t)| \leq b$ for all $t$. Since $\phi(z)$ is analytic, and thus harmonic, in the half-plane $\mathcal{H}_\delta := \{z : \operatorname{Re} z > \delta\}$ and bounded and continuous in $\overline{\mathcal{H}_\delta}$, $\phi$ is given by the Poisson integral of its boundary values [18, Lemma 3.4]:

$$\phi(x + \mathrm{i}y) = \int_{-\infty}^{\infty} P_{x-\delta}(y - t)\phi(\delta + \mathrm{i}t)\, dt, \quad x > \delta, \quad (2.2)$$

where $P_x(y) = x/(\pi(x^2 + y^2))$, the Poisson kernel for the right half-plane. Let $\omega(x + \mathrm{i}y) := \int_{-A}^{A} P_{x-\delta}(y - t)\, dt$, the harmonic measure of $[\delta - \mathrm{i}A, \delta + \mathrm{i}A]$; then (2.2) implies

$$|\phi(x + \mathrm{i}y)| \leq \int_{-\infty}^{\infty} P_{x-\delta}(y - t)|\phi(\delta + \mathrm{i}t)|\, dt \leq b\omega(x + \mathrm{i}y) + 1 - 2\varepsilon. \quad (2.3)$$

It is well-known, and easy to see, that the set $B := \{z \in \mathcal{H}_\delta : \omega(z) > \varepsilon/b\}$ is bounded; in fact, it is the intersection of $\mathcal{H}_\delta$ and a circular disc [18, p. 13]. Thus, $A_1 := \sup\{\operatorname{Im} z : z \in B\} < \infty$, and if $\operatorname{Re} x \geq \delta$ and $|y| > A_1$, then $\omega(z) \leq \varepsilon/b$ and (2.3) yields $|\phi(x + \mathrm{i}y)| \leq 1 - \varepsilon$. □

**Lemma 2.2** *If Condition* A *holds, then Condition* B($\delta$) *holds for every* $\delta \geq 0$.

This is given in [2, Lemma 2] and included here for completeness.

*Proof* We have $\mathbb{E}\, V_j^{\mathrm{i}t} = \mathbb{E}\left(e^{\mathrm{i}t \ln V_j} \mathbf{1}_{\{V_j>0\}}\right)$, the Fourier transform of the distribution of $\ln V_j$ (ignoring any point mass at 0), so by Condition A and the Riemann–Lebesgue lemma, $\mathbb{E}\, V_j^{it} \to 0$ as $t \to \infty$ for every $j$, and thus $\phi(\mathrm{i}t) \to 0$ as $t \to \infty$. Hence, Condition B(0) holds, and the result follows by Lemma 2.1. □

**Lemma 2.3** *If Condition* B($\delta$) *holds for some* $\delta > 0$*, then there is only a finite number of roots to* $\phi(\lambda) = 1$ *with* $\operatorname{Re}\lambda \geq \delta$.

*Proof* By Lemma 2.1, all such roots satisfy $|\operatorname{Im}\lambda| \leq C$ for some $C < \infty$. Furthermore, all roots satisfy $\operatorname{Re}\lambda \leq 1$, so if further $\operatorname{Re}\lambda \geq \delta$, $\lambda$ belongs to a compact rectangle $K$ in the open right half-plane. Since, $\phi(z) - 1$ is analytic and non-constant in this half-plane, it has only a finite number of roots in $K$. □

In particular, by the comments above, Condition B($\delta$) with $\delta \leq 1$ implies that $\mathbf{V}$ is non-lattice.

## 3 Mean and variance

We let $\Lambda$ denote the set of solutions to the characteristic equation (1.5), i.e.,

$$\Lambda := \{\lambda : \phi(\lambda) = 1\}; \tag{3.1}$$

we further define its subsets

$$\Lambda(s) := \{z \in \Lambda : \operatorname{Re}(z) = s\}. \tag{3.2}$$

In general, $\phi(\lambda)$ is defined only for $\operatorname{Re}\lambda \geq 0$, and we consider only such $\lambda$ in (3.1). However, in cases where $\phi$ extends to a meromorphic function in a larger domain (for example, when $\phi$ is rational), we may include such $\lambda$ too in $\Lambda$; this makes no difference in Theorem 3.1. (In Theorem 3.4, we include all roots in the complex plane.) We will use $\Lambda(s)$ only for $s \geq 0$, where there is no ambiguity.

Let $m(x) := \mathbb{E}\, N(x)$ and $\sigma^2(x) := \operatorname{Var} N(x)$. We will show the following asymptotics.

**Theorem 3.1** *Assume that Condition* B($\delta$) *holds with* $0 \leq \delta < 1$*, and let* $\lambda_1, \ldots, \lambda_{M(\delta)}$ *be the elements of* $\{\lambda \in \Lambda : \operatorname{Re}\lambda > \delta\}$*, ordered so that* $\lambda_1 = 1 > \operatorname{Re}\lambda_2 \geq \operatorname{Re}\lambda_3 \geq \cdots$*. (If* $M(\delta) = 1$*, let* $\lambda_2 = -\infty$*.) Then, as* $x \to \infty$*:*

(i) $m(x) \sim \alpha^{-1} x$.

(ii) *If further* $\phi'(\lambda_i) \neq 0$ *for* $i = 1, \dots, M(\delta)$, *i.e., each* $\lambda_i$ *is a simple root of* $\phi(\lambda) = 1$, *then, more precisely, for every* $\delta' > \delta$,

$$m(x) = \alpha^{-1}x + \sum_{i=2}^{M(\delta)} \frac{1}{-\lambda_i \phi'(\lambda_i)} x^{\lambda_i} + O\left(x^{\delta'}\right). \tag{3.3}$$

(iii) *If* $\delta < 1/2$ *and* $\operatorname{Re}\lambda_2 < 1/2$ (*including the case* $M(\delta) = 1$), *then* $\sigma^2(x) \sim \beta x$, *with*

$$\beta = \alpha^{-1} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\psi(1/2 + \mathrm{i}u, 1/2 - \mathrm{i}u)}{|1/2 + \mathrm{i}u|^2 |1 - \phi(1/2 + \mathrm{i}u)|^2} \, du \in (0, \infty). \tag{3.4}$$

(iv) *If* $\delta < 1/2 = \operatorname{Re}\lambda_2$, *and each* $\lambda_i$ *with* $\operatorname{Re}\lambda_i = 1/2$ *is a simple root of* $\phi(\lambda) = 1$, *then* $\sigma^2(x) = \beta x \ln x + o(x \ln x)$, *with*

$$\beta = \sum_{\lambda \in \Lambda(1/2)} \frac{1}{\alpha |\lambda \phi'(\lambda)|^2} \psi(\lambda, \overline{\lambda}) \geq 0. \tag{3.5}$$

*If, moreover, Condition* A$'$ *holds* (*or, more generally, for some* $\lambda_i \in \Lambda(1/2)$, (1.8) *does not hold*), *then* $\beta > 0$.

(v) *If* $\operatorname{Re}\lambda_2 > 1/2$, *and each* $\lambda_i$ *with* $\operatorname{Re}\lambda_i = \operatorname{Re}\lambda_2$ *is a simple root of* $\phi(\lambda) = 1$, *then*

$$\begin{aligned}\sigma^2(x) = &\sum_{\lambda_i, \lambda_k \in \Lambda(\operatorname{Re}\lambda_2)} \frac{1}{\lambda_i \lambda_k \phi'(\lambda_i) \phi'(\lambda_k) \left(1 - \phi(\lambda_i + \lambda_k)\right)} \psi(\lambda_i, \lambda_k) x^{\lambda_i + \lambda_k} \\ &+ o\left(x^{2 \operatorname{Re}\lambda_2}\right).\end{aligned}$$

*Remark 3.2* It follows from the proof that for (i) we do not need Condition B($\delta$); it is enough that **V** is non-lattice.

*Remark 3.3* The case when some $\phi'(\lambda_i) = 0$ is similar; now terms $x^{\lambda_i} \ln x$ (and possibly $x^{\lambda_i} \ln^d x$, $d \geq 2$) will appear in (3.3). We leave the details to the reader.

If $\phi$ is a rational function, then (3.3) can be improved to an exact formula. Furthermore, in case (iii) of Theorem 3.1 we then can give an alternative formula for $\beta$.

**Theorem 3.4** *Assume that* $\phi$ *is a rational function, and let* $\lambda_1, \dots, \lambda_M$ *be the roots of* $\phi(\lambda) = 1$ *in the complex plane, with* $\lambda_1 = 1$. *Suppose further that all these roots are simple.*

(i) *Then*

$$m(x) = \sum_{i=1}^{M} \frac{1}{-\lambda_i \phi'(\lambda_i)} x^{\lambda_i} - \frac{1}{\phi(0) - 1}, \quad x \geq 1. \tag{3.6}$$

(ii) *Assume further that* $\operatorname{Re}\lambda_i < 1/2$ *for* $i = 2, \dots, M$, *and that* $V_j > 0$ *a.s. for every* $j$. *Define, for notational convenience,* $\lambda_0 := 0$, $a_0 := -1/(b-1)$ *and* $a_i := -1/(\lambda_i \phi'(\lambda_i))$ *for* $i = 1, \dots, M$. *Then* $\sigma^2(x) \sim \beta x$, *with*

$$\begin{aligned}
\beta = {} & \alpha^{-1} \sum_{i,k \neq 1} \frac{a_i a_k}{1-\lambda_i-\lambda_k} \left( \sum_{j,l=1}^{b} \mathbb{E}\, V_j^{\lambda_i} V_l^{\lambda_k} (V_j \wedge V_l)^{1-\lambda_i-\lambda_k} - 2\phi(1-\lambda_k) + 1 \right) \\
& - 2\alpha^{-2} \sum_{i=2}^{M} \frac{a_i}{\lambda_i} \left( \sum_{j,l=1}^{b} \mathbb{E}\left( V_j^{\lambda_i} V_l^{1-\lambda_i} - V_l \right) \mathbf{1}_{\{V_l \le V_j\}} - \phi(1-\lambda_i) + 1 \right) \\
& - 2\alpha^{-2} a_0 \left( \sum_{j,l=1}^{b} \mathbb{E}\, V_l (\ln V_j - \ln V_l) \mathbf{1}_{\{V_l < V_j\}} - \alpha \right) \\
& + \alpha^{-3} \left( \sum_{j,l=1}^{b} \mathbb{E}(V_j \wedge V_l) - 1 \right) - \alpha^{-1}.
\end{aligned}$$

The proof of these theorems will occupy the remainder of this section. We first show that all moments of $N(x)$ are finite.

**Lemma 3.5** *For every* $m \ge 1$ *and* $x \ge 0$, $\mathbb{E}\, N(x)^m < \infty$. *Furthermore,* $\sup_{0 \le y \le x} \mathbb{E}\, N(y)^m < \infty$.

*Proof* For a string $J = j_1 \cdots j_k \in \mathcal{B}^*$ we denote by $|J| = k$ the depth of the corresponding node in $T_b$. Note that we have $|\{J \in \mathcal{B}^* : 0 \le |J| \le k\}| \le b^{k+1}$. Hence, if $N(x) > b^{k+1}$ for some $k \ge 0$ then by (1.2) there exists a $J = j_1 \cdots j_k \in \mathcal{B}^*$ with $x \prod_{i=1}^{k} V_{j_i}^{(j_1 \cdots j_{i-1})} \ge 1$. Markov's inequality implies that for all $q \ge 1$

$$\begin{aligned}
\mathbb{P}(N(x) > b^{k+1}) &\le \mathbb{P}\left( \bigcup_{J \in \mathcal{B}^* : |J| = k} \left\{ \prod_{i=1}^{k} V_{j_i}^{(j_1 \cdots j_{i-1})} \ge 1/x \right\} \right) \\
&\le \sum_{J \in \mathcal{B}^* : |J| = k} \mathbb{P}\left( \prod_{i=1}^{k} V_{j_i}^{(j_1 \cdots j_{i-1})} \ge 1/x \right) \\
&\le \sum_{J \in \mathcal{B}^* : |J| = k} x^q\, \mathbb{E} \prod_{i=1}^{k} \left( V_{j_i}^{(j_1 \cdots j_{i-1})} \right)^q \\
&= x^q \phi(q)^k.
\end{aligned}$$

Hence, for all $y \ge b$ we obtain with $k = \lfloor \log_b y \rfloor - 1$ and $\phi(q) \le 1$ that

$$\mathbb{P}(N(x) > y) \le x^q \phi(q)^k \le x^q \phi(q)^{\log_b y - 2} = \frac{x^q}{\phi(q)^2}\, y^{\log_b \phi(q)}. \tag{3.7}$$

We have $\phi(q) \to 0$ as $q \to \infty$ since $V_j < 1$ a.s. and by dominated convergence. Hence, for all $m \ge 1$ there exists a $q > 0$ with $\log_b \phi(q) < -m$. The tail bound (3.7) thus implies $\mathbb{E}\, N(x)^m < \infty$ for all $m \ge 1$ and all $x \ge 0$.

The final statement follows because $0 \le N(y) \le N(x)$ when $0 \le y \le x$. □

We find it convenient to switch from multiplicative to additive notion. We therefore define

$$X_j := -\ln V_j \in (0, \infty], \quad j = 1, \ldots, b,$$
$$N_*(t) := N(e^t), \quad -\infty \le t < \infty.$$

The definition (1.2) and the recursive equation (1.3) thus translate to

$$N_*(t) = \sum_{J \in \mathcal{B}^*} \mathbf{1}_{\left\{\sum_{i=1}^{k} X_{j_i}^{(j_1 \cdots j_{i-1})} \le t\right\}}, \tag{3.8}$$

$$N_*(t) \overset{\mathrm{d}}{=} 1 + \sum_{j=1}^{b} N_*^{(j)}(t - X_j), \quad t \ge 0, \tag{3.9}$$

where $N_*^{(j)}(\cdot)$ are independent copies of the process $N_*(\cdot)$, and $N_*(t) = 0$ for $-\infty \le t < 0$. Further define

$$m_*(t) := \mathbb{E}\, N_*(t) = m(e^t),$$
$$\sigma_*^2(t) := \operatorname{Var} N_*(t) = \sigma^2(e^t).$$

Thus $m_*(t) = \sigma_*^2(t) = 0$ for $t < 0$. Taking expectations in (3.9) we find

$$m_*(t) = 1 + \mathbb{E} \sum_{j=1}^{b} m_*(t - X_j), \quad t \ge 0. \tag{3.10}$$

Let $\mu_j$ be the distribution of $X_j$ on $(0, \infty)$; this is a measure of mass $1 - \mathbb{P}(V_j = 0)$; let further $\mu := \sum_{j=1}^{b} \mu_j$. Then (3.10) can be written as

$$m_*(t) = 1 + \sum_{j=1}^{b} \mu_j * m_*(t) = 1 + \mu * m_*(t), \quad t \ge 0, \tag{3.11}$$

where $\mu * f(t) = \int_0^\infty f(t - x)\, d\mu(x)$. This is the standard renewal equation, except that $\mu$ is not a probability measure.

Similarly, conditioning on $X_1, \dots, X_b$, for $t \geq 0$,

$$
\begin{aligned}
&\mathbb{E}\Big((N_*(t) - m_*(t))^2 \mid X_1, \dots, X_b\Big) \\
&\quad = \mathbb{E}\left(\left[\sum_{j=1}^{b} N_*^{(j)}(t - X_j) + 1 - m_*(t)\right]^2 \,\middle|\, X_1, \dots, X_b\right) \\
&\quad = \operatorname{Var}\left(\sum_{j=1}^{b} N_*^{(j)}(t - X_j) + 1 - m_*(t) \,\middle|\, X_1, \dots, X_b\right) \\
&\qquad + \left(\sum_{j=1}^{b} m_*(t - X_j) + 1 - m_*(t)\right)^2 \\
&\quad = \sum_{j=1}^{b} \sigma_*^2(t - X_j) + \left(\sum_{j=1}^{b} m_*(t - X_j) - m_*(t) + 1\right)^2.
\end{aligned}
$$

Taking the expectation we obtain

$$
\sigma_*^2(t) = \mathbb{E} \sum_{j=1}^{b} \sigma_*^2(t - X_j) + h(t) = \mu * \sigma_*^2(t) + h(t), \quad t \geq 0, \tag{3.12}
$$

where, recalling (3.10),

$$
\begin{aligned}
h(t) &:= \mathbb{E}\left(\sum_{j=1}^{b} m_*(t - X_j) - m_*(t) + 1\right)^2 \\
&= \mathbb{E}\left(\sum_{j=1}^{b} m_*(t - X_j) - m_*(t)\right)^2 + 2\left(\mathbb{E}\sum_{j=1}^{b} m_*(t - X_j) - m_*(t)\right) + 1 \\
&= \mathbb{E}\left(\sum_{j=1}^{b} m_*(t - X_j) - m_*(t)\right)^2 - 1. 
\end{aligned} \tag{3.13}
$$

Both (3.11) and (3.12) are instances of the general renewal equation (3.14) below, and from renewal theory we get the following result. We say that a function on $[0, \infty)$ is *locally bounded* if it is bounded on every finite interval.

**Lemma 3.6** *Assume that* $\mathbf{V}$ *is non-lattice. Let* $f$ *be a locally bounded measurable function on* $[0, \infty)$*. Then the renewal equation*

$$
F = f + \mu * F \tag{3.14}
$$

*has a unique locally bounded solution $F$ on* $[0,\infty)$. *We have the following asymptotical results, as $t\to\infty$,*

(i) *If $f$ is a.e. continuous and $\int_0^\infty f^*(t)\,dt<\infty$, where $f^*(t):=\sup_{u\ge t} e^{-u}|f(u)|$, then $F(t)=(\gamma+o(1))e^t$, with $\gamma=\alpha^{-1}\int_0^\infty f(t)e^{-t}\,dt$.*
(ii) *If $f(t)=e^t$, then $F(t)\sim\alpha^{-1}te^t$.*
(iii) *If $f(t)=e^{\lambda t}$ with* $\operatorname{Re}\lambda=1$ *and* $\operatorname{Im}\lambda\neq 0$, *then $F(t)=o\left(te^t\right)$.*
(iv) *If $f(t)=e^{\lambda t}$ with* $\operatorname{Re}\lambda>1$, *then $F(t)\sim(1-\phi(\lambda))^{-1}e^{\lambda t}$.*

*Proof* For a function $f$ on $(0,\infty)$ and $z\in\mathbb{C}$, we define, when the integral exists, the Laplace transform $\tilde f(z):=\int_0^\infty e^{-zt}f(t)\,dt$. Similarly, the Laplace transform of $\mu$ is

$$\widetilde{\mu}(z):=\int_0^\infty e^{-tz}\,d\mu(t)=\sum_{j=1}^b\mathbb{E}\,e^{-zX_j}=\sum_{j=1}^b\mathbb{E}\,V_j^z=\phi(z),\tag{3.15}$$

at least for $\operatorname{Re} z\ge 0$. (Using the original variable $\ln t$, the Laplace transforms become Mellin transforms, cf. Remark 1.8.)

Since $\mu$ is not a probability measure, we define another ("conjugate" or "tilted") measure $\nu$ on $[0,\infty)$ by

$$d\nu(u)=e^{-u}\,d\mu(u).\tag{3.16}$$

Then $\nu$ is a probability measure because, by (1.1),

$$\nu[0,\infty)=\int_0^\infty e^{-u}\,d\mu(u)=\sum_{j=1}^b\int_0^\infty e^{-u}\,d\mu_j(u)=\sum_{j=1}^b\mathbb{E}\,e^{-X_j}=\sum_{j=1}^b\mathbb{E}\,V_j=1.$$

Further, the mean of the distribution $\nu$ is

$$\mathbb{E}\,\nu=\int_0^\infty u\,d\nu(u)=\int_0^\infty ue^{-u}\,d\mu(u)=\sum_{j=1}^b\mathbb{E}\left(X_je^{-X_j}\right)=\sum_{j=1}^b\mathbb{E}\left((-\ln V_j)V_j\right)=\alpha\tag{3.17}$$

and the Laplace transform is, for $\operatorname{Re} z\ge 0$, recalling (3.15),

$$\widetilde{\nu}(z):=\int_0^\infty e^{-zu}\,d\nu(u)=\int_0^\infty e^{-u-zu}\,d\mu(u)=\widetilde{\mu}(z+1)=\phi(z+1).\tag{3.18}$$

Let $g(t) := e^{-t} f(t)$ and $G(t) := e^{-t} F(t)$. Then (3.14) translates to

$$G(t) = e^{-t} F(t) = e^{-t} f(t) + \int_0^\infty e^{-t} F(t-u)\, d\mu(u)$$
$$= g(t) + \int_0^\infty G(t-u) e^{-u}\, d\mu(u) = g(t) + \nu * G(t).$$

In other words, $G$ satisfies the renewal equation for the probability measure $\nu$, so we can use standard results from renewal theory.

First, it is well known that the equation $G = g + \nu * G$ has a unique locally bounded solution which is given by $G = \sum_{n=0}^\infty \nu^{*n} * g$, and thus $F = \sum_{n=0}^\infty \mu^{*n} * f$; see, e.g., [1, Theorem IV.2.4] (which also applies directly to $F$). If we let $Y_1, Y_2, \dots$ be i.i.d. random variables with the distribution $\nu$, and let $S_n := \sum_1^n Y_i$, this can be written as

$$G(t) = \sum_{n=0}^\infty \mathbb{E}\left(g(t - S_n)\mathbf{1}_{\{S_n \le t\}}\right) = \mathbb{E} \sum_{S_n \le t} g(t - S_n). \tag{3.19}$$

Under the assumptions of (i), $f^*$ is non-increasing and integrable; further, $\sup f^* \le \sup_{[0,1]} |f| + f^*(1) < \infty$, so $f^*$ is bounded too. Hence [1, Proposition IV.4.1(v),(iv)] shows that $f^*$ and $g$ are directly Riemann integrable. The key renewal theorem [1, Theorem IV.4.3] and (3.17) now yield $G(t) \to \alpha^{-1} \int_0^\infty g(u)\, du = \gamma$, which proves (i).

In case (ii) we have $g(t) = 1$, and thus $G(t) \sim \alpha^{-1} t$ by the elementary renewal theorem [1, IV.(1.5) and Theorem 2.4].

For (iii), $g(t) = e^{(\lambda-1)t} = e^{ibt}$ for some real $b \ne 0$. The solution to (3.14) may be written [1, Theorem IV.2.4] $G(t) = \int_0^t g(t-u)\, dU(u)$, where $U$ is the locally bounded solution to $U = 1 + \nu * U$ (i.e., $U = G$ for case (ii)). Since, in analogy with (3.17), $\int u^2\, d\nu(u) = \sum_j \mathbb{E}\left((\ln V_j)^2 V_j\right) < \infty$, the distribution $\nu$ has finite variance, and the renewal theorem has the sharper version [14, Theorem XI.3.1]

$$U(t) = \alpha^{-1} t + c + R(t),$$

where $c$ is a certain constant ($\int u^2\, d\nu(u)/2\alpha^2$) and $R(t) \to 0$ as $t \to \infty$. Hence, using integration by parts for one term,

$$G(t) = \int_0^t e^{ib(t-u)} \alpha^{-1}\, du + c e^{ibt} + \int_0^t e^{ib(t-u)}\, dR(u)$$
$$= O(1) + R(t) - R(0) e^{ibt} + ib \int_0^t e^{ib(t-u)} R(u)\, du = o(t).$$

For (iv), we have by (3.19) with $g(t) = e^{(\lambda-1)t}$, using dominated convergence and (3.18),

$$
\begin{aligned}
e^{-\lambda t} F(t) &= e^{(1-\lambda)t} G(t) = \sum_{n=0}^{\infty} \mathbb{E}\left(e^{(1-\lambda)t+(\lambda-1)(t-S_n)} \mathbf{1}_{\{t \le S_n\}}\right) \\
&\to \sum_{n=0}^{\infty} \mathbb{E}\left(e^{-(\lambda-1)S_n}\right) = \sum_{n=0}^{\infty} \mathbb{E}\left(e^{-(\lambda-1)Y_1}\right)^n = \sum_{n=0}^{\infty} \widetilde{\nu}(\lambda-1)^n \\
&= \sum_{n=0}^{\infty} \phi(\lambda)^n = (1-\phi(\lambda))^{-1}.
\end{aligned}
$$

□

*Proof of Theorem 3.1* We first apply Lemma 3.6 (i) to (3.11), with $f(t) = 1$ for $t \ge 0$, and obtain $\gamma = \alpha^{-1}$ and $m_*(t) \sim \alpha^{-1} e^t$, which proves Theorem 3.1 (i).

To obtain more refined asymptotics, we use Laplace transforms. Let $H(t) := \mathbf{1}_{\{t \ge 0\}}$ (the Heaviside function), and note that $\widetilde{H}(z) = \int_0^\infty e^{-tz}\,dt = 1/z$, $\operatorname{Re} z > 0$. Since the Laplace transform converts convolutions to products, the renewal equation (3.11) yields $\widetilde{m_*}(z) = \widetilde{H}(z) + \widetilde{\mu}(z)\widetilde{m_*}(z)$, and thus

$$
\widetilde{m_*}(z) = \frac{\widetilde{H}(z)}{1-\widetilde{\mu}(z)} = \frac{1}{z(1-\phi(z))}, \tag{3.20}
$$

for $z$ such that the transforms exist. By the estimate $m_*(t) \sim \alpha^{-1} e^t$ above, $m_*(t) = O(e^t)$ and thus $\widetilde{m_*}(z)$ exists for $\operatorname{Re} z > 1$. Consequently, (3.20) holds for $\operatorname{Re} z > 1$, and can be used to extend $\widetilde{m_*}(z)$ to a meromorphic function for $\operatorname{Re} z > 0$.

We want to invert the Laplace transform in (3.20). This is simple if $\phi$ is rational, yielding (3.6). (Note that $\phi(0) = \mathbb{E}\,|\{j : V_j > 0\}| > 1$.) In general, there are difficulties to doing this directly, because $\widetilde{m_*}(z)$ is not integrable along a vertical line $\operatorname{Re} z = s$; it decreases too slowly as $|\operatorname{Im} z| \to \infty$. We therefore regularize. Let $\varepsilon > 0$, and let $H_\varepsilon := H * \varepsilon^{-1}\mathbf{1}_{[0,\varepsilon]}$; thus

$$
H_\varepsilon(t) = \begin{cases} 0, & t < 0, \\ 1 - t/\varepsilon, & 0 \le t < \varepsilon, \\ 1, & t \ge \varepsilon. \end{cases}
$$

Let $m_{*\varepsilon} = \sum_{n=0}^\infty \mu^{*n} * H_\varepsilon$ be the locally bounded solution to $m_{*\varepsilon} = H_\varepsilon + \mu * m_{*\varepsilon}$. Note that $H_\varepsilon(t) \le H(t) \le H_\varepsilon(t+\varepsilon)$, and thus

$$
m_{*\varepsilon}(t) \le m_*(t) \le m_{*\varepsilon}(t+\varepsilon). \tag{3.21}
$$

We have

$$\widetilde{H_\varepsilon}(z) = \widetilde{H}(z)\varepsilon^{-1}\int_0^\varepsilon e^{-zt}\,dt = \frac{1-e^{-\varepsilon z}}{\varepsilon z^2}, \quad \operatorname{Re} z > 0,$$

and we find, arguing as for (3.20) above,

$$\widetilde{m_{*\varepsilon}}(z) = \frac{\widetilde{H_\varepsilon}(z)}{1-\widetilde{\mu}(z)} = \frac{1-e^{-\varepsilon z}}{\varepsilon z^2(1-\phi(z))},$$

first for $\operatorname{Re} z > 1$, and then for $\operatorname{Re} z > 0$, extending $\widetilde{m_{*\varepsilon}}$ to a meromorphic function in this domain. This function decreases (using Condition B($\delta$) and Lemma 2.1) as $|\operatorname{Im} z|^{-2}$ on vertical lines $\operatorname{Re} z = s \geq \delta$, and is thus integrable there. Hence, the Laplace inversion formula (a Fourier inversion) shows that for any $s > 1$ and $t \geq 0$,

$$m_{*\varepsilon}(t) = \frac{1}{2\pi\mathrm{i}}\int_{s-\mathrm{i}\infty}^{s+\mathrm{i}\infty} e^{tz}\,\widetilde{m_{*\varepsilon}}(z)\,dz. \tag{3.22}$$

We may, increasing $\delta$ a little if necessary, assume that $\phi(z) = 1$ has no roots with $\operatorname{Re} z = \delta$; in cases (iii), (iv) and (v) we may similarly assume that each $\lambda \in \Lambda$ with $\operatorname{Re}\lambda > \delta$ has $\phi'(\lambda) \neq 0$. It is then easy to show, using Condition B($\delta$) and Lemma 2.1, that we may shift the line of integration in (3.22) to $\operatorname{Re} z = \delta$ and obtain, for $0 < \varepsilon \leq 1$,

$$\begin{aligned}
m_{*\varepsilon}(t) &= \frac{1}{2\pi\mathrm{i}}\int_{\delta-\mathrm{i}\infty}^{\delta+\mathrm{i}\infty} e^{tz}\widetilde{m_{*\varepsilon}}(z)\,dz + \sum_{i=1}^{M(\delta)} \operatorname{Res}_{z=\lambda_i}\left(e^{tz}\widetilde{m_{*\varepsilon}}(z)\right)\\
&= O\left(e^{t\delta}\int_{\delta-\mathrm{i}\infty}^{\delta+\mathrm{i}\infty} \min\left(\frac{1}{|z|}, \frac{\varepsilon}{|z|^2}\right)|dz|\right) + \sum_{i=1}^{M(\delta)} \frac{e^{t\lambda_i}}{-\lambda_i\phi'(\lambda_i)}\,\frac{1-e^{-\varepsilon\lambda_i}}{\varepsilon\lambda_i}\\
&= \sum_{i=1}^{M(\delta)} \frac{e^{t\lambda_i}}{-\lambda_i\phi'(\lambda_i)}(1+O(\varepsilon)) + O\left(e^{t\delta}\left(1+\ln\frac{1}{\varepsilon}\right)\right).
\end{aligned}$$

Now choosing $\varepsilon := e^{-t}$ we obtain

$$m_{*\varepsilon}(t) = \sum_{i=1}^{M(\delta)} \frac{e^{t\lambda_i}}{-\lambda_i\phi'(\lambda_i)} + O(1) + O\left(e^{t\delta}(1+t)\right), \quad t \geq 0.$$

Replacing $t$ by $t+\varepsilon$, we obtain the same estimate for $m_{*\varepsilon}(t+\varepsilon)$, and thus (3.21) yields

$$m_*(t) = \sum_{i=1}^{M(\delta)} \frac{e^{t\lambda_i}}{-\lambda_i \phi'(\lambda_i)} + O\left(e^{t\delta'}\right), \quad t \geq 0, \tag{3.23}$$

which yields Theorem 3.1 (ii). (Recall that $-\lambda_1\phi'(\lambda_1) = -\phi'(1) = \alpha$.)

For the estimates of the variance, we use Lemma 3.6 and (3.12). It is easily seen (by dominated convergence) that $h$ in (3.13) is a.e. continuous. Choose $\delta' > \delta$ with $\delta < \delta' < \operatorname{Re}\lambda_{M(\delta)}$; in case (iii) with $M(\delta) = 1$, let further $\delta' < 1/2$.

Note that then (3.23) trivially holds for $t < 0$ too. Hence,

$$\begin{aligned}
\sum_{j=1}^{b} m_*(t-X_j) - m_*(t) &= \sum_{i=1}^{M(\delta)} \frac{\sum_{j=1}^{b} e^{(t-X_j)\lambda_i} - e^{t\lambda_i}}{-\lambda_i\phi'(\lambda_i)} + O\left(e^{t\delta'}\right) \\
&= \sum_{i=2}^{M(\delta)} \frac{e^{t\lambda_i}}{-\lambda_i\phi'(\lambda_i)} \left(\sum_{j=1}^{b} V_j^{\lambda_i} - 1\right) + O\left(e^{t\delta'}\right),
\end{aligned}$$

where we use the fact that $\lambda_1 = 1$ and thus $\sum_{j=1}^{b} V_j^{\lambda_1} - 1 = \sum_{j=1}^{b} V_j - 1 = 0$. Consequently, by (3.13) and (2.1), letting $\sigma_2 := \operatorname{Re}\lambda_2 > \delta'$ if $M(\delta) \geq 2$, and $\sigma_2 := \delta'$ if $M(\delta) = 1$,

$$h(t) = \sum_{i=2}^{M(\delta)} \sum_{k=2}^{M(\delta)} \frac{e^{(\lambda_i+\lambda_k)t}}{\lambda_i\lambda_k\phi'(\lambda_i)\phi'(\lambda_k)} \psi(\lambda_i,\lambda_k) + O\left(e^{t(\delta'+\sigma_2)}\right). \tag{3.24}$$

For Theorem 3.1 (iii), (3.24) yields $h(t) = O\left(e^{2\sigma_2 t}\right)$ with $\sigma_2 < 1/2$, and Lemma 3.6 (i) applies to (3.12), yielding $\sigma_*^2(t) \sim \gamma e^t$. We postpone the calculation of $\beta = \gamma$, verifying (3.4), to Lemma 3.7.

For Theorem 3.1 (iv) and (v), we treat the terms in (3.24) separately, using linearity; for the error term we also use monotonicity and comparison with the case $f(t) = e^{t(\delta'+\sigma_2)}$. In order to solve (3.12), we thus consider (3.14), with $f(t)$ replaced by the individual terms in (3.24), and apply Lemma 3.6, letting $t := \ln x$. For (iv), i.e., $\operatorname{Re}\lambda_2 = 1/2$, a term in (3.24) with $\operatorname{Re}\lambda_k = \operatorname{Re}\lambda_i = 1/2$ and $\lambda_k = \overline{\lambda_i}$, and thus $\lambda_i + \lambda_k = 1$, yields by Lemma 3.6 (ii) a contribution $(\alpha\lambda_i\lambda_k\phi'(\lambda_i)\phi'(\lambda_k))^{-1}\psi(\lambda_i,\lambda_k)te^t = \alpha^{-1}|\lambda_i\phi'(\lambda_i)|^{-2}\psi(\lambda_i,\overline{\lambda_i})te^t$. The contributions of all other terms in (3.24) are $o(te^t)$, by Lemma 3.6 (iii) (the other cases with $\operatorname{Re}\lambda_k = \operatorname{Re}\lambda_i = 1/2$) and Lemma 3.6 (iii) (the remaining cases).

Similarly, for (v), the leading terms come from the cases $\operatorname{Re}\lambda_k = \operatorname{Re}\lambda_i = \operatorname{Re}\lambda_2$ and Lemma 3.6 (iv).

Furthermore, by (3.5), $\beta = 0$ in (iv) only if for every $\lambda_i \in \Lambda(1/2)$, we have $\psi(\lambda_i,\bar\lambda_i) = \mathbb{E}\,|\sum_j V_j^{\lambda_i} - \phi(\lambda_i)|^2 = 0$, and thus, since $\phi(\lambda_i) = 1$, (1.8) holds. □

**Lemma 3.7** *Under the assumptions of* Theorem 3.1(iii), *with* $h(t)$ *as in* (3.13),

$$\int_0^\infty h(t)e^{-t}\,dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\psi(1/2+\mathrm{i}u,1/2-\mathrm{i}u)}{|1/2+\mathrm{i}u|^2|1-\phi(1/2+\mathrm{i}u)|^2}\,du > 0. \tag{3.25}$$

*Proof* Write $f(t) := \left(m_*(t)-\alpha^{-1}e^t\right)e^{-t/2}$, $-\infty < t < \infty$. Thus, by (3.23), $f(t) = O\left(e^{-(1/2-\sigma_2)t}\right)$ for $t \ge 0$ and $f(t) = -\alpha^{-1}e^{t/2} = O\left(e^{-|t|/2}\right)$ for $t < 0$. In particular, $f \in L^2(-\infty,\infty)$. Furthermore, the (two-sided) Laplace transform $\tilde{f}(z) := \int_{-\infty}^{\infty} f(t)e^{-tz}\,dt$ is analytic for $-(1/2-\sigma_2) < \operatorname{Re} z < 1/2$.

Define further $f_1(t) := f(t)e^{t/2} = m_*(t)-\alpha^{-1}e^t$ and $f_2(t) := f_1(t)\mathbf{1}_{\{t\ge 0\}}$. Then $f_2(t) = O\left(e^{\sigma_2 t}\right)$, and thus the Laplace transform $\tilde{f_2}(z)$ is analytic for $\operatorname{Re} z > \sigma_2$. For $\operatorname{Re} z > 1$ we have, by (3.20),

$$\begin{aligned}\tilde{f_2}(z) &= \int_0^\infty e^{-tz}\left(m_*(t)-\alpha^{-1}e^t\right)dt = \widetilde{m_*}(z) - \alpha^{-1}(z-1)^{-1}\\ &= \frac{1}{z(1-\phi(z)} - \frac{1}{\alpha(z-1)};\end{aligned}$$

by analytic continuation, this formula holds for $\operatorname{Re} z > \sigma_2$. Consequently, for $\sigma_2 < \operatorname{Re} z < 1$,

$$\tilde{f_1}(z) = \tilde{f_2}(z) + \int_{-\infty}^0 e^{-tz}(-\alpha^{-1}e^t)\,dt = \tilde{f_2}(z) - \alpha^{-1}(1-z)^{-1} = \frac{1}{z(1-\phi(z)}.$$

Since $\tilde{f}(z) = \tilde{f_1}(z+1/2)$, we find the Fourier transform

$$\hat{f}(u) := \int_0^\infty e^{-\mathrm{i}ut} f(t)\,dt = \tilde{f}(\mathrm{i}u) = \tilde{f_1}(\tfrac12+\mathrm{i}u) = \frac{1}{\left(\frac12+\mathrm{i}u\right)\left(1-\phi\left(\frac12+\mathrm{i}u\right)\right)}. \tag{3.26}$$

Next, since $\sum_j e^{-X_j} = \sum_j V_j = 1$,

$$\begin{aligned}\sum_{j=1}^b m_*(t-X_j) - m_*(t) &= \sum_{j=1}^b f_1(t-X_j) - f_1(t) + \alpha^{-1}\sum_{j=1}^b e^{t-X_j} - \alpha^{-1}e^t\\ &= \sum_{j=1}^b f_1(t-X_j) - f_1(t),\end{aligned}$$

so by (3.13), and defining $\Psi(w,y) := \int_{-\infty}^{\infty} f(t-w)f(t-y)\,dt$,

$$\int_0^\infty h(t)e^{-t}\,dt+1=\int_0^\infty (h(t)+1)\,e^{-t}\,dt=\mathbb{E}\int_{-\infty}^{\infty}\left(\sum_{j=1}^{b} f_1(t-X_j)-f_1(t)\right)^2 e^{-t}\,dt$$
$$=\mathbb{E}\int_{-\infty}^{\infty}\left|\sum_{j=1}^{b} e^{-X_j/2}f(t-X_j)-f(t)\right|^2 dt$$
$$=\mathbb{E}\sum_{j,k=1}^{b} e^{-X_j/2-X_k/2}\Psi(X_j,X_k)$$
$$-2\,\mathbb{E}\sum_{j=1}^{b} e^{-X_j/2}\Psi(X_j,0)+\Psi(0,0).$$

By Parseval's relation and $f=\overline{f}$,

$$\Psi(w,y)=\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\mathrm{i}uw}\hat{f}(u)\overline{e^{-\mathrm{i}uy}\hat{f}(u)}\,du=\frac{1}{2\pi}\int_{-\infty}^{\infty}|\hat{f}(u)|^2 e^{\mathrm{i}u(y-w)}\,du.$$

Hence,

$$\int_0^\infty h(t)e^{-t}\,dt+1=\mathbb{E}\,\frac{1}{2\pi}\int_{-\infty}^{\infty}|\hat{f}(u)|^2\left(\sum_{j,k=1}^{b} e^{-X_j/2-X_k/2+\mathrm{i}u(X_k-X_j)}\right.$$
$$\left.-\sum_{j=1}^{b} e^{-X_j/2+\mathrm{i}uX_j}-\sum_{k=1}^{b} e^{-X_k/2-\mathrm{i}uX_k}+1\right)du$$
$$=\mathbb{E}\,\frac{1}{2\pi}\int_{-\infty}^{\infty}|\hat{f}(u)|^2\,\mathbb{E}\left|\sum_{j=1}^{b} V_j^{1/2-\mathrm{i}u}-1\right|^2 du$$
$$=\mathbb{E}\,\frac{1}{2\pi}\int_{-\infty}^{\infty}|\hat{f}(u)|^2\Big(\psi(1/2+\mathrm{i}u,1/2-\mathrm{i}u)$$
$$+|\phi(1/2+\mathrm{i}u)-1|^2\Big)\,du.$$

Using (3.26),

$$\mathbb{E}\,\frac{1}{2\pi}\int_{-\infty}^{\infty}|\hat{f}(u)|^2|\phi(1/2+\mathrm{i}u)-1|^2\,du=\int_{-\infty}^{\infty}\frac{du}{|1/2+\mathrm{i}u|^2}=\int_{-\infty}^{\infty}\frac{du}{\frac{1}{4}+u^2}=2\pi,$$

and (3.25) follows by (3.26).

Since $\psi(z,\bar{z}) \geq 0$, and by dominated convergence is continuous for $\operatorname{Re} z \geq 0$, it follows from (3.4) that $\beta = 0$ only if for every $z$ with $\operatorname{Re} z = 1/2$, $\psi(z,\bar{z}) = \mathbb{E}\,|\sum_j V_j^z - \phi(z)|^2 = 0$, and thus

$$\sum_{j=1}^{b} V_j^{1/2+\mathrm{i}u} = \phi(1/2+\mathrm{i}u) \tag{3.27}$$

a.s., for every real $u$. Considering first rational $u$, we see that a.s. (3.27) holds for all real $u$.

However, for any realization $(V_1,\dots,V_b)$ and $\varepsilon > 0$, the Kronecker–Weyl theorem shows that $(1,\dots,1)$ is a cluster point of $(\exp(\mathrm{i}u\log V_1),\dots,\exp(\mathrm{i}u\log V_b))$ as $u \to \infty$ (even with $u \in \mathbb{N}$); thus it is possible to find arbitrarily large $u$ with $\operatorname{Re} V_j^{\mathrm{i}u} \geq (1-\varepsilon)$ and thus $\operatorname{Re} V_j^{1/2+\mathrm{i}u} \geq (1-\varepsilon)V_j^{1/2}$ for $j = 1,\dots,b$. Hence, (3.27) implies that $\limsup_{u\to\infty} |\phi(\frac{1}{2}+\mathrm{i}u)| \geq \phi(1/2) \geq 1$. This contradicts Condition B($\delta$) and Lemma 2.1. Hence $\beta > 0$. $\square$

*Proof of Theorem 3.4* (i): As remarked above, (3.6) follows by inverting the Laplace transform in (3.20), using a partial fraction expansion. (ii): Note first that $\phi(z) \to 0$ as $z \to +\infty$ (by dominated convergence); hence, $\phi$ being rational and thus continuous at $\infty$, $\phi(\infty) = 0$ and $\phi(z) \to 0$ as $|z| \to \infty$. Consequently, Condition B($\delta$) holds for every $\delta$. We thus see that the conditions of Theorem 3.1 (iii) are satisfied, and from the proof above we see that, with $h$ given by (3.13),

$$\beta = \alpha^{-1}\int_0^\infty h(t)e^{-t}\,dt = \alpha^{-1}\int_1^\infty h(\ln x)x^{-2}\,dx.$$

We have, by (3.6), $m(x) = \sum_{i=0}^{M} a_i x^{\lambda_i}\mathbf{1}_{\{x\geq 1\}}$ and thus

$$m_*(\ln x - X_j) = m(xe^{-X_j}) = m(xV_j) = \sum_{i=0}^{M} a_i (xV_j)^{\lambda_i}\mathbf{1}_{\{x\geq V_j^{-1}\}}.$$

Hence, letting $V_0 := 1$, $\varepsilon_0 = -1$ and $\varepsilon_j = 1$ for $j \geq 1$, and recalling (1.1),

$$\begin{aligned} H(x) &:= \sum_{j=1}^{b} m_*(\ln x - X_j) - m_*(\ln x) = \sum_{i=0}^{M} a_i x^{\lambda_i}\left(\sum_{j=1}^{b} V_j^{\lambda_i}\mathbf{1}_{\{x\geq V_j^{-1}\}} - \mathbf{1}_{\{x\geq 1\}}\right) \\ &= \sum_{i\neq 1} a_i x^{\lambda_i}\sum_{j=0}^{b} V_j^{\lambda_i}\varepsilon_j\mathbf{1}_{\{x\geq V_j^{-1}\}} - a_1 x\sum_{j=1}^{b} V_j\mathbf{1}_{\{1\leq x< V_j^{-1}\}}. \end{aligned}$$

By (3.13), this leads to

$$\int_1^\infty h(\ln x)x^{-2}\,dx + 1 = \int_1^\infty \mathbb{E}\,H(x)^2x^{-2}\,dx = \mathbb{E}\int_1^\infty H(x)^2x^{-2}\,dx$$

$$= \mathbb{E}\sum_{i,k\neq 1} a_i a_k \sum_{j,l=0}^{b} \varepsilon_j\varepsilon_l V_j^{\lambda_i} V_l^{\lambda_k} \int_{V_j^{-1}\vee V_l^{-1}}^{\infty} x^{\lambda_i+\lambda_k-2}\,dx$$

$$-2\,\mathbb{E}\sum_{i\neq 1} a_1 a_i \sum_{j=0}^{b}\sum_{l=1}^{b} \varepsilon_j V_j^{\lambda_i} V_l \int_{V_j^{-1}}^{V_l^{-1}} x^{\lambda_i-1}\,dx\,\mathbf{1}_{\{V_j^{-1}\le V_l^{-1}\}}$$

$$+\,\mathbb{E}\,a_1^2 \sum_{j,l=1}^{b} V_j V_l \int_1^{V_j^{-1}\wedge V_l^{-1}} dx$$

and the result follows by straightforward calculations, noting that $a_1 = \alpha^{-1}$. □

## 4 Zolotarev metric and minimal $L_s$ metric

In this section we collect properties of the minimal $L_s$ metric and the Zolotarev metric that are used subsequently.

We denote by $\mathcal{M}^d$ the space of probability measures on $\mathbb{R}^d$. The *minimal $L_s$ metric* $\ell_s$, $s > 0$, is defined on the subspace $\mathcal{M}_s^d \subset \mathcal{M}^d$ of probability measures with finite absolute moment of order $s$ by

$$\ell_s(\mu,\nu) := \inf\left\{\|X-Y\|_s^{s\wedge 1} \;:\; X \stackrel{d}{=} \mu,\; Y \stackrel{d}{=} \nu\right\}, \quad \mu,\nu\in\mathcal{M}_s^d,$$

where $\|X\|_s := (\mathbb{E}\,|X|^s)^{1/s}$ denotes the $L_s$ norm of $X$. The infimum is taken over all random vectors of $X, Y$ on a joint probability space with the given marginal distributions $\mu$ and $\nu$. (In other words, over all couplings $(X, Y)$ of $\mu$ and $\nu$.) We will also use the notation $\ell_s(X, Y) := \ell_s(\mathcal{L}(X), \mathcal{L}(Y))$.

For $s \ge 1$ and $\gamma \in \mathbb{R}^d$, we denote by $\mathcal{M}_s^d(\gamma) \subset \mathcal{M}_s^d$ the subspace of probability measures with expectation $\gamma$. The pairs $(\mathcal{M}_s^d, \ell_s)$, $s > 0$, and $(\mathcal{M}_s^d(\gamma), \ell_s)$, $s \ge 1$, are complete metric spaces and convergence in $\ell_s$ is equivalent to weak convergence plus convergence of the absolute moments of order $s$.

Random vectors $(X, Y)$ with $X \stackrel{d}{=} \mu$, $Y \stackrel{d}{=} \nu$, and $\ell_s(\mu,\nu) = \|X-Y\|_s^{s\wedge 1}$ are called optimal couplings of $(\mu,\nu)$. Such optimal couplings exist for all $\mu,\nu\in\mathcal{M}_s^d$. These properties can be found in Dall'Aglio [11], Major [32], Bickel and Freedman [5], and Rachev [34]. Similar properties hold for probability measures on $\mathbb{C}^d$ (because $\mathbb{C}^d \cong \mathbb{R}^{2d}$), where we use corresponding notations.

The *Zolotarev metric* $\zeta_s$, $s > 0$ is defined by

$$\zeta_s(X, Y) := \zeta(\mathcal{L}(X), \quad \mathcal{L}(Y)) := \sup_{f \in \mathcal{F}_s} |\mathbb{E}(f(X) - f(Y))|, \tag{4.1}$$

where $s = m + \alpha$ with $0 < \alpha \leq 1$, $m = \lceil s \rceil - 1 \geq 0$ is an integer, and

$$\mathcal{F}_s := \{f \in C^m(\mathbb{R}^d, \mathbb{R}) : \|f^{(m)}(x) - f^{(m)}(y)\| \leq \|x - y\|^\alpha\},$$

where $C^m(\mathbb{R}^d, \mathbb{R})$ denotes the space of $m$ times continuously differentiable functions $f$ on $\mathbb{R}^d$ and $f^{(m)}$ their $m$th derivative.

The expression $\zeta_s(X, Y)$ is finite if $X$ and $Y$ have finite absolute moments of order $s$ and all mixed moments of orders $1, \ldots, m$ of $X$ and $Y$ coincide.

The metric $\zeta_s$ is an ideal metric of order $s$, i.e., we have for $Z$ independent of $(X, Y)$ and any $d \times d$ square matrix $A$

$$\zeta_s(X + Z, \ Y + Z) \leq \zeta_s(X, Y), \quad \zeta_s(AX, AY) \leq \|A\|_{\mathrm{op}}^s \, \zeta_s(X, Y),$$

where $\|A\|_{\mathrm{op}} := \sup_{\|u\|=1} \|Au\|$ denotes the operator norm of the matrix. Convergence in $\zeta_s$ implies weak convergence. For general reference and properties of $\zeta_s$ we refer to Zolotarev [38,39] and Rachev [34].

## 5 General contraction theorems in continuous time

In this section we extend a general contraction theorem for recursive sequences $(Y_n)_{n \geq 0}$ of $d$-dimensional vectors as developed in Neininger and Rüschendorf [33] to families $(Y_t)_{t \geq 0}$ of $d$-dimensional vectors with continuous parameter $t \in [0, \infty)$. (For future applications, and since the proof is the same except for some minor notational differences, we state the result for random vectors. The reader may concentrate on the one-dimensional case, which is the only case needed in the rest of the paper.) We assume that we have

$$Y_t \stackrel{d}{=} \sum_{r=1}^{K} A_r(t) Y^{(r)}_{T_r^{(t)}} + b_t, \quad t \geq \tau_0, \tag{5.1}$$

where $K$ is a positive integer, $\tau_0 \geq 0$, and $T^{(t)} = (T_1^{(t)}, \ldots, T_K^{(t)})$ is a vector of random indices $T_r^{(t)} \in [0, t]$, the $A_r(t)$ are random $d \times d$ matrices for $r = 1, \ldots, K$ and $b_t$ is a random $d$-dimensional vector; further, $(Y_t^{(1)})_{t \geq 0}, \ldots, (Y_t^{(K)})_{t \geq 0}$ and $(A_1(t), \ldots, A_K(t), b_t, T^{(t)})_{t \geq 0}$ are mutually independent families of random variables, and for each $t \geq 0$, $Y_t$ and $Y_t^{(r)}$ are identically distributed for all $r = 1, \ldots, K$.

We assume that all $Y_t$ as well as $A_r(t)$, $b_t$ and $T^{(t)}$ are defined on some probability space $(\Omega, \mathcal{F}, \mu)$, and that they are measurable functions of $(t, \omega)$. (This is a technicality to ensure that the sum in (5.1) is well-defined. Note, however, that the joint distribution of $Y_t$ for different $t$ is irrelevant.)

We introduce the normalized random vectors

$$X_t := C_t^{-1/2}(Y_t - M_t), \quad t \geq 0, \tag{5.2}$$

where $M_t \in \mathbb{R}^d$ and $C_t$ is a symmetric, positive definite square matrix. We assume that $M_t$ and $C_t$ are measurable functions of $t$; further restrictions on $M_t$ and $C_t$ will be given in Convention C. The recurrence (5.1) implies a recurrence for $X_t$,

$$X_t \stackrel{d}{=} \sum_{r=1}^{K} A_r^{(t)} X^{(r)}_{T_r^{(t)}} + b^{(t)}, \quad t \geq \tau_0, \tag{5.3}$$

with independence relations as in (5.1) and

$$A_r^{(t)} = C_t^{-1/2} A_r(t) C^{1/2}_{T_r^{(t)}}, \quad b^{(t)} = C_t^{-1/2}\left(b_t - M_t + \sum_{r=1}^{K}\left(A_r(t) M_{T_r^{(t)}}\right)\right). \tag{5.4}$$

As for the case with integer indexed vectors we establish a transfer theorem of the following form: Appropriate convergence of the coefficients $A_r^{(t)} \to A_r^*$, $b^{(t)} \to b^*$ implies weak convergence of the quantities $X_t$ to a limit $X$. The distribution $\mathcal{L}(X)$ of $X$ is a fixed point of the limiting equation obtained from (5.3) by letting formally $t \to \infty$:

$$X \stackrel{d}{=} \sum_{r=1}^{K} A_r^* X^{(r)} + b^*, \tag{5.5}$$

where $(A_1^*, \ldots, A_K^*, b^*)$, $X^{(1)}, \ldots, X^{(K)}$ are independent and $X^{(r)} \stackrel{d}{=} X$ for $r = 1, \ldots, K$. To formalize this we introduce the map $T$ on the space $\mathcal{M}^d$ of probability measures on $\mathbb{R}^d$ by

$$T : \mathcal{M}^d \to \mathcal{M}^d, \quad \eta \mapsto \mathcal{L}\left(\sum_{r=1}^{K} A_r^* Z^{(r)} + b^*\right), \tag{5.6}$$

where $(A_1^*, \ldots, A_K^*, b^*)$, $Z^{(1)}, \ldots, Z^{(K)}$ are independent and $\mathcal{L}(Z^{(r)}) = \eta$ for $r = 1, \ldots, K$. Then $X$ is a solution of (5.5) if and only if $\mathcal{L}(X)$ is a fixed point of $T$.

We make use of Zolotarev's metric $\zeta_s$ with $0 < s \leq 3$. To ensure finiteness of the metric subsequently we make the following assumptions about the scaling imposed in (5.2):

**Convention C** For $1 < s \leq 3$ we assume that $M_t = \mathbb{E}\, Y_t$. For $2 < s \leq 3$ we assume that $\mathrm{Cov}(Y_t)$ is positive definite for all $t \geq \tau_1$ with a $\tau_1 \geq \tau_0$ and that $C_t = \mathrm{Id}_d$ for $0 \leq t < \tau_1$ and $C_t = \mathrm{Cov}(Y_t)$ for $t \geq \tau_1$.

This convention implies that $X_t$ is centered for $1 < s \leq 3$ and has $\mathrm{Id}_d$ as its covariance matrix for $2 < s \leq 3$ and $t \geq \tau_1$. (For $0 < s \leq 1$, Convention C is void.)

**Theorem 5.1** *Let $0 < s \le 3$ and let $(Y_t)_{t\ge 0}$ be a process of random vectors satisfying* (5.1) *such that $\|Y_t\|_s < \infty$ for every $t$. Denote by $X_t$ the rescaled quantities in* (5.2)*, assuming Convention* C. *Assume that $\|A_r^{(t)}\|_s < \infty$, $\|b^{(t)}\|_s < \infty$ and $\sup_{0\le u\le t}\|X_u\|_s < \infty$ for every $t \ge 0$, and*

$$\left(A_1^{(t)}, \dots, A_K^{(t)}, b^{(t)}\right) \xrightarrow{\ell_s} \left(A_1^*, \dots, A_K^*, b^*\right), \tag{5.7}$$

$$\mathbb{E}\sum_{r=1}^{K} \|A_r^*\|_{\mathrm{op}}^s < 1, \tag{5.8}$$

$$\mathbb{E}\left[\mathbf{1}_{\left\{T_r^{(t)}\le\tau\right\}} \left\|A_r^{(t)}\right\|_{\mathrm{op}}^s\right] \to 0 \tag{5.9}$$

*for every $\tau > 0$ and $r = 1, \dots, K$. Then $X_t$ converges in distribution to a limit $X$, and*

$$\zeta_s(X_t, X) \to 0, \quad t \to \infty, \tag{5.10}$$

*where $\mathcal{L}(X)$ is the unique fixed point of $T$ given in* (5.6) *subject to $\|X\|_s < \infty$ and*

$$\begin{cases} \mathbb{E}\, X = 0 & \text{for } 1 < s \le 2, \\ \mathbb{E}\, X = 0, \ \mathrm{Cov}(X) = \mathrm{Id}_d & \text{for } 2 < s \le 3. \end{cases} \tag{5.11}$$

*Proof* This proof is a continuous extension of the proof of Theorem 4.1 in Neininger and Rüschendorf [33] for the discrete time case. The existence and uniqueness of the fixed point of $T$ subject to (5.11) is obtained as follows: For $1 < s \le 3$ Eq. (5.3) implies $\mathbb{E}\, b^{(t)} = 0$ for all $t > 0$, thus by (5.7) we obtain $\mathbb{E}\, b^* = 0$. For $2 < s \le 3$ Eq. (5.3) implies that for all $t \ge \tau_1$

$$\begin{aligned} \mathrm{Id}_d &= \mathrm{Cov}(X_t) \\ &= \mathbb{E}\left[b^{(t)}\left(b^{(t)}\right)^{\mathrm{tr}}\right] + \mathbb{E}\left[\sum_{r=1}^{K}\left(\mathbf{1}_{\{T_r^{(t)}<\tau_1\}} A_r^{(t)} \tilde{C}_{T_r^{(t)}} \left(A_r^{(t)}\right)^{\mathrm{tr}} + \mathbf{1}_{\{T_r^{(t)}\ge\tau_1\}} A_r^{(t)} \left(A_r^{(t)}\right)^{\mathrm{tr}}\right)\right], \end{aligned}$$

where $b^{\mathrm{tr}}$ denotes the transpose of a vector or matrix and $\tilde{C}_t := \mathrm{Cov}(X_t)$; recall that $\tilde{C}_t = \mathrm{Id}$ when $t \ge \tau_1$.

By (5.7), (5.9) and Hölder's inequality this implies

$$\mathbb{E}\left[b^*(b^*)^{\mathrm{tr}}\right] + \mathbb{E}\left[\sum_{r=1}^{K} A_r^*(A_r^*)^{\mathrm{tr}}\right] = \mathrm{Id}_d.$$

Now, Corollary 3.4 in [33] implies existence and uniqueness of the fixed-point.

Since

$$\mathbb{E}\sum_{r=1}^{K} \|A_r^{(t)}\|_{\mathrm{op}}^s \to \mathbb{E}\sum_{r=1}^{K} \|A_r^*\|_{\mathrm{op}}^s = \xi < 1 \tag{5.12}$$

there exist $\xi_+ \in (\xi, 1)$ and $\tau_2 > \tau_1$ such that for all $t \geq \tau_2$ we have

$$\mathbb{E} \sum_{r=1}^{K} \| A_r^{(t)} \|_{\mathrm{op}}^s \leq \xi_+ < 1. \tag{5.13}$$

Now, we introduce the quantity

$$Q_t := \sum_{r=1}^{K} A_r^{(t)} \left( \mathbf{1}_{\left\{T_r^{(t)} < \tau_2\right\}} X_{T_r^{(t)}}^{(r)} + \mathbf{1}_{\left\{T_r^{(t)} \geq \tau_2\right\}} X^{(r)} \right) + b^{(t)}, \quad t \geq \tau_1, \tag{5.14}$$

where $(A_1^{(t)}, \dots, A_K^{(t)}, b^{(t)}, T^{(t)}), X^{(1)}, \dots, X^{(K)}, (X_t^{(1)}), \dots, (X_t^{(K)})$ are independent with $X^{(r)} \sim X$ and $X_t^{(r)} \sim X_t$ for $r = 1, \dots, K$ and $t \geq 0$. Comparing with (5.3) we obtain that $Q_t$ is centered for $1 < s \leq 3$ and has the covariance matrix $\mathrm{Id}_d$ for $2 < s \leq 3$ and $t \geq \tau_1$. Hence, $\zeta_s$ distances between $X_t$, $Q_t$ and $X$ are finite for all $t \geq \tau_1$. The triangle inequality implies

$$\Delta(t) := \zeta_s(X_t, X) \leq \zeta_s(X_t, Q_t) + \zeta_s(Q_t, X). \tag{5.15}$$

As in the proof for the discrete case we obtain $\zeta_s(Q_t, X) \to 0$ as $t \to 0$, where we use that $\sup_{0 \leq t \leq \tau_2} \|X_t\|_s < \infty$.

The first summand of (5.15) requires a continuous analog of the estimate in the discrete case. Using the properties of the $\zeta_s$ metric, we obtain, for $t \geq \tau_1$,

$$\zeta_s(X_t, Q_t) \leq \mathbb{E} \sum_{r=1}^{K} \mathbf{1}_{\left\{T_r^{(t)} \geq \tau_2\right\}} \left\| A_r^{(t)} \right\|_{\mathrm{op}}^s \Delta\left(T_r^{(t)}\right), \tag{5.16}$$

and, with (5.15), and $r_t := \zeta_s(Q_t, X)$ it follows

$$\Delta(t) \leq \mathbb{E} \sum_{r=1}^{K} \mathbf{1}_{\left\{T_r^{(t)} \geq \tau_2\right\}} \left\| A_r^{(t)} \right\|_{\mathrm{op}}^s \Delta\left(T_r^{(t)}\right) + r_t. \tag{5.17}$$

Now, we obtain $\Delta(t) \to 0$ in two steps, first showing that $(\Delta(t))_{t \geq 0}$ is bounded and then, using the bound, that $\Delta(t) \to 0$.

For the first step we introduce

$$\Delta^*(t) := \sup_{\tau_2 \leq u \leq t} \Delta(u). \tag{5.18}$$

We have $\Delta^*(t) < \infty$ for all $t \geq \tau_2$, since, for $\tau_2 \leq u \leq t$, we have $\zeta_s(X_u, X) \leq C_s(\|X\|_s^s + \|X_u\|_s^s) \leq C_s(\|X\|_s^s + \sup_{\tau_2 \leq u \leq t} \|X_u\|_s^s) < \infty$ with a constant $C_s > 0$, using [38, Lemma 2]. By definition, $\Delta^*$ is monotonically increasing. With $R := \sup_{t\geq\tau_2} r_t < \infty$ we obtain for $\tau_2 \leq u \leq t$, from (5.17), (5.18) and (5.13),

$$\begin{aligned}\Delta(u) &\leq \mathbb{E}\sum_{r=1}^{K} \mathbf{1}_{\left\{T_r^{(u)}\geq\tau_2\right\}} \left\| A_r^{(u)} \right\|_{\mathrm{op}}^{s} \Delta^*(u) + R \\ &\leq \xi_+ \Delta^*(t) + R.\end{aligned}$$

Hence, we obtain $\Delta^*(t) \leq \xi_+\Delta^*(t) + R$, thus $\Delta^*(t) \leq R/(1-\xi_+)$. This implies

$$\Delta^*(\infty) := \sup_{t\geq\tau_2} \Delta(t) \leq \frac{R}{1-\xi_+} < \infty. \tag{5.19}$$

For the second step we denote $L := \limsup_{t\to\infty} \Delta(t)$. For every $\varepsilon > 0$ there exists a $\tau_3 > \tau_2$ such that we have $\Delta(t) \leq L + \varepsilon$ for all $t \geq \tau_3$. Thus, from (5.17) we obtain

$$\Delta(t) \leq \mathbb{E}\sum_{r=1}^{K} \mathbf{1}_{\left\{\tau_2\leq T_r^{(t)}<\tau_3\right\}} \left\| A_r^{(t)} \right\|_{\mathrm{op}}^{s} \Delta^*(\infty) + \mathbb{E}\sum_{r=1}^{K} \mathbf{1}_{\left\{T_r^{(t)}\geq\tau_3\right\}} \left\| A_r^{(t)} \right\|_{\mathrm{op}}^{s} (L+\varepsilon) + r_t$$

and letting $t \to \infty$ we obtain by (5.9) and (5.12)

$$L \leq \xi(L+\varepsilon).$$

If $L > 0$, this is a contradiction for $0 < \varepsilon < L(1-\xi)/\xi$. Hence, we have $L = 0$. This proves (5.10). Finally, recall that convergence in $\zeta_s$ implies weak convergence. $\square$

As a corollary we formulate a univariate central limit theorem that corresponds to Neininger and Rüschendorf [33, Corollary 5.2] for the discrete time case. For this we assume that there are expansions, as $t \to \infty$,

$$\mathbb{E}\, Y_t = f(t) + o(g^{1/2}(t)), \quad \mathrm{Var}(Y_t) = g(t) + o(g(t)) \tag{5.20}$$

with functions $f : [0,\infty) \to \mathbb{R}$, $g : [0,\infty) \to [0,\infty)$, with

$$\sup_{u\leq t} |f(u)| < \infty \text{ for every } t > 0, \quad \lim_{t\to\infty} g(t) = \infty, \quad \sup_{u\leq t} g(u) = O(g(t)). \tag{5.21}$$

Thus, for some constant $C \geq 1$, $g(u) \leq Cg(t)$ when $0 \leq u \leq t$.

Then the following central limit law holds:

**Corollary 5.2** *Let* $2 < s \leq 3$ *and let* $Y_t$, $t \geq 0$, *be given* $s$*-integrable, univariate random variables satisfying* (5.1) *with* $A_r(t) = 1$ *for all* $r = 1, \ldots, K$ *and* $t \geq 0$.

*Assume that* $\sup_{u \le t} \mathbb{E}|Y_u|^s < \infty$ *for every* $t$*, and that the mean and variance of* $Y_t$ *satisfy* (5.20) *with* (5.21). *If, as* $t \to \infty$*,*

$$\left( \sqrt{\frac{g\left(T_1^{(t)}\right)}{g(t)}}, \ldots, \sqrt{\frac{g\left(T_K^{(t)}\right)}{g(t)}} \right) \xrightarrow{\ell_s} (A_1^*, \ldots, A_K^*), \tag{5.22}$$

$$\frac{1}{g^{1/2}(t)} \left( b_t - f(t) + \sum_{r=1}^{K} f\left(T_r^{(t)}\right) \right) \xrightarrow{\ell_s} 0, \tag{5.23}$$

*and furthermore*

$$\sum_{r=1}^{K} (A_r^*)^2 = 1 \; a.s., \quad \mathbb{P}\left( \bigcup_{r=1}^{K} \{A_r^* = 1\} \right) < 1, \tag{5.24}$$

*then*

$$\frac{Y_t - f(t)}{g^{1/2}(t)} \xrightarrow{\mathrm{d}} \mathcal{N}(0, 1). \tag{5.25}$$

*Proof* We begin by replacing $g(t)$ by $\max(g(t), 1)$; by (5.21), this does not affect $g(t)$ for large $t$, and it is easy to see that (5.20), (5.21), (5.22), (5.23) still hold. We may thus assume that $g(t) \ge 1$ for every $t$.

Denote $M_t := \mathbb{E}\, Y_t$ and $\sigma_t^2 := \mathrm{Var}(Y_t)$. By (5.20), $\sigma_t^2/g(t) \to 1$. (All unspecified limits are as $t \to \infty$.) Choose $\tau_1 \ge \tau_0$ such that $\frac{1}{4} g(t) \le \sigma_t^2 \le 4g(t)$ for $t \ge \tau_1$. Let, as in Convention C, $C_t := 1$ for $t < \tau_1$ and $C_t := \sigma_t^2$ for $t \ge \tau_1$, and write $\tilde{\sigma}_t := C_t^{1/2}$ and $\varepsilon(t) := \tilde{\sigma}_t / g(t)^{1/2} - 1 = (C_t/g(t))^{1/2} - 1$. For $t \ge \tau_1$, $\varepsilon(t) = (\mathrm{Var}\, Y_t / g(t))^{1/2} - 1$, so by (5.20),

$$\varepsilon(t) \to 0 \quad \text{as } t \to \infty. \tag{5.26}$$

Further, $C_t/g(t) = 1/g(t) \le 1$ for $t < \tau_1$, while $C_t/g(t) = \sigma_t^2/g(t) \le 4$ for $t \ge \tau_1$. Hence $|\varepsilon(t)| \le 1$ for all $t$. With (5.4) and $A_r(t) = 1$ we have, for $t \ge \tau_1$,

$$A_r^{(t)} = \frac{\tilde{\sigma}_{T_r^{(t)}}}{\sigma_t} = \frac{\left(1 + \varepsilon\left(T_r^{(t)}\right)\right) g\left(T_r^{(t)}\right)^{1/2}}{\sigma_t}, \tag{5.27}$$

$$b^{(t)} = \sigma_t^{-1} \left( b_t - M_t + \sum_{r=1}^{K} M_{T_r^{(t)}} \right). \tag{5.28}$$

Since $g(T_r^{(t)}) \leq Cg(t)$ by (5.21), we have, for $t \geq \tau_1$,

$$\left\| A_r^{(t)} - \frac{g\left(T_r^{(t)}\right)^{1/2}}{\sigma_t} \right\|_s = \left\| \varepsilon\left(T_r^{(t)}\right) \frac{g\left(T_r^{(t)}\right)^{1/2}}{\sigma_t} \right\|_s \leq \sup_{u \leq t} \left| \varepsilon(u) \frac{g(u)^{1/2}}{\sigma_t} \right|. \quad (5.29)$$

For any $\delta > 0$, there exists, by (5.26), $\tau(\delta) \geq \tau_1$ such that $|\varepsilon(t)| \leq \delta$ when $t \geq \tau(\delta)$. Thus, if $\tau(\delta) \leq u \leq t$, then

$$\left| \varepsilon(u) \frac{g(u)^{1/2}}{\sigma_t} \right| \leq \delta \frac{Cg(t)^{1/2}}{\sigma_t} \leq 2C\delta.$$

On the other hand, if $u \leq \tau(\delta)$, then

$$\left| \varepsilon(u) \frac{g(u)^{1/2}}{\sigma_t} \right| \leq \frac{Cg(\tau(\delta))^{1/2}}{\sigma_t} \to 0$$

as $t \to \infty$. Hence, $\sup_{u \leq t} |\varepsilon(u) g(u)^{1/2}/\sigma_t| \leq 2C\delta$ for sufficiently large $t$. Since $\delta > 0$ is arbitrary, it follows that the right hand side of (5.29) tends to 0 as $t \to \infty$, and thus (5.29) yields

$$\left\| A_r^{(t)} - \frac{g\left(T_r^{(t)}\right)^{1/2}}{\sigma_t} \right\|_s \to 0. \quad (5.30)$$

Since $g(t)^{1/2}/\sigma_t \to 1$, (5.22) yields $g(T_r^{(t)})^{1/2}/\sigma_t \xrightarrow{\ell_s} A_r^*$, which combined with (5.30) yields $A_r^{(t)} \xrightarrow{\ell_s} A_r^*$, jointly for $r = 1, \dots, k$.

Next, for any $\varepsilon > 0$, there exists by (5.20) $\tau_\varepsilon \geq \tau_1$ such that $|M_t - f(t)| \leq \varepsilon g(t)^{1/2}$ if $t \geq \tau_\varepsilon$. Consequently, if $T_r^{(t)} \geq \tau_\varepsilon$, then

$$|M_{T_r^{(t)}} - f\left(T_r^{(t)}\right)| \leq \varepsilon g\left(T_r^{(t)}\right)^{1/2} \leq C\varepsilon g(t)^{1/2}.$$

Since $\sup_{u \leq \tau_\varepsilon} |M_u|$ and $\sup_{u \leq \tau_\varepsilon} |f(u)|$ are finite, the same estimate holds for $T_r^{(t)} < \tau_\varepsilon$ too, provided $t$ is large. Consequently, $|M_{T_r^{(t)}} - f(T_r^{(t)})|/g(t)^{1/2} \leq C\varepsilon$ if $t$ is large enough. It follows that $\|M_{T_r^{(t)}} - f(T_r^{(t)})\|_s/g(t)^{1/2} \to 0$ as $t \to \infty$, so by (5.28), (5.23) and (5.20), $b^{(t)} \xrightarrow{\ell_s} 0$.

We apply Theorem 5.1 with $2 < s \leq 3$; we have shown that (5.7) holds with $b^* = 0$. The two assumptions in (5.24) and $s > 2$ ensure that we have $\mathbb{E} \sum_{r=1}^{K} |A_r^*|^s < 1$.

Finally, by (5.30), for every $\tau$ and $r$,

$$\left\|\mathbf{1}_{\left\{T_r^{(t)}\le\tau\right\}}A_r^{(t)}\right\|_s \le \left\|\mathbf{1}_{\left\{T_r^{(t)}\le\tau\right\}}\frac{g\left(T_r^{(t)}\right)^{1/2}}{\sigma_t}\right\|_s + \left\|A_r^{(t)} - \frac{g\left(T_r^{(t)}\right)^{1/2}}{\sigma_t}\right\|_s$$
$$\le \frac{Cg(\tau)^{1/2}}{\sigma_t} + o(1) \to 0.$$

Now, Theorem 5.1 implies $(Y_t - M_t)/\sigma_t \xrightarrow{\mathrm{d}} X$, where $\mathcal{L}(X)$ is characterized by $\|X\|_s < \infty$, $\mathbb{E}\,X = 0$, $\mathrm{Var}(X) = 1$, and

$$X \stackrel{d}{=} \sum_{r=1}^{K} A_r^* X^{(r)}, \tag{5.31}$$

with assumptions as in (5.5). Since $\sum_{r=1}^{K}(A_r^*)^2 = 1$ this is solved by $\mathcal{L}(X) = \mathcal{N}(0, 1)$. Consequently,

$$\frac{Y_t - M_t}{\sigma_t} \xrightarrow{\mathrm{d}} \mathcal{N}(0, 1),$$

which, in view of (5.20), implies the assertion. □

The following theorem covers cases where the previous central limit theorem of Corollary 5.2 fails due to the appearance of periodic behavior. For this we assume that there is an expansion of the mean, as $t \to \infty$,

$$\mathbb{E}\,Y_t = f(t) + \mathrm{Re}\left(\gamma t^\lambda\right) + o(t^\sigma), \tag{5.32}$$

with a function $f : [0, \infty) \to \mathbb{R}$, $\gamma \in \mathbb{C} \setminus \{0\}$, and $\lambda \in \mathbb{C}$ with $\sigma := \mathrm{Re}(\lambda) > 0$. We denote

$$A_r^{(t)} := \left(\frac{T_r^{(t)}}{t}\right)^\lambda, \quad r = 1, \ldots, K, \tag{5.33}$$

$$b^{(t)} := \frac{1}{t^\sigma}\left(b_t - f(t) + \sum_{r=1}^{K} f\left(T_r^{(t)}\right)\right). \tag{5.34}$$

Note that $A_r^{(t)}$ in general is complex, while $b^{(t)}$ is real.

**Theorem 5.3** *Let $Y_t$, $t \ge 0$, be given square-integrable, univariate random variables satisfying* (5.1) *with $A_r(t) = 1$ for all $r = 1, \ldots, K$ and $t \ge 0$. Assume that $\sup_{u\le t} \mathbb{E}\,|Y_u|^2 < \infty$ for every $t > 0$ and that the mean of $Y_t$ satisfies* (5.32) *with $\lambda = \sigma + \mathrm{i}\tau$ and $\sigma > 0$, and some locally bounded function $f(t)$. If, as $t \to \infty$,*

$$(A_1^{(t)}, \ldots, A_K^{(t)}) \xrightarrow{\ell_2} (A_1^*, \ldots, A_K^*) \quad \text{and} \quad \|b^{(t)}\|_2 \to 0, \tag{5.35}$$

*and furthermore*

$$\mathbb{E}\sum_{r=1}^{K}|A_r^*|^2 < 1, \tag{5.36}$$

*then, as $t \to \infty$,*

$$\ell_2\left(\frac{Y_t - f(t)}{t^\sigma}, \operatorname{Re}\left(Xe^{\mathrm{i}\tau \ln t}\right)\right) \to 0, \tag{5.37}$$

*where $\mathcal{L}(X)$ is the unique fixed point in $\mathcal{M}_2^{\mathbb{C}}(\gamma)$ of*

$$T : \mathcal{M}^{\mathbb{C}} \to \mathcal{M}^{\mathbb{C}}, \quad \eta \mapsto \mathcal{L}\left(\sum_{r=1}^{K} A_r^* Z^{(r)}\right), \tag{5.38}$$

*where $(A_1^*, \dots, A_K^*)$, $Z^{(1)}, \dots, Z^{(K)}$ are independent and $\mathcal{L}(Z^{(r)}) = \eta$ for $r = 1, \dots, K$.*

*Proof* We extend an approach based on the contraction method from Fill and Kapur [16]. We may assume that $\tau_0 \geq 1$.

First, for technical convenience we show that we further may assume $Y_t = 0$ and $f(t) = 0$ for $0 \leq t \leq 1$. Let $(Y_t^{(r)*})_t$, $r = 1, \dots, K$, be another set of copies of $(Y_t)_t$, independent of each other and of everything else. We may replace $Y_t^{(r)}$ in (5.1) by $Y_t^{(r)}\mathbf{1}_{\{t\geq 1\}} + Y_t^{(r)*}\mathbf{1}_{\{t<1\}}$, which has the same distribution and independence properties. Hence $Y_t^{(r)}\mathbf{1}_{\{t\geq 1\}}$ satisfies (5.1) (for $t \geq \tau_0 \geq 1$) with $b_t$ replaced by $\tilde{b}_t := b_t + \sum_r Y^{(r)*}_{T_r^{(t)}}\mathbf{1}_{\{T_r^{(t)}<1\}}$. This replaces $b^{(t)}$ by $\tilde{b}^{(t)}$ with

$$\left|\tilde{b}^{(t)} - b^{(t)}\right| \leq t^{-\sigma}\sum_r \left|Y^{(r)*}_{T_r^{(t)}}\right| \mathbf{1}_{\{T_r^{(t)}<1\}}$$

so $\|\tilde{b}^{(t)} - b^{(t)}\|_2 = O(t^{-\sigma})$ and (5.35) still holds. We may thus consider $Y_t^{(r)}\mathbf{1}_{\{t\geq 1\}}$ instead, and thus we may assume that $Y_t^{(r)} = 0$ when $t < 1$. Similarly, we may assume that $f(t) = 0$ for $t < 1$, changing $b^{(t)}$ by $O(t^{-\sigma})$.

With $X_t := (Y_t - f(t))/t^\sigma$ for $t > 0$ and $X_0 := 0$ we obtain

$$X_t \stackrel{\mathrm{d}}{=} \sum_{r=1}^{K}\left(\frac{T_r^{(t)}}{t}\right)^{\sigma} X^{(r)}_{T_r^{(t)}} + b^{(t)}, \quad t \geq \tau_0, \tag{5.39}$$

with $b^{(t)}$ as given in (5.34).

Next we prove that the restriction of $T$ defined in (5.38) to $\mathcal{M}_2^{\mathbb{C}}(\gamma)$ maps into $\mathcal{M}_2^{\mathbb{C}}(\gamma)$ and is Lipschitz in $\ell_2$ with Lipschitz constant bounded by $\left(\mathbb{E}\sum_{r=1}^{K}|A_r^*|^2\right)^{1/2} < 1$.

Note that (5.36) implies $\|A_r^*\|_2 < \infty$ for all $r = 1, \ldots, K$. This implies that $T(\eta)$ has a finite second moment for all $\eta \in \mathcal{M}_2^{\mathbb{C}}$. Next we claim that $\sum_{r=1}^K \mathbb{E}\, A_r^* = 1$. This implies that $T(\eta)$ has mean $\gamma$ for all $\eta \in \mathcal{M}_2^{\mathbb{C}}(\gamma)$. To prove $\sum_{r=1}^K \mathbb{E}\, A_r^* = 1$, note that (5.32) implies $\mathbb{E}\, X_t = \mathrm{Re}(\gamma t^{\mathrm{i}\tau}) + o(1)$ as $t \to \infty$. On the other hand, the right hand side of (5.39) has mean, using $\mathbb{E}\, b^{(t)} \to 0$,

$$\begin{aligned}
\sum_{r=1}^K \mathbb{E}\left[\left(\frac{T_r^{(t)}}{t}\right)^{\sigma} \mathrm{Re}\left(\gamma \left(T_r^{(t)}\right)^{\mathrm{i}\tau}\right)\right] + o(1) &= \mathrm{Re}\left(\gamma \sum_{r=1}^K \mathbb{E}\, \frac{\left(T_r^{(t)}\right)^{\lambda}}{t^{\sigma}}\right) + o(1) \\
&= \mathrm{Re}\left(\gamma t^{\mathrm{i}\tau} \sum_{r=1}^K \mathbb{E}\left(\frac{T_r^{(t)}}{t}\right)^{\lambda}\right) + o(1) \\
&= \mathrm{Re}\left(\gamma t^{\mathrm{i}\tau} \sum_{r=1}^K \mathbb{E}\, A_r^*\right) + o(1),
\end{aligned}$$

where we also used that $\mathbb{E}(T_r^{(t)}/t)^{\lambda} \to \mathbb{E}\, A_r^*$, see (5.35). Hence, together we obtain, as $t \to \infty$,

$$\mathrm{Re}(\gamma t^{\mathrm{i}\tau}) + o(1) = \mathrm{Re}\left(\gamma t^{\mathrm{i}\tau} \sum_{r=1}^K \mathbb{E}\, A_r^*\right) + o(1). \tag{5.40}$$

Thus, $\gamma \neq 0$ yields $\sum_{r=1}^K \mathbb{E}\, A_r^* = 1$. For the bound on the Lipschitz constant in $\ell_2$ of $T$ restricted to $\mathcal{M}_2^{\mathbb{C}}$ see Rösler and Rüschendorf [36, Lemma 1] and Fill and Kapur [16]: For $\mu, \nu \in \mathcal{M}_2^{\mathbb{C}}$ choose $(Z^{(1)}, W^{(1)}), \ldots, (Z^{(K)}, W^{(K)})$ as identically distributed vectors of optimal couplings of $\mu$ and $\nu$ and such that $(Z^{(1)}, W^{(1)}), \ldots, (Z^{(K)}, W^{(K)})$, $(A_1^*, \ldots, A_K^*)$ are independent. Then we have

$$\begin{aligned}
&\ell_2^2(T(\mu), T(\nu)) \\
&\quad = \ell_2^2\left(\sum_{r=1}^K A_r^* Z^{(r)}, \sum_{r=1}^K A_r^* W^{(r)}\right) \\
&\quad \leq \mathbb{E}\left|\sum_{r=1}^K A_r^* \left(Z^{(r)} - W^{(r)}\right)\right|^2 \\
&\quad = \mathbb{E}\left(\sum_{r=1}^K |A_r^*|^2 |Z^{(r)} - W^{(r)}|^2 + \sum_{r \neq s} A_r^* \left(Z^{(r)} - W^{(r)}\right) \overline{A_s^* \left(Z^{(s)} - W^{(s)}\right)}\right)
\end{aligned}$$

$$= \mathbb{E} \sum_{r=1}^{K} |A_r^*|^2 \ell_2^2(\mu, \nu) + 0$$
$$= \sum_{r=1}^{K} \mathbb{E} |A_r^*|^2 \ell_2^2(\mu, \nu).$$

Altogether we obtain that $T$ has a unique fixed point $\mathcal{L}(X)$ in $\mathcal{M}_2^{\mathbb{C}}(\gamma)$.

The fixed point property of $\mathcal{L}(X)$ implies

$$\frac{1}{t^\sigma} \operatorname{Re}\left(t^\lambda X\right) \stackrel{\mathrm{d}}{=} \frac{1}{t^\sigma} \operatorname{Re}\left(\sum_{r=1}^{K} t^\lambda A_r^* X^{(r)}\right). \tag{5.41}$$

where $(A_1^*, \dots, A_K^*)$, $X^{(1)}, \dots, X^{(b)}$ are independent and $\mathcal{L}(X^{(r)}) = \mathcal{L}(X)$ for $r = 1, \dots, K$. We may assume, e.g., by taking optimal couplings, that $\|A_r^{(t)} - A_r^*\|_2 \to 0$ as $t \to \infty$. We choose $X_t^{(r)}$ as optimal couplings to $\operatorname{Re}(t^{\mathrm{i}\tau} X^{(r)})$ (with the right distribution, i.e., the distribution of $X_t$) for $t \geq 0$ and $r = 1, \dots, K$. Clearly, we may assume that, as required, $X_t^{(r)}$, $r = 1, \dots, K$, are independent of each other and of $(T^{(t)}, b_t)_t$.

We denote, for $t > 0$,

$$\Delta(t) := \ell_2\left(\frac{Y_t - f(t)}{t^\sigma}, \operatorname{Re}\left(X e^{\mathrm{i}\tau \ln t}\right)\right) = \ell_2\left(X_t, \frac{1}{t^\sigma} \operatorname{Re}\left(t^\lambda X\right)\right).$$

Using (5.39) and (5.41) we obtain, for $t \geq \tau_0$,

$$\begin{aligned}
\Delta(t) &= \ell_2\left(\sum_{r=1}^{K} \left(\frac{T_r^{(t)}}{t}\right)^\sigma X_{T_r^{(t)}}^{(r)} + b^{(t)}, \frac{1}{t^\sigma} \operatorname{Re}\left(\sum_{r=1}^{K} t^\lambda A_r^* X^{(r)}\right)\right) \\
&\leq \left\| \sum_{r=1}^{K} \left(\left(\frac{T_r^{(t)}}{t}\right)^\sigma X_{T_r^{(t)}}^{(r)} - \frac{1}{t^\sigma} \operatorname{Re}\left(t^\lambda A_r^* X^{(r)}\right)\right) \right\|_2 + \left\| b^{(t)} \right\|_2 \\
&\leq \left\| \sum_{r=1}^{K} \left(\left(\frac{T_r^{(t)}}{t}\right)^\sigma X_{T_r^{(t)}}^{(r)} - \frac{1}{t^\sigma} \operatorname{Re}\left(\left(T_r^{(t)}\right)^\lambda X^{(r)}\right)\right) \right\|_2 + \left\| b^{(t)} \right\|_2 \\
&\quad + \left\| \sum_{r=1}^{K} \left(\frac{1}{t^\sigma} \operatorname{Re}\left(\left(T_r^{(t)}\right)^\lambda X^{(r)}\right) - \frac{1}{t^\sigma} \operatorname{Re}\left(t^\lambda A_r^* X^{(r)}\right)\right) \right\|_2 .
\end{aligned} \tag{5.42}$$

By (5.35) and (5.33) the second and third of the three latter summands tend to zero as $t \to \infty$. We abbreviate

$$W_r^{(t)} := \left(\frac{T_r^{(t)}}{t}\right)^\sigma X_{T_r^{(t)}}^{(r)} - \frac{1}{t^\sigma} \operatorname{Re}\left(\left(T_r^{(t)}\right)^\lambda X^{(r)}\right). \tag{5.43}$$

Hence, (5.42) implies

$$\begin{aligned}\Delta(t) &\le \left(\mathbb{E}\left(\sum_{r=1}^{K} W_r^{(t)}\right)^2\right)^{1/2} + o(1)\\ &= \left(\mathbb{E}\sum_{r=1}^{K}\left(W_r^{(t)}\right)^2 + \mathbb{E}\sum_{\substack{r,s=1\\ r\neq s}}^{K} W_r^{(t)} W_s^{(t)}\right)^{1/2} + o(1). \qquad (5.44)\end{aligned}$$

By the definition of $\Delta(t)$ and the fact that $\left(X_t^{(r)}, \mathrm{Re}(t^{\mathrm{i}\tau} X^{(r)})\right)$ are optimal couplings for all $t > 0$ and $r = 1, \ldots, K$ we obtain

$$\mathbb{E}(W_r^{(t)})^2 = \mathbb{E}\left[\left(\frac{T_r^{(t)}}{t}\right)^{2\sigma} \Delta^2\left(T_r^{(t)}\right)\right]. \qquad (5.45)$$

From (5.32) we obtain

$$\mathbb{E}\, X_t = \frac{1}{t^\sigma}\mathrm{Re}(\gamma t^\lambda) + R(t), \quad t > 0,$$

with $R(t) \to 0$ as $t \to \infty$. Since $\mathbb{E}\, X^{(r)} = \gamma$ and by the independence conditions we obtain $\mathbb{E}\, W_r^{(t)} = \mathbb{E}[(T_r^{(t)}/t)^\sigma R(T_r^{(t)})]$ and, for $r \neq s$,

$$\mathbb{E}\left[W_r^{(t)} W_s^{(t)}\right] = \mathbb{E}\left[\left(\frac{T_r^{(t)}}{t}\frac{T_s^{(t)}}{t}\right)^{\sigma} R\left(T_r^{(t)}\right) R\left(T_s^{(t)}\right)\right].$$

Splitting the latter integral into the events $\{T_r^{(t)} \le t_1 \text{ or } T_s^{(t)} \le t_1\}$ and $\{T_r^{(t)} > t_1$ and $T_s^{(t)} > t_1\}$ for some $t_1 > 0$ we obtain, for every $t_1 > 0$,

$$\left|\mathbb{E}\left[W_r^{(t)} W_s^{(t)}\right]\right| \le \left(\frac{t_1}{t}\right)^\sigma \|R\|_\infty^2 + \sup_{u \ge t_1} R^2(u),$$

where $\|R\|_\infty := \sup_t |R(t)| < \infty$. From this we obtain first, letting $t \to \infty$, $\limsup_{t\to\infty}\left|\mathbb{E}[W_r^{(t)} W_s^{(t)}]\right| \le \sup_{u \ge t_1} R^2(u)$, and then, letting $t_1 \to \infty$,

$$\mathbb{E}\left[W_r^{(t)} W_s^{(t)}\right] \to 0 \quad \text{as } t \to \infty. \qquad (5.46)$$

Now, (5.44), (5.45), and (5.46) imply, for $t > \tau_0$,

$$\Delta(t) \le \left(\mathbb{E}\left[\sum_{r=1}^{K}\left(\frac{T_r^{(t)}}{t}\right)^{2\sigma} \Delta^2\left(T_r^{(t)}\right)\right] + R_1(t)\right)^{1/2} + R_2(t), \qquad (5.47)$$

with $R_1(t), R_2(t) \to 0$ as $t \to \infty$.

We first show that $\|\Delta\|_\infty < \infty$. Define $\Delta^*(t) := \sup_{0<u\le t} \Delta(u)$. By the assumptions $\sup_{0\le u\le t} \mathbb{E}\,|Y_u|^2 < \infty$ and $\sup_{0\le u\le t} |f(u)| < \infty$, together with $Y_u = 0$ and $f(u) = 0$ for $u \le 1$, we have $\Delta^*(t) < \infty$ for all $t > 0$. Let $t_1 \ge \tau_0$ be such that $|R_1(t)| < 1$ and $|R_2(t)| < 1$ for $t \ge t_1$. Then with (5.47) we obtain, for $t \ge t_1$,

$$\Delta(t) \le \left(\mathbb{E}\left[\sum_{r=1}^K \left(\frac{T_r^{(t)}}{t}\right)^{2\sigma} (\Delta^*)^2(t)\right] + 1\right)^{1/2} + 1.$$

By (5.33), (5.35) and (5.36) there exists a $t_2 \ge t_1$ such that for all $t \ge t_2$ we have $\mathbb{E}\sum_{r=1}^K (T_r^{(t)}/t)^{2\sigma} \le \xi < 1$. Thus, for all $t \ge t_2$ we obtain, with $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}$ for $a, b \ge 0$,

$$\Delta(t) \le \sqrt{\xi}\,\Delta^*(t) + 2,$$

and thus

$$\Delta^*(t) \le \sqrt{\xi}\,\Delta^*(t) + 2 + \Delta^*(t_2),$$

which implies $\|\Delta\|_\infty \le (2 + \Delta^*(t_2))\,/(1 - \sqrt{\xi}) < \infty$.

In a second step we show that $\Delta(t) \to 0$ as $t \to \infty$. For this we assume that $L := \limsup_{t\to\infty} \Delta(t) > 0$. Let $\varepsilon > 0$. There exists a $t_3 \ge t_2$ such that for all $t \ge t_3$ we have $\Delta(t) \le L + \varepsilon$. Then (5.47) implies

$$\begin{aligned}
&\Delta(t)\\
&\le \left(\mathbb{E}\left[\sum_{r=1}^K \left(\frac{T_r^{(t)}}{t}\right)^{2\sigma} \left(\mathbf{1}_{\{T_r^{(t)}<t_3\}} + \mathbf{1}_{\{T_r^{(t)}\ge t_3\}}\right) \Delta^2\left(T_r^{(t)}\right)\right] + R_1(t)\right)^{1/2} + R_2(t)\\
&\le \left(\sum_{r=1}^K \left(\frac{t_3}{t}\right)^{2\sigma} \|\Delta\|_\infty^2 + \xi(L+\varepsilon)^2 + R_1(t)\right)^{1/2} + R_2(t).
\end{aligned}$$

Hence, $t \to \infty$ implies

$$L \le \sqrt{\xi}(L + \varepsilon),$$

which if $L > 0$ is a contradiction if we choose $\varepsilon$ small enough. Consequently, we have $L = 0$ yielding the assertion. □

*Remark 5.1* Note, that $\ell_2$ convergence implies convergence of second moments. Hence in the situation of Theorem 5.3 we also obtain the first order asymptotic term of the expansion of $\operatorname{Var} Y_t$:

$$\operatorname{Var} Y_t \sim t^{2\sigma} \operatorname{Var}\left(\operatorname{Re}\left(X e^{\mathrm{i}\tau \ln t}\right)\right) = \tfrac{1}{2} t^{2\sigma}\left(\mathbb{E}\,|X-\gamma|^2 + \operatorname{Re}\left(e^{2\mathrm{i}\tau \ln t}\,\mathbb{E}(X-\gamma)^2\right)\right).$$

## 6 Proof of Theorem 1.3

In this section we prove Theorem 1.3. The statements on mean and variance of $N(x)$ are proved in Sect. 3. It remains to identify the asymptotic distribution of $N(x)$ Note that recurrence (1.3) for $N(x)$ is covered by the general recurrence for $Y_t$ in (5.1) by making the choices $d = 1$, $K = b$, $\tau_0 = 1$, $A_r(t) = 1$, $T_r^{(t)} = V_r t$ and $b_t = 1$ for all $r = 1, \dots, K$ and $t \geq \tau_0$.

We consider the three cases (i)–(iii) appearing in Theorem 1.3 separately:

*Case (i)*: Theorem 3.1 yields $\mathbb{E}\, N(x) = \alpha^{-1}x + o(\sqrt{x})$ and $\mathrm{Var}(N(x)) \sim \beta x$ with $\beta > 0$. We apply Corollary 5.2 with the choices $f(t) = \alpha^{-1}t$ and $g(t) = \beta t$. The conditions (5.20) and (5.21) are satisfied. We have $\sup_{u \leq t} \mathbb{E}\, |Y_u|^s < \infty$ for $s = 3$ by Lemma 3.5. Condition (5.22) is satisfied with $A_r^* = \sqrt{V_r}$ for $r = 1, \dots, K$, condition (5.23) is trivially satisfied, and we have (5.24). Hence, Corollary 5.2 applies and yields

$$\frac{N(x) - \alpha^{-1}x}{\sqrt{\beta x}} \stackrel{\mathrm{d}}{\to} \mathcal{N}(0, 1),$$

which is the assertion.

*Case (ii)*: Theorem 3.1 yields $\mathbb{E}\, N(x) = \alpha^{-1}x + O(\sqrt{x})$ and $\mathrm{Var}(N(x)) \sim \beta x \ln x$ with $\beta > 0$. We apply Corollary 5.2 with the choices $f(t) = \alpha^{-1}t$ and $g(t) = \beta t \ln t$. Now we have $g(T_r^{(t)})/g(t) = V_r + V_r \ln(V_r)/\ln t$, hence we obtain, since $x \mapsto x \ln x$ is bounded on $[0, 1]$,

$$\left(\sqrt{\frac{g\left(T_1^{(t)}\right)}{g(t)}}, \dots, \sqrt{\frac{g\left(T_K^{(t)}\right)}{g(t)}}\right) \xrightarrow{\ell_3} (A_1^*, \dots, A_K^*),$$

with $A_r^* = \sqrt{V_r}$ for $r = 1, \dots, K$. All conditions of Corollary 5.2 are satisfied as in case (i) and we obtain

$$\frac{N(x) - \alpha^{-1}x}{\sqrt{\beta x \ln x}} \stackrel{\mathrm{d}}{\to} \mathcal{N}(0, 1).$$

*Case (iii)*: Theorem 3.1 yields $\mathbb{E}\, N(x) = \alpha^{-1}x + \mathrm{Re}(\gamma x^{\lambda_2}) + o(x^{\sigma})$, where $\sigma = \mathrm{Re}\,\lambda_2$, which verifies (5.32) (with $\lambda = \lambda_2$). We apply Theorem 5.3 with $f(x) = \alpha^{-1}x$. We have $A_r^{(t)} = (T_r^{(t)}/t)^{\lambda_2} = V_r^{\lambda_2}$ for all $t \geq 1$ and $r = 1, \dots, K$, so $A_r^* = V_r^{\lambda_2}$. Further, $b_t = 1$ and $f(t) = \sum_{r=1}^K f(T_r^{(t)})$, so $b^{(t)} = t^{-\sigma}$ and $\|b^{(t)}\|_2 = t^{-\sigma} \to 0$ as $t \to \infty$. Finally, $\mathbb{E} \sum_{r=1}^K |A_r^*|^2 = \mathbb{E} \sum_{r=1}^K V_r^{2\sigma} < 1$ since $\sigma > 1/2$. Thus all conditions of Theorem 5.3 are satisfied and we obtain

$$\ell_2\left(\frac{N(x) - \alpha^{-1}x}{x^{\mathrm{Re}\,\lambda_2}}, \mathrm{Re}\left(\Xi e^{\mathrm{i}\,\mathrm{Im}\,\lambda_2 \ln x}\right)\right) \to 0$$

as $x \to \infty$. This completes the proof except for the explicit rate of convergence in Theorem 1.3 (iii).

Now, we give a refined version of the proof of Theorem 5.3 for the special recurrence (1.3) which yields also the stated rate of convergence.

The restriction of $T$ defined in (1.7) to $\mathcal{M}_2^{\mathbb{C}}(\nu)$ is Lipschitz in $\ell_2$ with Lipschitz constant bounded by $\left(\mathbb{E}\sum_{r=1}^b V_r^{2\,\mathrm{Re}(\lambda_2)}\right)^{1/2}$; cf. the first part of the proof of Theorem 5.3. From $\sigma = \mathrm{Re}\,\lambda_2 > 1/2$ we obtain that $T$ has a unique fixed point $\mathcal{L}(\Xi)$ in $\mathcal{M}_2^{\mathbb{C}}(\nu)$.

For $X_t := N(t) - \alpha^{-1}t$ we obtain with (1.3)

$$X_t \stackrel{d}{=} \sum_{r=1}^b X_{V_r t}^{(r)} + 1, \tag{6.1}$$

where $X_t^{(r)}$ are independent distributional copies of $X_t$ also independent of $(V_1, \dots, V_b)$. With the fixed point property of $\Xi$ we have

$$t^\sigma\,\mathrm{Re}\left(\Xi e^{\mathrm{i}\tau \ln x}\right) = \mathrm{Re}(t^{\lambda_2}\Xi) \stackrel{d}{=} \mathrm{Re}\left(\sum_{r=1}^b (V_r t)^{\lambda_2}\Xi^{(r)}\right),$$

where $(V_1, \dots, V_b)$, $\Xi^{(1)}, \dots, \Xi^{(b)}$ are independent and $\mathcal{L}(\Xi^{(r)}) = \mathcal{L}(\Xi)$ for $r = 1, \dots, b$. We choose $X_t^{(r)}$ as optimal couplings to $\mathrm{Re}(t^{\lambda_2}\Xi^{(r)})$ for $t \geq 0$ and $r = 1, \dots, b$ and denote $\Delta(t) := \ell_2(X_t, \mathrm{Re}(t^{\lambda_2}\Xi))$. Note that in the definition of $X_t$ we did not rescale by $t^\sigma$, hence we have to show $\Delta(t) = O(t^\kappa)$.

With $W_r^{(t)} := X_{V_r t}^{(r)} - \mathrm{Re}((V_r t)^{\lambda_2}\Xi^{(r)})$ we obtain, for $t \geq 1$,

$$\begin{aligned}
\Delta(t) &= \ell_2\left(\sum_{r=1}^b X_{V_r t}^{(r)} + 1, \sum_{r=1}^b \mathrm{Re}\left((V_r t)^{\lambda_2}\Xi^{(r)}\right)\right) \\
&\leq \left\{\mathbb{E}\left(\sum_{r=1}^b W_r^{(t)}\right)^2\right\}^{1/2} + 1 \\
&= \left\{\sum_{r=1}^b \mathbb{E}\left(W_r^{(t)}\right)^2 + \sum_{\substack{r,s=1\\ r\neq s}}^b \mathbb{E}[W_r^{(t)}W_s^{(t)}]\right\}^{1/2} + 1.
\end{aligned}$$

Conditioning on $(V_1, \dots, V_b)$ yields $\mathbb{E}(W_r^{(t)})^2 = \mathbb{E}\,\Delta^2(V_r t)$. From $\mathbb{E}\,N(t) = \alpha^{-1}t + \mathrm{Re}(\gamma t^{\lambda_2}) + O(t^\kappa)$ and $\mathbb{E}\,\Xi = \gamma$ we obtain $\mathbb{E}\,W_r^{(t)} = O(t^\kappa)$. Since $W_r^{(t)}$ and $W_s^{(t)}$ are independent for $r \neq s$ conditionally on $(V_1, \dots, V_b)$, it follows that

$$\Delta(t) \leq \left\{\sum_{r=1}^b \mathbb{E}\,\Delta^2(V_r t) + O\left(t^{2\kappa}\right)\right\}^{1/2} + 1, \quad t \geq 1. \tag{6.2}$$

Now, we show that $\Delta(t)/t^\kappa = O(1)$. Note that this implies the assertion. We denote

$$\Psi^*(t) := \sup_{1\le u\le t} \frac{\Delta(u)}{u^\kappa}.$$

Then, (6.2) implies, that for appropriate $R > 0$

$$\Psi^*(t) \le \left\{ \sum_{r=1}^{b} \mathbb{E}\, V_r^{2\kappa} (\Psi^*)^2(t) + R \right\}^{1/2} + 1, \quad t \ge 1,$$

and, with $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}$ for $a, b \ge 0$ and $\xi = \mathbb{E} \sum_{r=1}^{b} V_r^{2\kappa} < 1$ this implies

$$\Psi^*(t) \le \frac{\sqrt{R}+1}{1-\sqrt{\xi}} < \infty.$$

The assertion follows.

## 7 Examples

*Example 7.1 (*Random splitting of intervals) Sibuya and Itoh [37] studied the tree defined by random splitting of intervals, with uniformly distributed splitting points; this is the case $b = 2$ and $\mathbf{V} = (U, 1-U)$, with $U \sim \mathrm{U}(0, 1)$. (See also Brennan and Durrett [6,7]; Kakutani [27] for other properties of such splittings.)

We have

$$\phi(z) = \mathbb{E}\, U^z + \mathbb{E}(1-U)^z = 2\int_0^1 u^z\, du = \frac{2}{1+z}, \quad \operatorname{Re} z > -1,$$

which is a rational function. The characteristic equation (1.5) is $2/(1+\lambda) = 1$, and has the single root $\lambda = 1$. Thus Theorem 1.3 (i) applies and shows asymptotic normality, as shown by Sibuya and Itoh [37]. Further, $\alpha = -\phi'(1) = 1/2$, so Theorem 3.1 (ii) yields $\mathbb{E}\, N(x) = m(x) = 2x + O(x^\delta)$ for every $\delta > 0$. More precisely, Theorem 3.4 yields

$$\mathbb{E}\, N(x) = m(x) = 2x - 1, \quad x \ge 1,$$

which also can be shown directly from (1.2) or from (3.11) [37].

For the asymptotic variance, we obtain from Theorem 3.4 (ii), since $M = 1$ and $a_0 = -1$, using symmetry,

$$\begin{aligned}
\beta &= \alpha^{-1}\left(\mathbb{E}\left(U + 2U \wedge (1-U) + 1 - U\right) - 2 + 1\right) \\
&\quad + 2\alpha^{-2}\left(2\,\mathbb{E}\left(U(\ln(1-U) - \ln U)\mathbf{1}_{\{U<1-U\}}\right)\right) - 2\alpha^{-1} \\
&\quad + \alpha^{-3}\left(2\,\mathbb{E}\left(U \wedge (1-U)\right)\right) - \alpha^{-1} \\
&= 20\,\mathbb{E}\left(U \wedge (1-U)\right) + 16\int_0^{1/2} u\left(\ln(1-u) - \ln(u)\right)\,du - 6 \\
&= 8\ln 2 - 5 \approx 0.545177.
\end{aligned}$$

This can also be obtained from Theorem 3.1 (iii); we have

$$\begin{aligned}
\psi(z,w) &= \mathbb{E}\left((U^z + (1-U)^z)(U^w + (1-U)^w)\right) - \phi(z)\phi(w) \\
&= \frac{2}{1+z+w} + 2B(z+1, w+1) - \frac{4}{(1+z)(1+w)} \\
&= \frac{2}{1+z+w} + 2\frac{\Gamma(z+1)\Gamma(w+1)}{\Gamma(z+w+2)} - \frac{4}{(1+z)(1+w)}
\end{aligned}$$

and thus

$$\begin{aligned}
\psi(1/2 + \mathrm{i}u, 1/2 - \mathrm{i}u) &= 1 + \Gamma(3/2 + \mathrm{i}u)\Gamma(3/2 - \mathrm{i}u) - \frac{4}{|3/2 + \mathrm{i}u|^2} \\
&= 1 + |1/2 + \mathrm{i}u|^2 \frac{\pi}{\cosh \pi u} - \frac{4}{|3/2 + \mathrm{i}u|^2},
\end{aligned}$$

and, since $1 - \phi(z) = (z-1)/(z+1)$,

$$\beta = \frac{1}{\pi}\int_{-\infty}^{\infty}\left(1 + \frac{\pi}{\cosh \pi u}|1/2 + \mathrm{i}u|^2 - \frac{4}{|3/2 + \mathrm{i}u|^2}\right)\frac{|3/2 + \mathrm{i}u|^2}{|1/2 + \mathrm{i}u|^4}\,du,$$

which can be integrated (with some effort) to yield $8\ln 2 - 5$.

Consequently, by Theorem 1.3, we recover the limit theorem by [37]:

$$\frac{N(x) - 2x}{\sqrt{x}} \xrightarrow{\mathrm{d}} \mathcal{N}(0, 8\ln 2 - 5).$$

*Example 7.2* *($m$-ary splitting of intervals)* We can generalize Example 7.1 by splitting each interval into $m$ parts, where $m \geq 2$ is fixed, using $m - 1$ independent, uniformly distributed cut points in each interval. This has been studied by Dean and Majumdar [12].

We have $b = m$, and $V_1, \dots, V_m$ have the same distribution with density $(m-1)(1-x)^{b-2}$, $0 < x < 1$. Hence,

$$\phi(z) = m\,\mathbb{E}\,V_1^z = m(m-1)\int_0^1 x^z(1-x)^z\,dx = m(m-1)B(z+1, m-1)$$
$$= \frac{\Gamma(z+1)m!}{\Gamma(m+z)} = \frac{m!}{(z+1)\cdots(z+m-1)}.$$

The characteristic equation $\phi(z) = 1$ becomes $\Gamma(z+m)/\Gamma(z+1) = m!$, or

$$(z+1)\cdots(z+m-1) = m!. \tag{7.1}$$

The same equation appears in the analysis of $m$-ary search trees. It is shown by Mahmoud and Pittel [31] and Fill and Kapur [17] that if $m \le 26$, then $\operatorname{Re}\lambda_2 < 1/2$, and thus (i) applies, but if $m \ge 27$, then $\operatorname{Re}\lambda_2 > 1/2$, see also, e.g., Chauvin and Pouyanne [8] and Chern and Hwang [10]. Theorem 3.4 yields an exact formula for $\mathbb{E}\,N(x)$ (although it is hardly useful except when $m$ is small). It further leads to a formula for the asymptotic variance, provided $m \le 26$.

We have, with $\psi(z) := \Gamma'(z)/\Gamma(z)$ and $H_z := \psi(z+1) - \psi(1)$ (for integer $z$, these are the harmonic numbers)

$$\alpha = -\phi'(1) = \psi(m+1) - \psi(2) = H_m - 1.$$

*Example 7.3 (*Random splitting of multidimensional intervals) Another generalization is to consider $d$-dimensional intervals, where an interval is split into $2^d$ subintervals by $d$ hyperplanes orthogonal to the coordinate axis and passing through a random, uniformly distributed point. This too has been studied by Dean and Majumdar [12].

We have $b = 2^d$. $V_1, \dots, V_b$ have the same distribution, $V_j \overset{\mathrm{d}}{=} U_1 \cdots U_d$, where $U_k \sim U(0,1)$ are i.i.d. Hence,

$$\phi(z) = 2^d\,\mathbb{E}\,V_1^z = 2^d\left(\mathbb{E}\,U_1^z\right)^d = \left(\frac{2}{1+z}\right)^d.$$

Again, $\phi$ is rational. The characteristic equation may be written $((1+\lambda)/2)^d = 1$, with the roots

$$\Lambda = \{2e^{2\pi\mathrm{i}k/d} - 1 : 0 \le k \le d-1\}.$$

Thus $\sigma_2 := \operatorname{Re}\lambda_2 = 2\cos\frac{2\pi}{d} - 1$, and the condition $\operatorname{Re}\lambda_2 < 1/2$ is equivalent to $\cos(2\pi/d) < 3/4$, which holds for $d \le 8$, while $\operatorname{Re}\lambda_2 > 1/2$ for $d \ge 9$. This justifies the claims in Dean and Majumdar [12].

The same characteristic equation, and the same phase transition, appears for quad trees, see Chern et al. [9].

We further observe that $\alpha = -\phi'(1) = d/2$.

The random trees in these three examples have also been studied by [20], [25] and [26], where the properties of a randomly selected branch are investigated. This problem is quite different, and there is no phase transition. See also [23].

*Example 7.4 (*Random splitting of simplices) Consider $d$-dimensional simplices, where a simplex is split into $d+1$ new simplices by choosing a uniform random point $X$ in the interior and connecting it to the vertices of the original simplex; each new simplex has as vertices $X$ and $d$ of the original $d+1$ vertices.

It is easily seen that this is equivalent to $d+1$-ary splitting as in Example 7.2, see [13, Lemma 3], so we have the same results as there, with $m = d+1$. In particular, $N(x)$ is asymptotically normal if $d \le 25$.

*Example 7.5 (*Non-uniform splitting of intervals) Returning to binary splitting of intervals, we can generalize Example 7.1 by taking another distribution for the cut points; we thus have $b = 2$ and $\mathbf{V} = (V, 1-V)$, where $V$ has any distribution on $(0, 1)$. An interesting case is when $V$ has a beta distribution $V \sim B(a, a')$ with $a, a' > 0$; then

$$\begin{aligned}\mathbb{E}\, V^z &= B(a,a')^{-1} \int_0^1 x^{z+a-1}(1-x)^{a'-1}\,dx = \frac{B(a+z,a')}{B(a,a')}\\ &= \frac{\Gamma(z+a)}{\Gamma(z+a+a')}\frac{\Gamma(a+a')}{\Gamma(a)};\end{aligned}$$

$\mathbb{E}(1-V)^z$ is obtained by interchanging $a$ and $a'$. In particular, if $a$ and $a'$ are integers, then $\phi$ is rational.

We consider two special cases.

(i) The symmetric case with $a' = a$, $V \sim B(a, a)$. Then

$$\phi(z) = 2\frac{\Gamma(z+a)}{\Gamma(z+2a)}\frac{\Gamma(2a)}{\Gamma(a)} = \frac{\Gamma(z+a)}{\Gamma(z+2a)}\frac{\Gamma(1+2a)}{\Gamma(1+a)}.$$

We have $\alpha = -\phi'(1) = H_{2a} - H_a$, with $H_x$ as in Example 7.2. Numerical solution of the characteristic equation seems to show that $\operatorname{Re}\lambda_2 < 1/2$ if and only if $a < a_0$, where $a_0 \approx 59.547$.

(ii) The case $a' = 1$, $V \sim B(a, 1)$. Then

$$\begin{aligned}\phi(z) &= \frac{\Gamma(z+a)}{\Gamma(z+a+1)}\frac{\Gamma(a+1)}{\Gamma(a)} + \frac{\Gamma(z+1)}{\Gamma(z+a+1)}\Gamma(a+1)\\ &= \frac{a}{z+a} + \frac{\Gamma(z+1)}{\Gamma(z+a+1)}\Gamma(a+1).\end{aligned}$$

One finds $\alpha = H_a/(a+1)$. The characteristic equation $\phi(\lambda) = 1$ is equivalent to $\Gamma(a+1)\Gamma(\lambda+1)/\Gamma(\lambda+a+1) = \lambda/(\lambda+a)$ or

$$\frac{\Gamma(a+\lambda)}{\Gamma(\lambda)} = \Gamma(a+1).$$

When $a = m$ is an integer, this is the same as (7.1), so $\operatorname{Re}\lambda_2 < 1/2$ for integer $a$ if and only if $a \le 26$. In general, numerical solution of the characteristic equation seems to show that $\operatorname{Re}\lambda_2 < 1/2$ if and only if $a < a_0$, where $a_0 \approx 26.9$.

## 8 Non-examples

In this section, we give a few examples where our theorems are not valid.

*Example 8.1* (Lattice). In the lattice case, there exists $R > 1$ such that every $V_j \in \{R^{-k} : k \ge 1\} \cup \{0\}$ a.s. In this case, $\phi$ is periodic with period $2\pi\mathrm{i}/\ln R$; in particular, the characteristic equation (1.5) has infinitely many roots $1 + 2\pi\mathrm{i}n/\ln R$ on $\{\lambda : \operatorname{Re}\lambda = 1\}$, and thus Condition B($\delta$) fails. Indeed, it is obvious from (1.2) that $N(x) = N(R^m)$ when $R^m \le x < R^{m+1}$, so $\mathbb{E}\,N(x)/x$ oscillates and does not converge as $x \to \infty$. The natural approach is to consider only $x \in \{R^m : m \ge 0\}$. It is then straightforward to prove an analogue of Theorem 3.1, using the lattice versions of the renewal theory theorems that were used in Sect. 3. An analogue of Theorem 1.3 then follows by the usual (discrete) contraction method, as in [33]. We leave the details to the reader.

*Example 8.2* (Deterministic). If $V = (V_1, \dots, V_b)$ is deterministic, then so is $N(x)$, and it is meaningless to ask for an asymptotic distribution. However, it makes sense to study the asymptotics of $N(x) = m(x)$. (Clearly, $\sigma^2(x) = 0$.)

If $V$ is non-lattice, then $N(x)/x \to \alpha$ by Theorem 3.1 and Remark 3.2. If $V$ is lattice, we consider, as in Example 8.1, only $x = R^m$, $m \ge 1$.

We may assume that $V_j > 0$ for each $j$. By the Kronecker–Weyl theorem, for every $\varepsilon > 0$, there exist arbitrarily large $t$ such that $|V_j^{\mathrm{i}t} - 1| < \varepsilon$ for $j = 1, \dots, b$; thus $\limsup_{t\to\infty} |\phi(1 + \mathrm{i}t)| = 1$. Hence Condition B(1) does not hold, and therefore, by Lemma 2.1, Condition B($\delta$) does not hold for any $\delta \le 1$.

More precisely, if $|V_j^{\mathrm{i}t} - 1| < \varepsilon$ for $j = 1, \dots, b$, let $z_0 = 1 + \mathrm{i}t$. Then $|\phi(z_0) - 1| < \varepsilon$ and

$$|\phi'(z_0) + \alpha| = \left|\sum_{j=1}^{b} \ln V_j (V_j^{1+it} - V_j)\right| \le \varepsilon\alpha.$$

Since further $|\phi''(z)| \le \sum_j |\ln V_j|^2$ for $\operatorname{Re} z \ge 0$, it follows easily that if $\varepsilon$ is small enough, then $\phi(z) - 1$ has a zero in the disc $B := \{z : |z - z_0| < 2\varepsilon/\alpha\}$. (Use the Newton–Raphson method, or Rouché's theorem and a comparison with the linear function $\phi(z_0) + (z - z_0)\phi'(z_0)$.) It follows that there exists a sequence $\lambda_n \in \Lambda$ with $\operatorname{Re}\lambda_n \to 1$ and $\operatorname{Im}\lambda_n \to +\infty$.

We give some concrete examples:

$V = (1/2, 1/2)$ is lattice with $R = 2$ and $N(2^n) = 2^n$.

$V = (\tau^{-1}, \tau^{-2})$ where $\tau = (1 + \sqrt{5})/2$ (the golden ratio) is lattice with $R = \tau$ and $N(\tau^n) = F_{n+3} - 1$, $n \ge 0$, as is easily proven by induction. ($F_n$ denotes the Fibonacci numbers.) Thus, $N(\tau^n) \sim 5^{-1/2}\tau^{n+3}$.

$V = (1/3, 2/3)$ is non-lattice and thus $N(x) \sim \alpha^{-1}x$, where $\alpha = \frac{1}{3}\ln 3 + \frac{2}{3}\ln(3/2) = \ln 3 - \frac{2}{3}\ln 2$.

## 9 Some related models

The basic model may be varied in various ways. We mention here some variations that we find interesting. We do *not* consider these versions in the present paper; we leave the possibility of extensions of our results as an open problem, hoping that these remarks will be an inspiration for future research.

*Remark 9.1* By our assumptions, the label of a node equals the sum of the labels of its children. Another version would be to allow a (possibly random) loss at each node. One important case is *Rényi's parking problem* [35], where a node with label $x$ is interpreted as an interval of length $x$ on a street, where cars of length 1 park at random. Each car splits an interval of length $x \geq 1$ into two free intervals with the lengths $U(x-1)$ and $(1-U)(x-1)$, where $U \sim \mathrm{U}(0,1)$. An obvious generalization is to split $(x-1)$ using an arbitrary random vector $(V_1, \dots, V_b)$. (The one-sided version, where we study only one branch of the tree, is studied in [20], [23].)

*Remark 9.2* Krapivsky et al. [29] have studied a fragmentation process where fragmentation stops stochastically, with a probability $p(x)$ of further fragmentation that in general depends on the mass $x$ of the fragment. Our process is the case $p(x) = \mathbf{1}_{\{x\geq 1\}}$. Another interesting case is $p(x) = 1 - e^{-x}$, see Remark 9.3 below. A different stochastic stopping rule is treated by Gnedin and Yakubovich [19].

*Remark 9.3* Our model is a continuous version of the split trees studied by Devroye [13], where the labels are integers (interpreted as numbers of balls to be distributed in the corresponding subtree) and each label $n$ is, except at the leaves, randomly split according to a certain procedure into $b$ integers summing to $n - s_0$; here $s_0$ is a small positive integer (for example 1) that represents the number of balls stored at the node. Typical examples are binary search trees, $m$-ary search trees and quadtrees. We can regard the continuous model as an approximation of the discrete, or conversely, and it is easy to guess that many properties will have similar asymptotics for the two models. This has been observed in several examples by various authors, see [12] and [9]. For example, the results for Example 7.2 parallel those found for $m$-ary search trees by [8], [10], [17], [31] and others. Similarly, the results in Example 7.3 parallel those found for quadtrees by [9].

We study only the continuous version in this paper. It would be very interesting to be able to rigorously transfer results from the continuous to the discrete version (or conversely); we will, however, not attempt this here.

Note that for binary search trees, we have $n$ random (uniformly distributed) points in an interval, split the interval by the first of these points, and continue recursively splitting each subinterval that contains at least one of the points. If we scale the initial interval to have length $n$, then the probability that a subinterval of length $x$ contains at least one point is $\approx 1 - e^{-x}$. Thus it seems likely that the binary search tree is well approximated by a fragmentation tree, with $\mathbf{V}$ as in Example 7.1, with a fragmentation probability $1 - e^{-x}$ as in Remark 9.2. The same goes for random quadtrees and simplex trees corresponding to Examples 7.3 and 7.4.

**Acknowledgements** This research was initiated and largely done during conferences in Oberwolfach, Frankfurt and Vienna in August and September 2004. We thank Luc Devroye, Jim Fill, Alexander Gnedin, Allan Gut and Hsien-Kuei Hwang for valuable comments, and a referee for careful reading and constructive criticism that has improved our paper.

## References

1. Asmussen, S.: Applied Probability and Queues. Wiley, Chichester (1987)
2. Baryshnikov, Y., Gnedin, A.: Counting intervals in the packing process. Ann. Appl. Probab. **11**, 863–877 (2001)
3. Bertoin, J.: Random Fragmentation and Coagulation Processes. Cambridge University Press, Cambridge (2006)
4. Bertoin, J., Gnedin, A.: Asymptotic laws for nonconservative self-similar fragmentations. Electron. J. Probab. **9**, 575–593 (2004)
5. Bickel, P.J., Freedman, D.A.: Some asymptotic theory for the bootstrap. Ann. Statist. **9**, 1196–1217 (1981)
6. Brennan, M.D., Durrett, R.: Splitting intervals. Ann. Probab. **14**, 1024–1036 (1986)
7. Brennan, M.D., Durrett, R.: Splitting intervals. II. Limit laws for lengths. Probab. Theory Related Fields **75**, 109–127 (1987)
8. Chauvin, B., Pouyanne, N.: $m$-ary search trees when $m \geq 27$: a strong asymptotics for the space requirements. Random Struct. Algorithms **24**, 133–154 (2004)
9. Chern, H.-H., Fuchs, M., Hwang, H.-K.: Phase changes in random point quadtrees. ACM Trans. Algorithms (to appear) (2006)
10. Chern, H.-H., Hwang, H.-K.: Phase changes in random $m$-ary search trees and generalized quicksort. Random Struct. Algorithms **19**, 316–358 (2001)
11. Dall'Aglio, G.: Sugli estremi dei momenti delle funzioni di ripartizione doppia. Ann. Scuola Norm. Sup. Pisa **10**, 35–74 (1956)
12. Dean, D.S., Majumdar, S.N.: Phase transition in a random fragmentation problem with applications to computer science. J. Phys. A: Math. Gen. **35**, L501–L507 (2002)
13. Devroye, L.: Universal limit laws for depths in random trees. SIAM J. Comput. **28**, 409–432 (1999)
14. Feller, W.: An Introduction to Probability Theory and its Applications, vol. II, 2nd edn. Wiley, New York (1971)
15. Fill, J.A., Janson, S.: A characterization of the set of fixed points of the quicksort transformation. Electronic Comm. Probab. **5**(9), 77–84 (2000)
16. Fill, J.A., Kapur, N.: The space requirement of $m$-ary search trees: distributional asymptotics for $m \geq 27$. In: Proceedings of the 7th Iranian Statistical Conference, Tehran (2004)
17. Fill, J.A., Kapur, N.: Transfer theorems and asymptotic distributional results for $m$-ary search trees. Random Struct. Algorithms **26**, 359–391 (2005)
18. Garnett, J.B.: Bounded Analytic Functions. Academic, New York (1981)
19. Gnedin, A.V., Yakubovich, Y.: Recursive partition structures. Ann. Probab. **34**, 2203–2218 (2006)
20. Itoh, Y., Mahmoud, H.: One-sided variations on interval trees. J. Appl. Prob. **40**, 654–670 (2003)
21. Jagers, P.: Branching Processes with Biological Applications. Wiley, Chichester (1975)
22. Janson, S.: Functional limit theorems for multitype branching processes and generalized Pólya urns. Stochastic Processes Appl. **110**, 177–245 (2004)
23. Janson, S.: One-sided interval trees. J. Iranian Stat. Soc. **3**, 149–164 (2004)
24. Janson, S.: Rounding of continuous random variables and oscillatory asymptotics. Ann. Probab. **34**, 1807–1826 (2006)
25. Javanian, M., Mahmoud, H., Vahidi-Asl, M.: Paths in $m$-ary interval trees. Discrete Math. **287**, 45–53 (2004)
26. Javanian, M., Vahidi-Asl, M.: Multidimensional interval trees. In: Drmota, M., Flajolet, P., Gardy, D., Gittenberger, B. (eds.) Mathematics and Computer Science III, Algorithms, Trees, Combinatorics and Probabilities (Vienna 2004), pp. 255–256. Birkhäuser, Basel (2004)

27. Kakutani, S.: A problem of equidistribution on the unit interval [0, 1]. Measure theory (Oberwolfach, 1975), Lecture Notes in Math., vol. 541, pp. 369–375. Springer, Berlin (1976)
28. Kolmogoroff, A.N.: Über das logarithmisch normale Verteilungsgesetz der Dimensionen der Teilchen bei Zerstückelung. C. R. (Doklady) Acad. Sci. URSS (N. S.) **31**, 99–101 (1941)
29. Krapivsky, P.L., Ben-Naim, E., Grosse, I.: Stable distributions in stochastic fragmentation. J. Phys. A: Math. Gen. **37**, 2863–2880 (2004)
30. Krapivsky, P.L., Grosse, I., Ben-Naim, E.: Scale invariance and lack of self-averaging in fragmentation. Phys. Rev. E **61**, R993–R996 (2000)
31. Mahmoud, H.M., Pittel, B.: Analysis of the space of search trees under the random insertion algorithm. J. Algorithms **10**, 52–75 (1989)
32. Major, P.: On the invariance principle for sums of independent identically distributed random variables. J. Multivariate Anal. **8**, 487–517 (1978)
33. Neininger, R., Rüschendorf, L.: A general limit theorem for recursive algorithms and combinatorial structures. Ann. Appl. Probab. **14**, 378–418 (2004)
34. Rachev, S.T.: Probability Metrics and the Stability of Stochastic Models. Wiley, New York (1991)
35. Rényi, A.: On a one-dimensional random space-filling problem. (Hungarian.) Magyar Tud. Akad. Mat. Kutató Int. Közl. 3, 109–127 (1958). English transl. in Selected papers of Alfréd Rényi, vol. II: 1956–1961. Ed. Pál Turán, (1976) pp. 173–188. Akadémiai Kiadó, Budapest
36. Rösler, U., Rüschendorf, L.: The contraction method for recursive algorithms. Algorithmica **29**, 3–33 (2001)
37. Sibuya, M., Itoh, Y.: Random sequential bisection and its associated binary tree. Ann. Inst. Statist. Math. **39**, 69–84 (1987)
38. Zolotarev, V.M.: Approximation of the distributions of sums of independent random variables with values in infinite-dimensional spaces. (Russian.) Teor. Verojatnost. i Primenen. 21, no. 4, 741–758 (1976). Erratum ibid 22 (1977), no. 4, 901. English transl. Theor. Probability Appl. 21, no. 4, 721–737; 22 (1977), no. 4, 679–691 (1978)
39. Zolotarev, V.M.: Ideal metrics in the problem of approximating the distributions of sums of independent random variables (Russian). Teor. Verojatnost. i Primenen. **22**(3), 449–465 (1977)