The Annals of Applied Probability
2003, Vol. 0, No. 0, 1–41

GENERAL LIMIT THEOREM FOR RECURSIVE ALGORITHMS AND COMBINATORIAL STRUCTURES

BY RALPH NEININGER[1] AND LUDGER RÜSCHENDORF

McGill University and Universität Freiburg

Limit laws are proven by the contraction method for random vectors of a recursive nature as they arise as parameters of combinatorial structures such as random trees or recursive algorithms, where we use the Zolotarev metric. In comparison to previous applications of this method, a general transfer theorem is derived which allows us to establish a limit law on the basis of the recursive structure and to use the asymptotics of the first and second moments of the sequence. In particular, a general asymptotic normality result is obtained by this theorem which typically cannot be handled by the more common ℓ_2 metrics. As applications we derive quite automatically many asymptotic limit results ranging from the size of tries or m-ary search trees and path lengths in digital structures to mergesort and parameters of random recursive trees, which were previously shown by different methods one by one. We also obtain a related local density approximation result as well as a global approximation result. For the proofs of these results we establish that a smoothed density distance as well as a smoothed total variation distance can be estimated from above by the Zolotarev metric, which is the main tool in this article.

1. Introduction. This work gives a systematic approach to limit laws for sequences of random vectors that satisfy distributional recursions as they appear under various models of randomness for parameters of trees, characteristics of divide-and-conquer algorithms or, more generally, for quantities related to recursive structures. Although there are also strong analytic techniques for the subject, we extend and systematize a more probabilistic approach—the contraction method. This method was first introduced for the analysis of Quicksort in [50] and further developed independently in [51] and [49]; see also the survey article by Rösler and Rüschendorf [53]. The name of the method refers to the fact that the analysis makes use of an underlying map of measures, which is a contraction with respect to some probability metric.

Our article is a continuation of the article by Neininger [43], who used the ℓ_2 metric approach to establish a general limit theorem for multivariate divide-and-conquer recursions, thus extending the one-dimensional results in [52]. Although the ℓ_2 approach works well for many problems that lead to nonnormal limit

---

Received October 2001; revised January 2003.

[1]Supported by NSERC Grant A3450 and the Deutsche Forschungsgemeinschaft.

AMS 2000 subject classifications. Primary 60F05, 68Q25; secondary 68P10.

Key words and phrases. Contraction method, multivariate limit law, asymptotic normality, random trees, recursive algorithms, divide-and-conquer algorithm, random recursive structures, Zolotarev metric.

distributions, its main defect is that it does not work for an important class of problems that lead to normal limit laws. We discuss this problem in more detail in Section 2 and, to overcome this problem, we propose to use, as alternative metrics, the Zolotarev metrics ζ_s, which are more flexible and at the same time still manageable. The advantage of alternative metrics such as the Zolotarev metrics for the analysis of algorithms has been demonstrated for some examples in [49] and [10].

The flexibility of the ζ_s metrics is in the fact that although for $s = 2$, we reobtain the common ℓ_2 theory, we also can use ζ_s with $s > 2$, which gives access to normal limit laws, or with $s < 2$, which leads to results where we can weaken the assumption of finite second moments—an assumption that is usually present in the ℓ_2 approach.

In his 1999 article, Pittel [46] stated as a heuristic principle that various global characteristics of large size combinatorial structures such as graphs and trees are asymptotically normal if the mean and variance are "nearly linear" in n. As a technical reason, he argued that the normal distribution with the same two moments "almost" satisfies the recursion. He exemplified this idea by the independence number of uniformly random trees. An essential step in the proof of our limit theorem is the introduction of an accompanying sequence which fulfills approximatively a recursion of the same form as the characteristics do and is formulated essentially in terms of the limiting distribution. This is similar to the technical idea proposed by Pittel [46].

We obtain a general limit theorem for divide-and-conquer recursions where the conditions are formulated in terms of relationships of moments and a condition that ensures the asymptotic stability of the recursive structure. These conditions can quite easily be checked in a series of examples and allow us to (re)derive many examples from the literature. In fact, for the special case of normal limit laws, we need—according to Pittel's principle—the first and second moment to apply the method; see Corollary 5.2.

Several further metrics can be estimated from above by the Zolotarev metric. We prove that, in any dimension, a smoothed density distance and the smoothed total variation distance are estimated from above by a Zolotarev metric. As a consequence we obtain a local density approximation result and a global approximation property for general recursive algorithms.

We investigate sequences of d-dimensional vectors $(Y_n)_{n\in\mathbb{N}_0}$, which satisfy the distributional recursion

$$Y_n \stackrel{\mathcal{D}}{=} \sum_{r=1}^{K} A_r(n) Y^{(r)}_{I_r^{(n)}} + b_n, \qquad n \ge n_0, \tag{1}$$

where $(A_1(n), \dots, A_K(n), b_n, I^{(n)}), (Y_n^{(1)}), \dots, (Y_n^{(K)})$ are independent, $A_1(n), \dots, A_K(n)$ are random $d \times d$ matrices, b_n is a random d-dimensional vector, $I^{(n)}$ is a vector of random cardinalities $I_r^{(n)} \in \{0, \dots, n\}$ and $(Y_n^{(1)}), \dots, (Y_n^{(K)})$

are identically distributed as (Y_n). The $\stackrel{\mathcal{D}}{=}$ denotes equality in distribution and we have $n_0 \geq 1$. Note that we do not define the sequence (Y_n) by (1); we only assume that (Y_n) satisfies the recurrence (1). In our discussion, the number $K \geq 1$ is, for simplicity of presentation, considered first to be fixed. However, in Section 4.3 we state the extension of the main result to random K depending on n. Later, in Section 5.2, we also treat the situation of $K = K_n$ being random with $K_n \to \infty$ almost surely.

This situation is present for many parameters of random structures of a recursive nature like random trees or recursive algorithms. Many examples are given in Section 5. In this context the $I_r^{(n)}$ are the cardinalities of the subproblems generated by the divide-and-conquer algorithm and b_n is the cost to subdivide and merge, also called the toll function. For more background and reference to related work, see [43].

We normalize the Y_n by

$$X_n := C_n^{-1/2}(Y_n - M_n), \qquad n \geq 0, \tag{2}$$

where $M_n \in \mathbb{R}$ and C_n is a positive-definite square matrix. If first or second moments for Y_n are finite, we essentially choose M_n and C_n as the mean and covariance matrix of Y_n, respectively. The X_n satisfy

$$X_n \stackrel{\mathcal{D}}{=} \sum_{r=1}^{K} A_r^{(n)} X_{I_r^{(n)}}^{(r)} + b^{(n)}, \qquad n \geq n_0, \tag{3}$$

with

$$A_r^{(n)} := C_n^{-1/2} A_r(n) C_{I_r^{(n)}}^{1/2}, \qquad b^{(n)} := C_n^{-1/2}\left(b_n - M_n + \sum_{r=1}^{K}\left(A_r(n) M_{I_r^{(n)}}\right)\right) \tag{4}$$

and independence relationships as in (1).

We use the contraction method to prove a limit theorem for the sequence (X_n). Our aim is to establish a transfer theorem of the following form: Appropriate convergence of the coefficients $A_r^{(n)} \to A_r^*$, $b^{(n)} \to b^*$ implies weak convergence of the parameters (X_n) to a limit X. The limit distribution $\mathcal{L}(X)$ satisfies a fixed point equation obtained from (3) by letting formally $n \to \infty$:

$$X \stackrel{\mathcal{D}}{=} \sum_{r=1}^{K} A_r^* X^{(r)} + b^*. \tag{5}$$

Here $(A_1^*, \ldots, A_K^*, b^*)$, $X^{(1)}, \ldots, X^{(K)}$ are independent and $X^{(r)} \sim X$ for $r = 1, \ldots, K$, where $X \sim Y$ denotes equality of the distributions of X and Y.

We show convergence for (X_n) with respect to a Zolotarev metric. This class of metrics is introduced in the next section, where we also explain the necessity for a change from the more common ℓ_2 metric to these metrics by using an example

AAP ims v.2003/08/05 Prn:1/09/2003; 11:54 F:aap165.tex; (Aiste) p. 4

of a fixed-point equation of the form (5) related to the normal distribution. Then we study contraction properties of fixed-point equations (5) with respect to these metrics in Section 3 and give a transfer theorem as desired in Section 4. The rest of the article is devoted to applications of our general transfer result, where we (re)derive various central limit laws for random recursive structures, ranging from the size of m-ary search trees or random tries path lengths in digital search trees, tries and Patricia tries, via top-down mergesort, and the maxima in right triangles to parameters of random recursive trees and plane-oriented versions thereof.

2. The Zolotarev metric. The contraction method applied in this article is based on certain regularity properties of the probability metrics used for proving convergence of the parameters as well as on some lower and upper bounds for these metrics. A probability metric $\tau = \tau(X, Y)$ defined for random vectors X and Y in general depends on the joint distribution of (X, Y). The probability metric τ is called simple if $\tau(X, Y) = \tau(\mu, \nu)$ depends only on the marginal distributions μ and ν of X and Y. Most of the metrics used in this article are simple and therefore induce a metric on (a subclass of) all probability measures on $\mathbb{R}^d$. A imple metric τ is called ideal of order $s > 0$ if

$$\tau(X + Z, Y + Z) \leq \tau(X, Y) \tag{6}$$

for all Z independent of X and Y, and

$$\tau(cX, cY) = |c|^s \tau(X, Y)$$

for all $c \neq 0$. Note that $\tau(X, Y) = \tau(\mu, \nu)$ depends only on the marginal distributions μ and ν of X and Y.

The ℓ_2 metric defined by

$$\ell_2(\mu, \nu) = \inf\{\|X - Y\|_2 : X \sim \mu,\ Y \sim \nu\}$$

has been used frequently in the analysis of algorithms since its introduction in this context by Rösler [50] for the analysis of Quicksort (see, e.g., [41, 43, 44]); note that ℓ_2 is ideal of order 1. This implies that ℓ_2 typically cannot be used for fixed-point equations that occur for the normal distribution such as

$$X \stackrel{d}{=} \frac{1}{\sqrt{2}} X_1 + \frac{1}{\sqrt{2}} X_2 \stackrel{d}{=}: TX, \tag{7}$$

where X_r are independent copies of X. Here we consider T as a map $T : \mathcal{M} \to \mathcal{M}$ on the space $\mathcal{M}$ of univariate probability measures with $T\mu := \mathcal{L}(Z_1/\sqrt{2} + Z_2/\sqrt{2})$, where Z_1 and Z_2 are independent with distribution μ, and we abbreviate $TX := T\mathcal{L}(X)$. For centered X and Y with finite second moment, we may choose independent optimal couplings $(X_1, Y_1), (X_2, Y_2)$ of (X, Y), that is, vectors that satisfy $\ell_2(X,Y) = \|X_r - Y_r\|_2$, $r = 1, 2$. Then we have

$$\begin{aligned}\ell_2^2(TX,TY) &\le \left\|\frac{1}{\sqrt{2}}(X_1 - Y_1)\right\|_2^2 + \left\|\frac{1}{\sqrt{2}}(X_2 - Y_2)\right\|_2^2\\ &= \left(\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2\right)\ell_2^2(X,Y)\\ &= \ell_2^2(X,Y).\end{aligned}$$

This suggests that restriction of T to the space $\mathcal{M}_2^1(0)$ of univariate centered probability measures with finite second moment is not a strict contraction in ℓ_2, which would be essential for application of the contraction method. In fact, T is not a contraction on $\mathcal{M}_2^1(0)$ in any metric. This results from the fact that all the centered normal distributions are (exactly) the fixed points of T. On the other hand, strict contraction would imply uniqueness of the fixed point.

The basic idea is to refine the working space. We restrict T to the subspace $\mathcal{M}_s^1(0,\sigma^2) \subset \mathcal{M}_2^1(0)$, $2 < s \le 3$, where the variance of the measures is fixed to be $\sigma^2 > 0$ and a finite absolute sth moment is assumed. In our example (7) the fixed point is then unique and, in fact, we can prove the contraction property in the Zolotarev metrics ζ_s.

Zolotarev [59] found the following metric ζ_s, which is ideal of order $s > 0$, defined for d-dimensional vectors by

$$\zeta_s(X,Y) = \sup_{f \in \mathcal{F}_s} \big|E\big(f(X) - f(Y)\big)\big|, \tag{8}$$

where for $s = m + \alpha$, $0 < \alpha \le 1$, $m \in \mathbb{N}_0$,

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

Convergence in ζ_s implies weak convergence and moreover, for some $c > 0$,

$$c\zeta_s(X,Y) \ge \begin{cases} E(\|X\|^s - \|Y\|^s), \\ \pi^{1+s}(\|X\|, \|Y\|), \end{cases}$$

where π is the Prohorov metric (see [60]). There are upper bounds for ζ_s in terms of difference pseudomoments

$$\kappa_s(X,Y) = \sup\big\{|Ef(X) - f(Y)| : \|f(x) - f(y)\| \le \big\|\|x\|^{s-1}x - \|y\|^{s-1}y\big\|\big\}.$$

Note that κ_s is the minimal metric of the compound metric

$$\tau_s(X,Y) = E\big\|\|X\|^{s-1}X - \|Y\|^{s-1}Y\big\|, \tag{9}$$

that is, $\kappa_s(\mu,\nu) = \inf\{\tau_s(X,Y) : X \sim \mu, Y \sim \nu\}$, and, therefore, allows estimates in terms of moments. We also use that κ_s and ℓ_s are topologically equivalent on spaces of random variables with uniformly bounded absolute sth moment. Finiteness of $\zeta_s(X,Y)$ implies that X and Y have identical mixed moments up to order m. Mixed moments are the expectations of products of powers of coordinates of a multivariate random variable. The order of a mixed moment is the sum of the exponents in such a product. For X and Y such that all mixed moments up to order m are zero and moments of order s are finite, we have that

$$\zeta_s(X,Y) \le \frac{\Gamma(1+\alpha)}{\Gamma(1+s)} \Lambda_s(X,Y), \tag{10}$$

where

$$\Lambda_s(X,Y) = 2m\kappa_s(X,Y) + (2\kappa_s(X,Y))^{\alpha}\big[\min(M_X^s, M_Y^s)\big]^{1-\alpha}. \tag{11}$$

Here M_X^s and M_Y^s are the absolute moments of X and Y of order s. From (11) we obtain, in particular, for $\alpha = 1$ (i.e., $s \in \mathbb{N}$),

$$\zeta_s(X,Y) \le (2m+s)\kappa_s(X,Y).$$

For $s \in \mathbb{N}$, finiteness of $\zeta_s(X,Y)$ does not need finiteness of sth absolute moments and so ζ_s also can be applied to stable distributions, for example.

For proving convergence to a normal distribution, ζ_s is well suited for $s > 2$. The reason for this is that the operator T in (7) that characterizes the normal distribution is a contraction on $\mathcal{M}_s^1(0,\sigma^2)$ with respect to ζ_s,

$$\begin{aligned} \zeta_s(TX, TY) &= \zeta_s\left(\frac{1}{\sqrt{2}}X_1 + \frac{1}{\sqrt{2}}X_2, \frac{1}{\sqrt{2}}Y_1 + \frac{1}{\sqrt{2}}Y_2\right) \\ &\le \zeta_s\left(\frac{1}{\sqrt{2}}X_1, \frac{1}{\sqrt{2}}Y_1\right) + \zeta_s\left(\frac{1}{\sqrt{2}}X_2, \frac{1}{\sqrt{2}}Y_2\right) \\ &\le \left(\left(\frac{1}{\sqrt{2}}\right)^s + \left(\frac{1}{\sqrt{2}}\right)^s\right)\zeta_s(X,Y), \end{aligned}$$

and we have $2(1/\sqrt{2})^s < 1$. Note, however, that by normalization typically only the first two moments can be matched and so the range of application is restricted to $s \le 3$. For linear transformations A, one obtains

$$\zeta_s(AX, AY) \le \|A\|_{\mathrm{op}}^s \zeta_s(X,Y), \tag{12}$$

where $\|A\|_{\mathrm{op}} := \sup_{\|x\|=1} \|Ax\|$ is the operator norm of A (see [60]). Some further properties are established throughout the article.

3. Contraction and fixed-point properties. In this section random affine transformations of multivariate measures are studied. Given a vector $(A_1, \dots, A_K, b)$ of random $d \times d$ matrices $A_1, \dots, A_K$ and a random d-dimensional vector b, we associate the transformation

$$T : \mathcal{M}^d \to \mathcal{M}^d, \qquad \mu \mapsto \mathcal{L}\left(\sum_{r=1}^{K} A_r Z^{(r)} + b\right). \tag{13}$$

Here $(A_1,\dots,A_K,b)$, $Z^{(1)},\dots,Z^{(K)}$ are assumed to be independent, $Z^{(r)}\sim\mu$ for $r=1,\dots,K$ and $\mathcal{M}^d$ denotes the space of d-dimensional probability measures. If $(A_1,\dots,A_K,b)$ has components with finite absolute sth moments and $\|\mu\|_s := (\mathbb{E}|Z|^s)^{(1/s)\vee 1} < \infty$, then, by independence, $\|T\mu\|_s<\infty$. In the following discussion, Lipschitz and contraction properties of T with respect to the Zolotarev metric ζ_s are crucial. To have $\zeta_s(\mu,\nu)<\infty$ we assume that μ and ν have finite absolute sth moments and that all mixed moments of μ,ν of orders less than s are equal. In this case the following Lipschitz property, which is an extension of (12), holds:

LEMMA 3.1. *Let* $(A_1,\dots,A_K,b)$ *and* T *be given as in* (13), *and let* $\mu,\nu\in\mathcal{M}^d$ *with* $\|\mu\|_s,\|\nu\|_s<\infty$ *and identical mixed moments of orders less than* s. *Let* $(A_1,\dots,A_K,b)$ *be* s-*integrable. Then we have*

$$\zeta_s(T\mu,T\nu)\le\left(\mathbb{E}\sum_{r=1}^K\|A_r\|_{\mathrm{op}}^s\right)\zeta_s(\mu,\nu). \tag{14}$$

PROOF. By independence we have $\|T\mu\|_s,\|T\nu\|_s<\infty$. For given $(A_1,\dots,A_K,b)$ the mixed moments of order less than s of $T\mu$ depend only on the mixed moments of μ of order less than s. Thus $T\mu$ and $T\nu$ have identical mixed moments of order less than s. This implies $\zeta_s(T\mu,T\nu)<\infty$. The s-homogeneity of ζ_s with respect to linear transformations given in (12) implies, with the notation $\Upsilon=\mathcal{L}(A_1,\dots,A_K,b)$ and $\alpha=(\alpha_1,\dots,\alpha_K)$,

$$\begin{aligned}
&\zeta_s(T\mu,T\nu)\\
&\quad=\zeta_s\left(\sum_{r=1}^K A_rZ^{(r)}+b,\sum_{r=1}^K A_rW^{(r)}+b\right)\\
&\quad=\sup_{f\in\mathcal{F}_s}\left\{\left|\mathbb{E}\left[f\left(\sum_{r=1}^K A_rZ^{(r)}+b\right)-f\left(\sum_{r=1}^K A_rW^{(r)}+b\right)\right]\right|\right\}\\
&\quad=\sup_{f\in\mathcal{F}_s}\left\{\left|\int\mathbb{E}\left[f\left(\sum_{r=1}^K \alpha_rZ^{(r)}+\beta\right)-f\left(\sum_{r=1}^K \alpha_rW^{(r)}+\beta\right)\right]d\Upsilon(\alpha,\beta)\right|\right\}\\
&\quad\le\int\sup_{f\in\mathcal{F}_s}\left\{\left|\mathbb{E}\left[f\left(\sum_{r=1}^K \alpha_rZ^{(r)}+\beta\right)-f\left(\sum_{r=1}^K \alpha_rW^{(r)}+\beta\right)\right]\right|\right\}d\Upsilon(\alpha,\beta)\\
&\quad=\int\zeta_s\left(\sum_{r=1}^K\alpha_rZ^{(r)}+\beta,\sum_{r=1}^K\alpha_rW^{(r)}+\beta\right)d\Upsilon(\alpha,\beta)\\
&\quad\le\int\sum_{r=1}^K\|\alpha_r\|_{\mathrm{op}}^s\zeta_s(\mu,\nu)\,d\Upsilon(\alpha,\beta)=\left(\mathbb{E}\sum_{r=1}^K\|A_r\|_{\mathrm{op}}^s\right)\zeta_s(\mu,\nu),
\end{aligned} \tag{15}$$

AAP ims v.2003/08/05 Prn:1/09/2003; 11:54 F:aap165.tex; (Aiste) p. 8

where $Z^{(1)}, \ldots, Z^{(K)}, W^{(1)}, \ldots, W^{(K)}, (A_1, \ldots, A_K, b)$ are independent with $Z^{(r)} \sim \mu$ and $W^{(r)} \sim \nu$. $\square$

REMARK. With respect to the ℓ_2 metric, the corresponding contraction property is (see [5, 43])

$$\ell_2(T\mu, T\nu) \leq \left\| \mathbb{E} \sum_{r=1}^{K} A_r^t A_r \right\|_{\mathrm{op}} \ell_2(\mu, \nu).$$

Since $\|\mathbb{E}\sum_{r=1}^K (A_r^t A_r)\|_{\mathrm{op}} \leq \mathbb{E}\sum_{r=1}^K \|A_r^t A_r\|_{\mathrm{op}} = \mathbb{E}\sum_{r=1}^K \|A_r\|_{\mathrm{op}}^2$, the contraction condition for ℓ_2 is weaker than that for the comparable ζ_2 and, therefore, it is preferable to apply ℓ_2 compared to ζ_2 if it applies. Note, however, that in the one-dimensional case both conditions are identical and only first moments have to be controlled for the application of ζ_2. In comparison to the ℓ_p metrics, the ζ_s metrics allow the exponent s in (14) to vary and thus are a much more flexible tool compared to the ℓ_p metrics.

From the point of view of applications, we can scale only the first and second mixed moments. Therefore, in particular, the cases $0 < s \leq 3$ are of interest. For $2 < s \leq 3$ we have to control the mean and the covariances to obtain the finiteness of the ζ_s metric. We define for $2 < s \leq 3$, a vector $m \in \mathbb{R}^d$, and for a symmetric positive semidefinite $d \times d$ matrix C, the space

$$\mathcal{M}_s^d(m, C) := \{\mu \in \mathcal{M}^d : \|\mu\|_s < \infty, \mathbb{E}\mu = m, \operatorname{Cov}(\mu) = C\}, \quad 2 < s \leq 3. \tag{16}$$

Then ζ_s is finite on $\mathcal{M}_s^d(m, C) \times \mathcal{M}_s^d(m, C)$ for all $2 < s \leq 3$, $m \in \mathbb{R}^d$ and symmetric positive semidefinite C. For the sake of short notation, we also write $\mathcal{M}_s^d(m, C)$ for $0 < s \leq 2$; this same term for $1 < s \leq 2$ denotes the subspace of probability distributions with finite sth moment and mean m (the C has no meaning). For $0 < s \leq 1$, the m and C have no meaning, and $\mathcal{M}_s^d(m, C)$ then denotes the space of probability distributions on $\mathbb{R}^d$ with finite sth moment; thus,

$$\mathcal{M}_s^d(m, C) := \{\mu \in \mathcal{M}^d : \|\mu\|_s < \infty, \mathbb{E}\mu = m\}, \qquad 1 < s \leq 2, \tag{17}$$

$$:= \{\mu \in \mathcal{M}^d : \|\mu\|_s < \infty\}, \qquad 0 < s \leq 1. \tag{18}$$

A direct calculation yields the ranges of the restriction of T to the set $\mathcal{M}_s^d(m, C)$:

LEMMA 3.2. *Let* $(A_1, \ldots, A_K, b)$ *and* T *be given as in* (13) *with* $(A_1, \ldots, A_K, b)$ *being* s*-integrable for some* $0 < s \leq 3$. *Then it holds that*

$$T\big(\mathcal{M}_s^d(m, C)\big) \subset \mathcal{M}_s^d(m_T, C_T)$$

AAP ims v.2003/08/05 Prn:1/09/2003; 11:54 F:aap165.tex; (Aiste) p. 9

with

$$m_T := \left(\mathbb{E}\sum_{r=1}^K A_r\right)m + \mathbb{E}b \tag{19}$$

and

$$\begin{aligned} C_T &:= E[bb^t] + \mathbb{E}\sum_{r=1}^K (A_r C A_r^t) \\ &\quad + \mathbb{E}\left[\left(\sum_{r=1}^K A_r\right)mb^t\right] + \mathbb{E}\left[\left(\sum_{r=1}^K A_r\right)mb^t\right]^t. \end{aligned} \tag{20}$$

We are interested in fixed points of the map T in certain subsets of $\mathcal{M}^d$. To this aim we consider, for a fixed $(A_1, \dots, A_k, b)$, some $m \in \mathbb{R}^d$ and some symmetric positive semidefinite $d \times d$ matrix C such that

$$m = m_T, \qquad C = C_T.$$

Then, by Lemma 3.2, T maps $\mathcal{M}_s^d(m, C)$ into $\mathcal{M}_s^d(m, C)$ for all $0 < s \le 3$. Moreover, by Lemma 3.1, T is a contraction on $(\mathcal{M}_s^d(m, C), \zeta_s)$ if

$$\mathbb{E}\sum_{r=1}^K \|A_r\|_{\mathrm{op}}^s < 1.$$

Next we prove the existence and uniqueness of a fixed point of T in $\mathcal{M}_s^d(m, C)$.

LEMMA 3.3. *Let* $(A_1, \dots, A_K, b)$ *and* T *be given as in* (13) *with* $(A_1, \dots, A_K, b)$ *being* s*-integrable for some* $0 < s \le 3$. *Let* $m \in \mathbb{R}^d$ *and let a symmetric positive semidefinite* $d \times d$ *matrix* C *be given such that* $m = m_T$ *and* $C = C_T$ [*defined in* (19) *and* (20)]. *If the contraction condition*

$$\xi := \mathbb{E}\sum_{r=1}^K \|A_r\|_{\mathrm{op}}^s < 1$$

is satisfied, *then the restriction of* T *to* $\mathcal{M}_s^d(m, C)$ *has a unique fixed point*.

PROOF. We choose $\mu_0 \in \mathcal{M}_s^d(m, C)$ and define $\mu_n := T(\mu_{n-1}) = T^{(n)}(\mu_0)$ for $n \ge 1$. Then for all $p \in \mathbb{N}$ we have, by Lemma 3.1,

$$\begin{aligned} \zeta_s(\mu_n, \mu_{n+p}) &\le \sum_{i=0}^{p-1} \zeta_s(\mu_{n+i}, \mu_{n+i+1}) \\ &\le \zeta_s(\mu_0, \mu_1) \sum_{i=0}^{p-1} \xi^{n+i} \\ &\le \zeta_s(\mu_0, \mu_1) \frac{\xi^n}{1-\xi} \to 0 \end{aligned}$$

as $n \to \infty$. Thus (μ_n) is a Cauchy sequence in the metric space $(\mathcal{M}_s^d(m, C), \zeta_s)$. By the lower estimate of ζ_s in the Lévy metric, we obtain that (μ_n) converges weakly to a $\mu \in \mathcal{M}^d$. The weak convergence, the independence properties in the definition of T and the continuity of the affine transformation imply $T\mu_n \to T\mu$ weakly; thus we have that $\mu = T\mu$ is a fixed point of T. For random variables X_n and X with $\mathcal{L}(X_n) = \mu_n$ and $\mathcal{L}(X) = \mu$, and with the estimate $|\mathbb{E}\|X_n\|^s - \mathbb{E}\|X_m\|^s| \le \mathrm{const}\,\zeta_s(\mu_n, \mu_m)$ [see (9)], we have that the real sequence $(\|\mu_n\|_s)$ is a Cauchy sequence and, therefore, bounded. This implies the uniform integrability of $\{|X_n|^{\tilde{s}} : n \in \mathbb{N}\}$ for all $0 < \tilde{s} < s$, which together with the convergence in distribution implies $X_n \to X$ in $L_{\tilde{s}}$. Since $\mathbb{E}X_n = m$ for all $n \in \mathbb{N}$ in the case $s > 1$ and additionally $\mathrm{Cov}(X_n) = C$ for all $n \in \mathbb{N}$ in the case $s > 2$, this implies $\mathbb{E}X = m$ and $\mathrm{Cov}(X) = C$ in these cases, respectively. By the lemma of Fatou, $\mathbb{E}\|X\|^s < \infty$; thus $\mathcal{L}(X) \in \mathcal{M}_s^d(m, C)$. For the uniqueness, let $\mu, \nu \in \mathcal{M}_s^d(m, C)$ be fixed points of T. Then

$$\zeta_s(T\mu, T\nu) \le \left(\mathbb{E}\sum_{r=1}^K \|A_r\|_{\mathrm{op}}^s\right)\zeta_s(\mu, \nu);$$

thus $\zeta_s(\mu, \nu) = 0$ and $\mu = \nu$. $\square$

REMARK. Svante Janson pointed out to us that the metric spaces $(\mathcal{M}_s^d(m, C), \zeta_s)$ are complete, which, by Banach's fixed-point theorem, implies the assertion of Lemma 3.3 as well.

The previous considerations yield:

COROLLARY 3.4. *Let* $(A_1, \dots, A_K, b)$ *and* T *be given as in* (13) *with* $(A_1, \dots, A_K, b)$ *being* s*-integrable*, $0 < s \le 3$ *and* $\mathbb{E}\sum_{r=1}^K \|A_r\|_{\mathrm{op}}^s < 1$. *Assume*

$$\mathbb{E}b = \begin{cases} 0 \quad \textit{and} \quad \mathbb{E}[bb^t] + \mathbb{E}\sum_{r=1}^K (A_r A_r^t) = \mathrm{Id}_d, & \textit{if } 2 < s \le 3, \\ 0, \qquad \textit{if } 1 < s \le 2. \end{cases}$$

Then T *has a unique fixed point in* $\mathcal{M}_s^d(0, \mathrm{Id}_d)$.

4. The main convergence theorem.

4.1. *Convergence in the Zolotarev metrics.* We return to the situation outlined in the Introduction. Given is a sequence (Y_n) of random vectors satisfying the recurrence

$$Y_n \stackrel{\mathcal{D}}{=} \sum_{r=1}^K A_r(n) Y_{I_r^{(n)}}^{(r)} + b_n, \qquad n \ge n_0, \tag{21}$$

AAP ims v.2003/08/05 Prn:1/09/2003; 11:54 F:aap165.tex; (Aiste) p. 11

with relationships as in (1). In the case $2 < s \le 3$, we assume that $\mathrm{Cov}(Y_n)$ is positive definite for $n \ge n_1 \ge n_0$. The fact that for the application of ζ_s we have to control mixed moments up to order less than s allows the following flexibility in the normalization: We set

$$X_n := C_n^{-1/2}(Y_n - M_n), \qquad n \ge 0, \tag{22}$$

where

$$M_n := \begin{cases} \mathbb{E}Y_n, & C_n := \begin{cases} \mathrm{Id}_d & \text{for } 0 \le n < n_1, \\ \mathrm{Cov}(Y_n) & \text{for } n \ge n_1, \end{cases} \qquad \text{if } 2 < s \le 3, \\ \mathbb{E}Y_n, & C_n \text{ is positive definite} \qquad \text{if } 1 < s \le 2, \end{cases}$$

$$C_n \text{ is positive definite} \qquad \text{if } 0 < s \le 1.$$

The normalized quantities (X_n) then satisfy the modified recurrence

$$X_n \stackrel{\mathcal{D}}{=} \sum_{r=1}^{K} A_r^{(n)} X^{(r)}_{I_r^{(n)}} + b^{(n)}, \qquad n \ge n_1, \tag{23}$$

with $A_r^{(n)}$ and $b^{(n)}$ as in (4). The following theorem gives a transfer of the type that convergence of the coefficients in (23) yields under appropriate conditions, that is, convergence of the X_n to the fixed point of the associated limiting equation (5).

THEOREM 4.1. *Let (X_n) be given as in* (22) *and be s-integrable*, $0 < s \le 3$. *We assume that*

$$\bigl(A_1^{(n)}, \dots, A_K^{(n)}, b^{(n)}\bigr) \stackrel{L_s}{\to} (A_1^*, \dots, A_k^*, b^*), \tag{24}$$

$$\mathbb{E} \sum_{r=1}^{K} \|A_r^*\|_{\mathrm{op}}^s < 1, \tag{25}$$

$$\mathbb{E}\Bigl[\mathbf{1}_{\{I_r^{(n)} \le l\} \cup \{I_r^{(n)} = n\}} \|A_r^{(n)}\|_{\mathrm{op}}^s\Bigr] \to 0 \tag{26}$$

for all $l \in \mathbb{N}$ and $r = 1, \dots, K$. Then (X_n) converges to a limit X,

$$\zeta_s(X_n, X) \to 0, \qquad n \to \infty,$$

where $\mathcal{L}(X) \in \mathcal{M}_s^d(0, \mathrm{Id}_d)$ is given as the unique fixed point in $\mathcal{M}_s^d(0, \mathrm{Id}_d)$ of the equation

$$X \stackrel{\mathcal{D}}{=} \sum_{r=1}^{K} A_r^* X^{(r)} + b^*. \tag{27}$$

Here $(A_1^*, \dots, A_k^*, b^*), X^{(1)}, \dots, X^{(K)}$ are independent and $X^{(r)} \sim X$ for $r = 1, \dots, K$.

PROOF. For $2 < s \leq 3$, the sequence (X_n) in (23) is standardized; thus we obtain the relationships

$$\mathbb{E}b^{(n)} = 0, \qquad \mathbb{E}b^{(n)}(b^{(n)})^t + \mathbb{E}\sum_{r=1}^{K}\big(A_r^{(n)}(A_r^{(n)})^t\big) = \mathrm{Id}_d, \qquad n \geq n_1. \tag{28}$$

For $1 < s \leq 2$, the sequence (X_n) is centered; thus $\mathbb{E}b^{(n)} = 0$. Thus from the convergence in L_s in (24) we obtain

$$\mathbb{E}b^* = 0, \qquad \mathbb{E}[b^*(b^*)^t] + \mathbb{E}\sum_{r=1}^{K}\big(A_r^*(A_r^*)^t\big) = \mathrm{Id}_d$$

in the case $2 < s \leq 3$ and $\mathbb{E}b^* = 0$ for $1 < s \leq 2$. Therefore, by Corollary 3.4, there exists a unique fixed point $\mathcal{L}(X) \in \mathcal{M}_s^d(0, \mathrm{Id}_d)$ of (27) in $\mathcal{M}_s^d(0, \mathrm{Id}_d)$ for all $0 < s \leq 3$.

We introduce the accompanying sequence

$$Q_n :\stackrel{\mathcal{D}}{=} \sum_{r=1}^{K} A_r^{(n)}\Big(\mathbf{1}_{\{I_r^{(n)} < n_1\}} X^{(r)}_{I_r^{(n)}} + \mathbf{1}_{\{I_r^{(n)} \geq n_1\}} X^{(r)}\Big) + b^{(n)}, \qquad n \geq n_1,$$

where $(A_1^{(n)}, \dots, A_K^{(n)}, b^{(n)}, I^{(n)}), X^{(1)}, \dots, X^{(K)}, (X_n^{(1)}), \dots, (X_n^{(K)})$ are independent with $X^{(r)} \sim X$ and $X_j^{(r)} \sim X_j$ for $r = 1, \dots, K$, $j = 0, \dots, n_1 - 1$. Since $\mathcal{L}(X) \in \mathcal{M}_s^d(0, \mathrm{Id}_d)$ and comparing the definition of Q_n for $2 < s \leq 3$ with X_n in (23), we deduce $\mathrm{Cov}(Q_n) = \mathrm{Cov}(X_n)$; hence $\mathcal{L}(Q_n) \in \mathcal{M}_s^d(0, \mathrm{Id}_d)$ for all $n \geq n_1$ and thus ζ_s distances between X_n, Q_n and X are finite for $n \geq n_1$. We obtain from the triangle inequality

$$\zeta_s(X_n, X) \leq \zeta_s(X_n, Q_n) + \zeta_s(Q_n, X). \tag{29}$$

First we show that $\zeta_s(Q_n, X) \to 0$. This is a consequence of the upper bound in (10) and of the convergence of the pseudomoments $\kappa_s(Q_n, X) \to 0$ which follows from $\ell_s(Q_n, X) \to 0$, since absolute moments of order s are bounded for (Q_n): With the representation $X \stackrel{\mathcal{D}}{=} \sum_{r=1}^K A_r^* X^{(r)} + b^*$, we obtain

$$\begin{aligned}
\ell_s(Q_n, X) &\leq \Bigg\|\sum_{r=1}^{K}\Big(A_r^* - \mathbf{1}_{\{I_r^{(n)} \geq n_1\}} A_r^{(n)}\Big) X^{(r)}\Bigg\|_s + \|b^{(n)} - b^*\|_s \\
&\quad + \Bigg\|\sum_{r=1}^{K} \mathbf{1}_{\{I_r^{(n)} < n_1\}} A_r^{(n)} X^{(r)}_{I_r^{(n)}}\Bigg\|_s \\
&\leq \sum_{r=1}^{K}\Big(\|A_r^* - A_r^{(n)}\|_s + \Big\|\mathbf{1}_{\{I_r^{(n)} < n_1\}} \|A_r^{(n)}\|_{\mathrm{op}}\Big\|_s\Big)\|X\|_s + \|b^{(n)} - b^*\|_s \qquad (30) \\
&\quad + \sum_{r=1}^{K}\Bigg\|\sum_{j=0}^{n_1-1} \mathbf{1}_{\{I_r^{(n)} = j\}} A_r^{(n)} X_j^{(r)}\Bigg\|_s. \qquad (31)
\end{aligned}$$

The three summands in (30) converge to zero by (24) and (26). The summand in (31) tends to zero using (26) by

$$\begin{aligned}
&\Big\| \mathbf{1}_{\{I_r^{(n)}=j\}} A_r^{(n)} X_j^{(r)} \Big\|_s \\
&\qquad \le \Big\| \mathbf{1}_{\{I_r^{(n)}=j\}} \|A_r^{(n)}\|_{\mathrm{op}} \|X_j^{(r)}\| \Big\|_s \\
&\qquad \le \Big\| \mathbf{1}_{\{I_r^{(n)}<n_1\}} \|A_r^{(n)}\|_{\mathrm{op}} \Big\|_s \sup_{0\le j<n_1} \|X_j\|_s .
\end{aligned}$$

The first summand in (29) is estimated similarly to (15). Let Υ_n denote the joint distribution of $(A_1^{(n)},\dots,A_K^{(n)},b^{(n)},I^{(n)})$ and $\alpha=(\alpha_1,\dots,\alpha_K)$, $j=(j_1,\dots,j_K)$. Then with

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

we obtain, for $n \ge n_1$,

$$\begin{aligned}
&\zeta_s(X_n, Q_n) \\
&\quad = \zeta_s\Bigg(\sum_{r=1}^K A_r^{(n)} X_{I_r^{(n)}}^{(r)} + b^{(n)}, \\
&\qquad\qquad \sum_{r=1}^K A_r^{(n)} \Big(\mathbf{1}_{\{I_r^{(n)}<n_1\}} X_{I_r^{(n)}}^{(r)} + \mathbf{1}_{\{I_r^{(n)}\ge n_1\}} X^{(r)} \Big) + b^{(n)} \Bigg) \\
&\quad \le \int \zeta_s\Bigg(\sum_{r=1}^K \alpha_r X_{j_r}^{(r)}, \sum_{r=1}^K \alpha_r \big(\mathbf{1}_{\{j_r<n_1\}} X_{j_r}^{(r)} + \mathbf{1}_{\{j_r\ge n_1\}} X^{(r)} \big) \Bigg)\, d\Upsilon_n(\alpha,\beta,j) \\
&\quad \le \int \sum_{r=1}^K \mathbf{1}_{\{j_r\ge n_1\}} \|\alpha_r\|_{\mathrm{op}}^s \zeta_s(X_{j_r}, X)\, d\Upsilon_n(\alpha,\beta,j) \\
&\quad \le p_n \zeta_s(X_n, X) + \Bigg(\mathbb{E}\sum_{r=1}^K \|A_r^{(n)}\|_{\mathrm{op}}^s \Bigg) \sup_{n_1\le j\le n-1} \zeta_s(X_j, X).
\end{aligned} \tag{32}$$

Thus with (29) it follows that

$$\zeta_s(X_n, X) \le \frac{1}{1-p_n}\Bigg[\Bigg(\mathbb{E}\sum_{r=1}^K \|A_r^{(n)}\|_{\mathrm{op}}^s \Bigg) \sup_{n_1\le j\le n-1} \zeta_s(X_j, X) + o(1) \Bigg].$$

This implies that $(\zeta_s(X_n, X))$ is bounded. Let $\bar\eta := \sup_{n\ge n_1} \zeta_s(X_n, X)$ and $\eta := \limsup_{n\to\infty} \zeta_s(X_n, X)$, and let $\varepsilon > 0$ be arbitrary. There exists an $l \in \mathbb{N}$ with

AAP ims v.2003/08/05 Prn:1/09/2003; 11:54 F:aap165.tex; (Aiste) p. 14

$\zeta_s(X_n, X) \le \eta + \varepsilon$ for all $n \ge l$. We deduce, using (33), (29) and (24),

$$\begin{aligned}
\zeta_s(X_n, X) &\le \frac{1}{1-p_n}\biggl[\int \sum_{r=1}^{K} \mathbf{1}_{\{n_1 \le j_r \le l\}} \|\alpha_r\|_{\mathrm{op}}^s \zeta_s(X_{j_r}, X)\, d\Upsilon_n(\alpha, \beta, j) \\
&\qquad\qquad + \int \sum_{r=1}^{K} \mathbf{1}_{\{l < j_r < n\}} \|\alpha_r\|_{\mathrm{op}}^s \zeta_s(X_{j_r}, X)\, d\Upsilon_n(\alpha, \beta, j)\biggr] \\
&\le \frac{\bar{\eta}}{1-p_n}\mathbb{E}\sum_{r=1}^{K}\Bigl(\mathbf{1}_{\{n_1 \le I_r^{(n)} \le l\}} \|A_r^{(n)}\|_{\mathrm{op}}^s\Bigr) + \frac{\eta+\varepsilon}{1-p_n}\mathbb{E}\sum_{r=1}^{K}\|A_r^{(n)}\|_{\mathrm{op}}^s \\
&\le \Biggl(\mathbb{E}\sum_{r=1}^{K}\|A_r\|_{\mathrm{op}}^s\Biggr)(\eta+\varepsilon) + o(1).
\end{aligned}$$

With $n \to \infty$ we obtain

$$\eta \le \Biggl(\mathbb{E}\sum_{r=1}^{K}\|A_r\|_{\mathrm{op}}^s\Biggr)(\eta+\varepsilon).$$

Since $\varepsilon > 0$ is arbitrary and $\mathbb{E}\sum_{r=1}^{K}\|A_r\|_{\mathrm{op}}^s < 1$, we obtain $\eta = 0$. Hence $\zeta_s(X_n, X) \to 0$. $\square$

4.2. *Other metrics.* Theorem 4.1 yields convergence of X_n to X w.r.t. the ζ_s metric, where X is the unique fixed point of (27) in $\mathcal{M}_s^d(0, \mathrm{Id}_d)$ under the contraction condition $\mathbb{E}\sum_{r=1}^{K}\|A_r^*\|_{\mathrm{op}}^s < 1$. It is of interest that several related convergence results are obtainable for further metrics by similar arguments or by upper bounds for these metrics in terms of the Zolotarev metric ζ_s. We show that this, in particular, leads to local and global approximation results.

For random vectors X and Y with densities f_X and f_Y, let

$$\ell(X, Y) = \operatorname*{ess\,sup}_{x \in \mathbb{R}^d} |f_X(x) - f_Y(x)|$$

denote the sup distance of the densities. Let θ be a random vector with a smooth density f_θ. We say that f_θ satisfies the Hölder condition (H_r), $r = m + \alpha$, $0 < \alpha \le 1$, if

$$(H_r) \qquad \bigl\| f_\theta^{(m)}(x) - f_\theta^{(m)}(y)\bigr\| \le C_r(\theta)\|x - y\|^\alpha. \tag{33}$$

By means of θ we introduce a smoothed version $\tilde{\ell}_r$ of the distance ℓ. Define $\tilde{\ell}_r = \tilde{\ell}_{r,\theta}$ by

$$\tilde{\ell}_r(X, Y) = \sup_{h \in \mathbb{R}} |h|^r \ell(X + h\theta, Y + h\theta).$$

Smoothing metrics have been used for proving central limit theorems for normalized sums and for martingales in probability theory (cf. [47, 48, 55]).

PROPOSITION 4.2 (Regularity of $\tilde{\ell}_r$). *Let $r > d$ and θ satisfy condition (H_{r-d}). Then $\tilde{\ell}_r$ is a probability metric ideal of order $r-d$ and*

$$\tilde{\ell}_r(X,Y) \le C_{r-d}(\theta)\zeta_{r-d}(X,Y), \tag{34}$$

where $C_r(\theta)$ is the Hölder constant in (33).

PROOF. The probability metric property and property (6) are easy to establish. To see that $\tilde{\ell}_r$ is ideal of order $r-d$, let $c \neq 0$. Then

$$\begin{aligned}\tilde{\ell}_r(cX,cY) &= \sup_{h\in\mathbb{R}} |h|^r \operatorname*{ess\,sup}_{x\in\mathbb{R}^d} \left| f_{cX+h\theta}(x) - f_{cY+h\theta}(x)\right| \\ &= \sup_{h\in\mathbb{R}} |h|^r \operatorname*{ess\,sup}_{x\in\mathbb{R}^d} \left| f_{c(X+(h/c)\theta)}(x) - f_{c(Y+(h/c)\theta)}(x)\right| \\ &= \frac{1}{|c|^d} \sup_{h\in\mathbb{R}} |h|^r \operatorname*{ess\,sup}_{x\in\mathbb{R}^d} \left| f_{X+(h/c)\theta}\left(\frac{x}{c}\right) - f_{Y+(h/c)\theta}\left(\frac{x}{c}\right)\right| \\ &= \frac{1}{|c|^d} \sup_{h\in\mathbb{R}} |h|^r \operatorname*{ess\,sup}_{x\in\mathbb{R}^d} \left| f_{X+(h/c)\theta}(x) - f_{Y+(h/c)\theta}(x)\right| \\ &= |c|^{r-d}\tilde{\ell}_r(X,Y).\end{aligned}$$

To prove (34) note that

$$\begin{aligned}\tilde{\ell}_r(X,Y) &= \sup_{h\in\mathbb{R}} |h|^r \operatorname*{ess\,sup}_{x\in\mathbb{R}^d} \left| f_{h(X/h+\theta)}(x) - f_{h(Y/h+\theta)}(x)\right| \\ &= \sup_{h\in\mathbb{R}} |h|^{r-d} \operatorname*{ess\,sup}_{x\in\mathbb{R}^d} \left| f_{X/h+\theta}\left(\frac{x}{h}\right) - f_{Y/h+\theta}\left(\frac{x}{h}\right)\right| \\ &= \sup_{h\in\mathbb{R}} |h|^{r-d} \ell\left(\frac{X}{h}+\theta, \frac{Y}{h}+\theta\right).\end{aligned}$$

If θ satisfies condition (H_{r-d}), then we obtain with $H = \mathbb{P}_X - \mathbb{P}_Y$, where $\mathbb{P}_X$ denotes the distribution of X,

$$\begin{aligned}\ell(X+\theta, Y+\theta) &= \operatorname*{ess\,sup}_{x\in\mathbb{R}^d} \left| \int f_\theta(x-y)\,dH(y)\right| \\ &\le C_{r-d}(\theta)\zeta_{r-d}(X,Y),\end{aligned}$$

since for any x, $f_\theta(x-\cdot)$ is a Hölder function of order $r-d$ with Hölder constant $C_{r-d}(\theta)$. Therefore,

$$\begin{aligned}|h|^r\ell(X+h\theta, Y+h\theta) &= |h|^{r-d}\ell\left(\frac{X}{h}+\theta, \frac{Y}{h}+\theta\right) \\ &\le |h|^{r-d}C_{r-d}(\theta)\zeta_{r-d}\left(\frac{X}{h}, \frac{Y}{h}\right) \\ &= C_{r-d}(\theta)\zeta_{r-d}(X,Y).\end{aligned}$$

This implies that $\tilde{\ell}_r(X,Y) \le C_{r-d}(\theta)\zeta_{r-d}(X,Y)$. $\square$

REMARK. The smoothed metric $\tilde{\ell}_r = \tilde{\ell}_{r,\theta}$ was introduced in [47], 14.2.12, and used to prove central limit theorems. The statement of the regularity properties and bounds there is correct, however, only for $d = 1$ and is corrected here.

As a consequence, we obtain the following local convergence result.

COROLLARY 4.3 (Local convergence theorem). *Let θ satisfy the Hölder condition (H_s). Then under the conditions of Theorem* 4.1 *we obtain the local convergence result*

$$\alpha_n := \tilde{\ell}_{s+d}(X_n, X) \to 0.$$

In particular, for any sequence h_n such that $\alpha_n / h_n^{s+d} \to 0$, we obtain that

$$\ell(X_n + h_n\theta, Y + h_n\theta) \le \frac{\alpha_n}{h_n^{s+d}} \to 0.$$

If θ has a bounded support and $h_n \to 0$, then we obtain a local density convergence result with smoothing over a shrinking neighborhood. If, moreover, Y has a continuous density, we have $\ell(Y + h_n\theta, Y) = O(h_n)$. Then with the triangle inequality, we obtain $\ell(X_n + h_n\theta, Y) \to 0$, which is uniform convergence of the density of $X_n + h_n\theta$ to the density of Y for an appropriate sequence (h_r).

As a second example we consider the global smoothed total variation metric. Let

$$\sigma(X,Y) = 2\sup_A |\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)|$$

denote the total variation metric and let

$$\tilde{\sigma}_r(X,Y) = \tilde{\sigma}_{r,\theta}(X,Y) = \sup_{h\in\mathbb{R}} |h|^r \sigma(X + h\theta, Y + h\theta)$$

denote the smoothed total variation metric (see [47], page 267).

PROPOSITION 4.4 (Regularity of $\tilde{\sigma}_r$). *Assume that θ satisfies the Hölder condition (H_r). Then $\tilde{\sigma}_r$ is a probability metric, ideal of order r, and*

$$\tilde{\sigma}_r(X,Y) \le C_r(\theta)\zeta_r(X,Y).$$

PROOF. The proof is similar to that of Proposition 4.2. Note that σ is regular of order zero, that is,

$$\sigma(cX, cY) = \sigma(X,Y).$$

AAP ims v.2003/08/05 Prn:1/09/2003; 11:54 F:aap165.tex; (Aiste) p. 17

Therefore, $\sigma(X+h\theta, Y+h\theta) = \sigma(X/h+\theta, Y/h+\theta)$. Furthermore,

$$\begin{aligned}\sigma(X+\theta, Y+\theta) &= \sup_{|f|\le 1}\left|\int \big(f(X+\theta) - f(Y+\theta)\big)\,dP\right|\\ &= \sup_{|f|\le 1}\left|\int \bar{f}_\theta(x)\,dH(x)\right|,\end{aligned}$$

where $H = \mathbb{P}_X - \mathbb{P}_Y$ and $\bar{f}_\theta(x) = \int f_\theta(y-x) f(y)\,dy$. Since f_θ satisfies the Hölder condition of order r, $\bar{f}_\theta$ also satisfies the Hölder condition of order r and Hölder constant $C_r(\theta)$. Therefore, $\sigma(X+\theta, Y+\theta) \le C_r(\theta)\zeta_r(X, Y)$. Thus

$$\begin{aligned}\tilde{\sigma}_r(X, Y) &= \sup_{h\in\mathbb{R}} |h|^r \sigma(X+h\theta, Y+h\theta)\\ &\le \sup_{h\in\mathbb{R}} |h|^r \zeta_r\left(\frac{X}{h}, \frac{Y}{h}\right) C_r(\theta)\\ &= C_r(\theta)\zeta_r(X, Y).\end{aligned}$$

This yields the assertion. $\square$

As a consequence, we therefore obtain the following global convergence result.

COROLLARY 4.5 (Global convergence). *Let θ satisfy the Hölder condition (H_s). Then under the conditions of Theorem* 4.1, *we obtain the global convergence result*

$$\alpha_n = \tilde{\sigma}_s(X_n, X) \to 0.$$

In particular, for any sequence h_n such that $\alpha_n/h_n^s \to 0$, we obtain

$$\sigma(X_n + h_n\theta, Y + h_n\theta) \le \frac{\alpha_n}{h_n^s} \to 0.$$

Similar convergence results also hold true for further metrics like the smoothed ℓ_1 metric. For dimension $d = 1$,

$$\begin{aligned}(\tilde{\ell_1})_s(X, Y) &= \sup_{h\in\mathbb{R}} |h|^s \ell_1(X+h\theta, Y+h\theta)\\ &= \sup_{h\in\mathbb{R}} |h|^s \int |F_{X+h\theta}(x) - F_{Y+h\theta}(x)|\,dx,\end{aligned}$$

where F_X denotes the distribution function of X. The corresponding local result also holds true. It concerns the smoothed Kolmogorov metric, $d = 1$:

$$\begin{aligned}\tilde{\rho}_s(X, Y) &= \sup_{h\in\mathbb{R}} |h|^s \rho(X+h\theta, Y+h\theta)\\ &= \sup_{h\in\mathbb{R}} |h|^s \sup_{x\in\mathbb{R}} |F_{X+h\theta}(x) - F_{Y+h\theta}(x)|.\end{aligned}$$

4.3. *Random K depending on n.* The arguments presented to develop Theorem 4.1 can be extended to cover recurrences, where the number of copies K of the costs Y_n and X_n in recurrences (1) and (3), respectively, may be random and depend on n. However, we assume subsequently that K_n tends to a proper random variate, whereas in Section 5.2, we handle cases where $K_n \to \infty$ almost surely. We assume that $(Y_n)_{n\in\mathbb{N}_0}$ satisfies

$$Y_n \stackrel{\mathcal{D}}{=} \sum_{r=1}^{K_n} A_r(n) Y^{(r)}_{I_r^{(n)}} + b_n, \qquad n \geq n_0,$$

where $((A_r(n))_{r\in\mathbb{N}}, b_n, I^{(n)}, K_n), (Y_n^{(1)}), (Y_n^{(2)}), \dots$ are independent, $A_1(n)$, $A_2(n), \dots$ are random $d \times d$ matrices, b_n is a random d-dimensional vector, $(I_r^{(n)})$ are random cardinalities with $I_r^{(n)} \in \{0, \dots, n\}$, K_n is a positive integer-valued random variable and $(Y_n^{(1)}), (Y_n^{(2)}), \dots$ are identically distributed as (Y_n). We scale

$$X_n := C_n^{-1/2}(Y_n - M_n), \qquad n \geq 0, \tag{35}$$

as in (22), where in the case $2 < s \leq 3$ we assume that $\mathrm{Cov}(Y_n)$ is positive definite of all $n \geq n_1 \geq n_0$. Then we have

$$X_n \stackrel{\mathcal{D}}{=} \sum_{r=1}^{K_n} A_r^{(n)} X^{(r)}_{I_r^{(n)}} + b^{(n)}, \qquad n \geq n_1, \tag{36}$$

with $A_r^{(n)}$ and $b^{(n)}$ as in (4), and with K in the definition of $b^{(n)}$ replaced by K_n. For the formulation of a corresponding limit equation, let $((A_r^*)_{r\in\mathbb{N}}, b^*, K)$ be a tuple of random $d \times d$ matrices A_r^*, a random d-dimensional vector b^* and a random variate K in the positive integers such that

$$\|b^*\|_s < \infty, \qquad \left\| \sum_{r=1}^{K} \|A_r^*\|_{\mathrm{op}} \right\|_s < \infty. \tag{37}$$

In the conditions of the subsequent theorem we use the convention

$$A_r^{(n)}(\omega) = 0 \qquad \text{for } r > K_n(\omega), \qquad A_r^*(\omega) = 0 \qquad \text{for } r > K(\omega)$$

for all ω of the underlying probability space. Then we have:

THEOREM 4.6. *Let (X_n) be given as in* (35) *and be s-integrable*, $0 < s \leq 3$. *We assume the conditions* (37) *and*

$$\left\| \sum_{r=1}^{K_n \vee K} \|A_r^{(n)} - A_r^*\|_{\mathrm{op}} \right\|_s \to 0, \qquad n \to \infty,$$

$$\mathbb{E} \sum_{r=1}^{K} \|A_r^*\|_{\mathrm{op}}^s < 1, \qquad \left\| \sum_{r=1}^{K_n} \mathbf{1}_{\{I_r^{(n)} \leq l\} \cup \{I_r^{(n)} = n\}} \|A_r^{(n)}\|_{\mathrm{op}} \right\|_s \to 0$$

AAP ims v.2003/08/05 Prn:1/09/2003; 11:54 F:aap165.tex; (Aiste) p. 19

for all $l \in \mathbb{N}$. Then (X_n) converges to a limit X,

$$\zeta_s(X_n, X) \to 0, \qquad n \to \infty,$$

where $\mathcal{L}(X) \in \mathcal{M}_s^d(0, \mathrm{Id}_d)$ is given as the unique fixed point in $\mathcal{M}_s^d(0, \mathrm{Id}_d)$ of the equation

$$X \stackrel{\mathcal{D}}{=} \sum_{r=1}^{K} A_r^* X^{(r)} + b^*.$$

Here $(A_1^*, A_2^*, \dots, b^*, K), X^{(1)}, X^{(2)}, \dots$ are independent and $X^{(r)} \sim X$ for $r = 1, 2, \dots$.

While we are mainly concerned with the analysis of combinatorial structures data structures and recursive algorithms, where we typically have fixed K or $K_n \to \infty$ almost surely, the theorem may prove useful for the analysis of branching processes such as Galton–Watson trees. In particular, the theorem covers a proof of Yaglom's exponential limit law as given for an application of the contraction method in the ℓ_2 setting by Geiger [17]. The exponential limit distribution is characterized as the fixed point of $X \stackrel{\mathcal{D}}{=} U(X + X^*)$, with X, X^*, U independent and U unif[0, 1] distributed. The K_n is the number of children of the most recent common ancestor of the population at generation n.

5. Applications: central limit laws. In this section we link expansions of moments to our transfer theorem, Theorem 4.1, in a general setup. This leads to both, asymptotic normality and nonnormal cases. The normal distribution appears as the fixed point of the maps T given in (13) with

$$b = 0, \qquad \sum_{r=1}^{K} A_r A_r^t = \mathrm{Id}_d \tag{38}$$

almost surely. It is easy to check by characteristic functions that $\mathcal{N}(0, \mathrm{Id}_d)$ is then a fixed point of T and by Corollary 3.4 that this solution is unique in the space $M_s^d(0, \mathrm{Id}_d)$ for $s > 2$ if $\mathbb{E} \sum_{r=1}^{K} \|A_r\|_{\mathrm{op}}^s < 1$. In all the central limit laws in Section 5.3, the normal distribution as a limit distribution comes up as the fixed point of a transformation T satisfying (38).

We first focus on the univariate case $d = 1$ and link the expansion of the moments to our transfer theorem, Theorem 4.1, in a general setup that is not restricted to normal limit distributions. This leads first to a general transfer theorem that is capable of rederiving results involving nonnormal limit laws that also could be proven via the usage of ℓ_r metrics. We give some applications from the analysis of algorithms to illustrate our general theorem. Second, we give a convenient specialization to asymptotic normality, which covers many examples of central limit laws in the field of combinatorial structures. This is of particular importance because the competitive ℓ_r metric approach does not lead to such a transfer theorem. In the last part we discuss the multivariate case together with examples, where the number K of copies of the parameter on the right side may be random, depend on n and satisfy $K_n \to \infty$ almost surely.

5.1. *Univariate central limit laws.* A univariate situation quite common in combinatorial structures encompasses recursions of the type

$$Y_n \stackrel{\mathcal{D}}{=} \sum_{r=1}^{K} Y^{(r)}_{I_r^{(n)}} + b_n, \qquad n \geq n_0, \tag{39}$$

where $(Y_n^{(1)}), \dots, (Y_n^{(K)}), (I^{(n)}, b_n)$ are independent, $Y_j^{(r)} \sim Y_j$, $j \geq 0$, $\mathbb{P}(I_r^{(n)} = n) \to 0$ for $n \to \infty$ and $r = 1, \dots, K$ and such that $\mathrm{Var}(Y_n) > 0$ for $n \geq n_1$. Assume that for functions $f, g : \mathbb{N}_0 \to \mathbb{R}_0^+$ with $g(n) > 0$ for n sufficiently large we have the *stabilization condition* in L_s,

$$\begin{gathered} \left(\frac{g(I_r^{(n)})}{g(n)}\right)^{1/2} \to A_r^*, \qquad r = 1, \dots, K, \\ \frac{1}{g^{1/2}(n)}\left(b_n - f(n) + \sum_{r=1}^{K} f(I_r^{(n)})\right) \to b^*, \end{gathered} \tag{40}$$

and the *contraction condition*

$$\mathbb{E} \sum_{r=1}^{K} (A_r^*)^s < 1. \tag{41}$$

THEOREM 5.1 (Univariate transfer theorem). *Let (Y_n) be s-integrable and satisfy the recursive equation* (39), *and let f and g be given with stabilization and contraction conditions* (40) *and* (41) *for some $0 < s \leq 3$. Assume*

$$\mathbb{E}Y_n = \begin{cases} f(n) + o(g^{1/2}(n)), & \mathrm{Var}(Y_n) = g(n) + o(g(n)), \qquad \text{if } 2 < s \leq 3, \\ f(n) + o(g^{1/2}(n)), & \qquad\qquad\qquad\qquad\qquad\qquad\quad \text{if } 1 < s \leq 2. \end{cases}$$

Then

$$\frac{Y_n - f(n)}{g^{1/2}(n)} \xrightarrow{\mathcal{L}} X,$$

where X is the unique fixed point of

$$X \stackrel{d}{=} \sum_{r=1}^{K} A_r^* X^{(r)} + b^* \tag{42}$$

in $\mathcal{M}_s^1(0,1)$, where $(A_1^*, \dots, A_K^*, b^*), X^{(1)}, \dots, X^{(K)}$ are independent with $X^{(r)} \sim X$.

PROOF. We denote $M_n := \mathbb{E}Y_n$ for $1 < s \le 3$, $M_n := f(n)$ for $0 < s \le 1$, $\sigma_n := \sqrt{\mathrm{Var}(Y_n)}$ for $2 < s \le 3$ and $\sigma_n := \sqrt{g(n)}$ for $0 < s \le 2$. First note that $(Y_n - f(n))/g^{1/2}(n) \stackrel{\mathcal{L}}{\to} X$ for some X if and only if $X_n := (Y_n - M_n)/\sigma_n \stackrel{\mathcal{L}}{\to} X$. We have

$$X_n \stackrel{\mathcal{D}}{=} \sum_{r=1}^{K} \frac{\sigma_{I_r^{(n)}}}{\sigma_n} X_{I_r^{(n)}} + \frac{1}{\sigma_n}\left(b_n - M_n + \sum_{r=1}^{K} M_{I_r^{(n)}}\right), \qquad n \ge n_1. \tag{43}$$

By (40) we obtain the following convergences in L_s:

$$\lim_{n\to\infty} \frac{\sigma_{I_r^{(n)}}}{\sigma_n} = \lim_{n\to\infty}\left(\frac{g(I_r^{(n)})}{g(n)}\right)^{1/2} = A_r^*, \qquad r = 1, \dots, K,$$

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{\sigma_n}&\left(b_n - M_n + \sum_{r=1}^{K} M_{I_r^{(n)}}\right)\\ &= \lim_{n\to\infty} \frac{1}{g^{1/2}(n)}\left(b_n - f(n) + \sum_{r=1}^{K} f(I_r^{(n)})\right) = b^*.\end{aligned}$$

Furthermore, we have, for all $l \ge 0$,

$$\mathbb{E}\left[\mathbf{1}_{\{I_r^{(n)} \le l\} \cup \{I_r^{(n)} = n\}}\left(\frac{I_r^{(n)}}{n}\right)^{s/2}\right] \le \left(\frac{l}{n}\right)^{s/2} + \mathbb{P}\big(I_r^{(n)} = n\big) \to 0$$

for $n \to \infty$. Since all the conditions of Theorem 4.1 are satisfied and X is the unique solution of

$$X \stackrel{\mathcal{D}}{=} \sum_{r=1}^{K} A_r^* X^{(r)} + b^*$$

in $\mathcal{M}_s^d(0,1)$ subject to $(A_1, \dots, A_K), X^{(1)}, \dots, X^{(K)}$ independent and $X \sim X^{(r)}$ for all $r = 1, \dots, K$, the assertion follows. □

REMARK. Note that for $s = 2$ the order of the variance can be guessed from convergence in (40) and need not be known for the application of Theorem 5.1. Moreover, if the theorem can be applied for some g and $s = 2$, we obtain from the ζ_2 convergence in Theorem 4.1,

$$\mathrm{Var}(Y_n) = \mathrm{Var}(X) g(n) + o(g(n)) \tag{44}$$

since $h(x) := x^2/2$, $x \in \mathbb{R}$, is in $\mathcal{F}_2$. Similarly if the theorem can be applied for some f, g and $s = 1$, we obtain

$$\mathbb{E}Y_n = \mathbb{E}[X] g^{1/2}(n) + f(n) + o\big(g^{1/2}(n)\big) \tag{45}$$

AAP ims v.2003/08/05 Prn:1/09/2003; 11:54 F:aap165.tex; (Aiste) p. 22

since the identity map is in $\mathcal{F}_1$. The properties that the variance (or mean) can be guessed and proved by the application of the method are known for the ℓ_2 and ℓ_1 metric approaches. From this point of view ζ_2 and ζ_1 are as powerful as ℓ_2 and ℓ_1, respectively. However, the range $2 < s \leq 3$ for ζ_s leads to applications especially including normal limit laws which are out of reach for the ℓ_p metrics for all $p \geq 1$.

The case $0 < s \leq 2$ leads to examples which previously have been treated by the ℓ_p metrics. We first give some applications for these cases and focus then on the range $2 < s \leq 3$, where we use the following specialization of Theorem 5.1 for the application to normal limit laws, thus extending the previously known framework of the contraction method.

COROLLARY 5.2 (Central limit theorem). *Let* (Y_n) *be* s*-integrable*, $s > 2$, *and satisfy the recursive equation* (39) *with*

$$\mathbb{E}Y_n = f(n) + o\big(g^{1/2}(n)\big), \qquad \mathrm{Var}(Y_n) = g(n) + o(g(n)).$$

Assume for all $r = 1, \dots, K$ *and some* $2 < s \leq 3$,

$$\left(\frac{g(I_r^{(n)})}{g(n)}\right)^{1/2} \to A_r^*,$$
$$\frac{1}{g^{1/2}(n)}\left(b_n - f(n) + \sum_{r=1}^{K} f(I_r^{(n)})\right) \to 0 \qquad \text{in } L_s \tag{46}$$

and

$$\sum_{r=1}^{K} (A_r^*)^2 = 1, \qquad \mathbb{P}(\exists r : A_r^* = 1) < 1. \tag{47}$$

Then

$$\frac{Y_n - f(n)}{g^{1/2}(n)} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1),$$

where $\mathcal{N}(0, 1)$ *denotes the standard normal distribution*.

PROOF. The conditions of the univariate transfer theorem, Theorem 5.1, are satisfied with $b^* = 0$. The solution of the fixed-point equation (42) in $\mathcal{M}_s^1(0, 1)$ is $\mathcal{N}(0, 1)$ since $\sum_{r=1}^K (A_r^*)^2 = 1$, $b^* = 0$. $\square$

In the following sections we give applications of Theorem 5.1 and Corollary 5.2 for various choices of the parameters. We do not recall the underlying structures, for example, the various types of random graphs, but do introduce the essential recursive equations satisfied by the parameters of interest to cover the particular algorithm or data structure. References to basic accounts on them are given.

5.2. *Nonnormal limit laws.*

Quickselect. The number of key comparisons Y_n of the selection algorithm Quickselect, also known as FIND (for definition and mathematical analysis, see [18, 24, 32, 41]) when selecting the smallest order statistic in a set of n data satisfies $Y_0 = Y_1 = 0$ and the recurrence

$$Y_n \stackrel{\mathcal{D}}{=} Y_{I_n} + n - 1, \qquad n \geq 2,$$

where (Y_n) and I_n are independent with I_n unif$\{0, \dots, n-1\}$ distributed. Here, I_n corresponds to the size of the left sublist generated by the first splitting step of the algorithm. We can directly apply Theorem 5.1 with $K = 1$, $s = 1$, $f(n) = 0$, $g(n) = n^2$, $I_1^{(n)} = I_n$ and $b_n = n - 1$. We have $(g(I_n)/g(n))^{1/2} \to U$ and $b_n/g^{1/2}(n) \to 1$ in L_1 for an appropriate unif[0, 1] random variate U. Condition (41) is satisfied because we have $\mathbb{E}U = 1/2 < 1$. Hence we obtain

$$\frac{Y_n}{n} \xrightarrow{\mathcal{L}} X, \qquad X \stackrel{\mathcal{D}}{=} UX + 1, \tag{48}$$

where X and U are independent. This limit distribution is also known as the Dickman distribution, which arises in number theory (see [58] and [24]). This can easily be rederived by checking that the Dickman distribution satisfies the fixed-point relationship (48).

For the case where we apply the Quickselect algorithm to select an order statistic of $o(n)$ from a set of n data, we obtain the same limit distribution. This can be derived via a slight generalization of Theorem 5.1 and is as well covered with different approaches by all four references given above.

COROLLARY 5.3. *The normalized number of comparisons Y_n/n of Quickselect when selecting an order statistic of $o(n)$ from a set of n data converges in distribution to the Dickman distribution given as the unique solution of the fixed point equation in* (48).

Quicksort. The number of key comparisons Y_n of the sorting algorithm Quicksort applied to a randomly permuted list of n numbers (see [36]) satisfies $Y_0 = Y_1 = 0$ and the recurrence

$$Y_n \stackrel{\mathcal{L}}{=} Y_{I_n} + Y^*_{n-I_n} + n - 1, \qquad n \geq 2, \tag{49}$$

with (Y_n), (Y_n^*) and I_n independent and I_n unif$\{0, \dots, n-1\}$ distributed. It is well known that $\mathbb{E}Y_n = 2n \ln n + cn + o(n)$ for a constant $c \in \mathbb{R}$. Theorem 5.1 can be applied with $K = 2, s = 2$, $f(n) = 2n \ln n + cn$, $g(n) = n^2$, $I_1^{(n)} = I_n$, $I_2^{(n)} = n - I_n$ and $b_n = n - 1$. We have $(g(I_n)/g(n))^{1/2} \to U$ and $(g(n - I_n)/g(n))^{1/2} \to 1 - U$ in L_2 for an appropriate unif[0, 1] random variate U. A direct calculation,

AAP ims v.2003/08/05 Prn:1/09/2003; 11:54 F:aap165.tex; (Aiste) p. 24

(see [50]) shows that

$$\begin{aligned}\frac{1}{g^{1/2}(n)}&\bigl(b_n - f(n) + f(I_n) + f(n-I_n)\bigr)\\ &\to \mathcal{E}(U)\\ &= 1 + 2U\ln(U) + 2(1-U)\ln(1-U)\end{aligned}$$

in L_2. Condition (41) is satisfied because we have $\mathbb{E}[U^2 + (1-U)^2] = 2/3 < 1$. Thus Theorem 5.1 yields the convergence

$$\frac{Y_n - 2n\ln n - cn}{n} \xrightarrow{\mathcal{L}} X \in \mathcal{M}_2^1(0,1), \qquad X \stackrel{\mathcal{D}}{=} UX + (1-U)X^* + \mathcal{E}(U), \tag{50}$$

with X, X^* and U independent and $X \sim X^*$. This was first obtained by Rösler [50].

COROLLARY 5.4. *The normalized number of comparisons* $(Y_n - \mathbb{E}Y_n)/n$ *of Quicksort when sorting a set of* n *randomly permuted data converges in distribution to the unique solution in* $\mathcal{M}_2^1(0,1)$ *of the fixed point equation in* (50).

This convergence holds as well in the ζ_2 metric following from Theorem 4.1. Note that Corollaries 4.3 and 4.5 imply local and global convergence theorems for sequences (h_n) tending to zero sufficiently slowly. The exact order of the rate of convergence of the standardized cost of Quicksort for the ζ_3 metric has been identified to be of the order $\Theta(\ln(n)/n)$; see [45]. Hence, on the basis of this refinement we obtain as well rates of convergence for global and local convergences by applying Corollaries 4.3 and 4.5. This was made explicit by Neininger and Rüschendorf [45] for the case of Quicksort given in (49). Following this scheme, the general inequalities of Section 4.2 allow similar local and global approximation results for all the examples mentioned in Section 5.

Path length in m*-ary search trees.* The internal path length Y_n of random m-ary search trees, $m \geq 2$, containing n data satisfies $Y_j = j$ for $j = 0, \ldots, m-1$ and

$$Y_n \stackrel{\mathcal{D}}{=} \sum_{r=1}^{m} Y^{(r)}_{I_r^{(n)}} + n - m + 1, \qquad n \geq m,$$

with independence conditions as in recurrence (39). Let $V = (U_{(1)}, U_{(2)} - U_{(1)}, \ldots, 1 - U_{(m-1)})$ denote the vector of spacings of independent unif$[0,1]$ random variables $U_1, \ldots, U_{m-1}$. For $u \in [0,1]^m$ with $\sum u_r = 1$ the conditional distribution of $I^{(n)}$ given $V = u$ is multinomial $M(n-m+1, u)$. It is known that $\mathbb{E}Y_n = (H_m - 1)^{-1} n\ln(n) + c_m n + o(n)$ with constants $c_m \in \mathbb{R}$. We apply Theorem 5.1 with $K = m$, $s = 2$, $f(n) = (H_m - 1)^{-1} n\ln(n) + c_m n$, $g(n) = n^2$

and $b_n = n - m + 1$. The conditional distribution of $I^{(n)}$ given $V = u$ implies $I_r^{(n)}/n \to V_r$ in L_2. A calculation similar to the Quicksort case (see [44]) yields

$$\frac{1}{g^{1/2}(n)}\left(b_n - f(n) + \sum_{r=1}^{m} f(I_r^{(n)})\right) \to 1 + \frac{1}{H_m - 1}\sum_{r=1}^{m} V_r \ln(V_r).$$

Since $\mathbb{E}\sum V_r^2 < 1$, we rederive a limit law from [44].

COROLLARY 5.5. *The normalized internal path length $(Y_n - \mathbb{E}Y_n)/n$ of random m-ary search trees converges in distribution to the unique solution in $\mathcal{M}_2^1(0,1)$ of the fixed point equation*

$$X \stackrel{\mathcal{D}}{=} \sum_{r=1}^{m}(V_r X^{(r)}) + 1 + \frac{1}{H_m - 1}\sum_{r=1}^{m} V_r \ln(V_r),$$

with $X^{(1)}, \dots, X^{(m)}$, V independent and $X^{(r)} \sim X$ for $r = 1, \dots, m$.

5.3. *Normal limit laws.*

5.3.1. *Linear mean and variance.*

Size of random m-ary search trees. The size Y_n of the random m-ary search tree (see [34]) containing n data satisfies $Y_0 = 0$, $Y_1 = \cdots = Y_{m-1} = 1$ and the recursion

$$Y_n \stackrel{\mathcal{D}}{=} \sum_{r=1}^{m} Y^{(r)}_{I_r^{(n)}} + 1, \qquad n \ge m.$$

Let $V = (U_{(1)}, U_{(2)} - U_{(1)}, \dots, 1 - U_{(m-1)})$ denote the vector of spacings of independent unif[0, 1] random variables $U_1, \dots, U_{m-1}$ as in the previous example. For $u \in [0,1]^m$ with $\sum u_r = 1$ the conditional distribution of $I^{(n)}$ given $V = u$ is multinomial $M(n - (m-1), u)$. Thus we obtain

$$\frac{I^{(n)}}{n} \to \bigl(U_{(1)}, U_{(2)} - U_{(1)}, \dots, 1 - U_{(m-1)}\bigr) \tag{51}$$

in $L_{1+\varepsilon}$. The mean and the variance satisfy, for $3 \le m \le 26$ (see [2, 7, 31, 38]),

$$\mathbb{E}Y_n = \frac{1}{2(H_m - 1)}n + O(1 + n^{\alpha - 1}), \qquad \operatorname{Var}(Y_n) = \gamma_m n + o(n),$$

with $\gamma_m > 0$ and $\alpha < 3/2$ depending as well on m. Thus with Corollary 5.2 we rederive the limit law (see [7, 33, 38]):

COROLLARY 5.6. *The normalized size $(Y_n - \mathbb{E}Y_n)/\sqrt{\operatorname{Var}(Y_n)}$ of a random m-ary search tree with $3 \le m \le 26$ converges in distribution to the standard normal distribution.*

PROOF. We apply Corollary 5.2 with $f(n) = (2(H_m - 1))^{-1}n$ and $g(n) = \gamma_m n$. Condition (46) follows from (51); (47) holds since $\alpha < 3/2$. $\square$

The case of linear mean and variance is ubiquitous for parameters that satisfy (39), where the toll function b_n is appropriately small. Such examples covered by Corollary 5.2 are the number of certain patterns in random binary search trees, secondary cost parameters of Quicksort or the cost of certain tree traversal algorithms, Quicksort, m-ary search tree or generalized Quicksort recursions with *small* toll functions; see [7, 8, 11, 12, 14, 16, 23].

5.3.2. *Periodic functions in the mean and variance.*

Size of random tries. The number Y_n of internal nodes of a random trie with n keys in the symmetric Bernoulli model (see [34]) satisfies $Y_0 = 0$ and

$$Y_n \stackrel{\mathcal{D}}{=} Y^{(1)}_{I_1^{(n)}} + Y^{(2)}_{I_2^{(n)}} + 1, \qquad n \geq 1,$$

where $I_1^{(n)}$ is $\mathrm{B}(n, 1/2)$ distributed and $I_2^{(n)} = n - I_1^{(n)}$. The mean and variance satisfy ([22])

$$\mathbb{E}Y_n = n\varpi_1(\log_2 n) + O(1), \qquad \mathrm{Var}(Y_n) = n\varpi_2(\log_2 n) + O(1), \tag{52}$$

where ϖ_1 and ϖ_2 are positive C^∞ functions with period 1. We obtain a limit law due to Jacquet and Régnier [26]:

COROLLARY 5.7. *The normalized size $(Y_n - \mathbb{E}Y_n)/\sqrt{\mathrm{Var}(Y_n)}$ of a random trie in the symmetric Bernoulli model converges in distribution to the standard normal distribution.*

PROOF. For the application of Corollary 5.2 we first check (47). We have

$$f(n) = n\varpi_1(\log_2 n), \qquad g(n) = n\varpi_2(\log_2 n),$$

where ϖ_1 and ϖ_2 are given in (52). Here and in the following discussion we use the convention $\log_2 n := 0$ for $n = 0$. We split the sample space into the ranges $A = \{|I_1^{(n)} - n/2| < n^{3/4}\}$ and $B = \{|I_1^{(n)} - n/2| \geq n^{3/4}\}$. Note that on A and B also $|I_2^{(n)} - n/2| < n^{3/4}$ and $|I_2^{(n)} - n/2| \geq n^{3/4}$ hold, respectively.

For the estimate on the set A we assume n to be sufficiently large so that $n^{-1/4} < 1/4$. Then by the mean value theorem we obtain, for $r = 1, 2$,

$$1 + \log_2 \frac{I_r^{(n)}}{n} = 1 + \log_2\left(\frac{1}{2} + \frac{I_r^{(n)} - n/2}{n}\right) = \Psi_{n,r}\frac{I_r^{(n)} - n/2}{n}, \tag{53}$$

where the random $\Psi_{n,r}$ satisfy, on A,

$$\left\| \Psi_{n,r} - \frac{2}{\ln 2} \right\|_\infty \leq \frac{8}{\ln 2} n^{-1/4} \tag{54}$$

for n sufficiently large; in particular, $|\Psi_{n,r}| \le 4/\ln 2$, $r = 1, 2$. Thus on A we have $|1 + \log_2(I_r^{(n)}/n)| \le (4/\ln 2)n^{-1/4}$, which implies, using the periodicity of ϖ_1,

$$
\begin{aligned}
\varpi_1(\log_2 I_r^{(n)}) &= \varpi_1\left(\log_2 n + \left(1 + \log_2 \frac{I_r^{(n)}}{n}\right)\right) \\
&= \varpi_1(\log_2 n) + \Phi_{n,r}\left(1 + \log_2 \frac{I_r^{(n)}}{n}\right)
\end{aligned}
\tag{55}
$$

with random

$$
\Phi_{n,r} \in \left\{\varpi_1'(x + \log_2 n) : |x| \le \frac{4}{\ln 2} n^{-1/4}\right\} =: B_n, \qquad r = 1, 2.
$$

By the continuity of ϖ_1' it follows that the sets B_n are intervals. Since ϖ_1' is also periodic, we obtain that ϖ_1' is uniformly continuous and therefore the length of the intervals B_n tends to zero. This implies, on A,

$$
\|\Phi_{n,1} - \Phi_{n,2}\|_\infty \to 0, \qquad n \to \infty. \tag{56}
$$

Now we verify condition (47) to apply Corollary 5.2. Combining (53) and (55) we have, on A with n sufficiently large,

$$
\frac{1}{\sqrt{n}}\big(I_1^{(n)} \varpi_1(\log_2 I_1^{(n)}) + I_2^{(n)} \varpi_1(\log_2 I_2^{(n)}) - n\varpi_1(\log_2 n)\big) \tag{57}
$$

$$
= \frac{1}{\sqrt{n}}\left(\Phi_{n,1}\Psi_{n,1} I_1^{(n)} \frac{I_1^{(n)} - n/2}{n} + \Phi_{n,2}\Psi_{n,2} I_2^{(n)} \frac{I_2^{(n)} - n/2}{n}\right) \tag{58}
$$

$$
= \Phi_{n,1}\Psi_{n,1} \frac{1}{n^{3/2}}\Big(I_1^{(n)}\big(I_1^{(n)} - n/2\big) + I_2^{(n)}\big(I_2^{(n)} - n/2\big)\Big) \tag{59}
$$

$$
+ (\Phi_{n,2}\Psi_{n,2} - \Phi_{n,1}\Psi_{n,1}) \frac{I_2^{(n)}(I_2^{(n)} - n/2)}{n^{3/2}}. \tag{60}
$$

First we estimate summand (60) on A. By (54) and (56) we have $\|\Phi_{n,2}\Psi_{n,2} - \Phi_{n,1}\Psi_{n,1}\|_\infty \to 0$; thus

$$
\begin{aligned}
&\int_A \left|(\Phi_{n,2}\Psi_{n,2} - \Phi_{n,1}\Psi_{n,1}) \frac{I_2^{(n)}(I_2^{(n)} - n/2)}{n^{3/2}}\right|^3 d\mathbb{P} \\
&\quad \le \|\Phi_{n,2}\Psi_{n,2} - \Phi_{n,1}\Psi_{n,1}\|_\infty^3 \int \left|\frac{I_2^{(n)} - n/2}{\sqrt{n}}\right|^3 d\mathbb{P} \\
&\quad = o(1)\frac{1}{8}\mathbb{E}|\mathcal{N}(0, 1)|^3 \to 0, \qquad n \to \infty.
\end{aligned}
$$

For the first summand (59) we write $I_1^{(n)}/n = 1/2 + R_n$, so that on A we have $\|R_n\|_\infty \le n^{-1/4}$. Note that $I_2^{(n)}/n = 1/2 - R_n$. This yields

$$\frac{1}{n^{3/2}}\Big(I_1^{(n)}(I_1^{(n)} - n/2) + I_2^{(n)}(I_2^{(n)} - n/2)\Big) = 2R_n \frac{I_1^{(n)} - n/2}{\sqrt{n}}.$$

Since $|\Phi_{n,1}\Psi_{n,1}|$ remains bounded, say bounded by C, we obtain

$$\begin{aligned}\int_A \bigg|\Phi_{n,1}\Psi_{n,1}\frac{1}{n^{3/2}}\Big(I_1^{(n)}(I_1^{(n)} - n/2) + I_2^{(n)}(I_2^{(n)} - n/2)\Big)\bigg|^3 d\mathbb{P}\\ \le 2C\|R_n\|_\infty^3 \int \bigg|\frac{I_1^{(n)} - n/2}{\sqrt{n}}\bigg|^3 dP \to 0.\end{aligned}$$

Putting this altogether, we obtain

$$\int_A \bigg|\frac{1}{g^{1/2}(n)}(1 - f(n) + f(I_1^{(n)}) + f(I_2^{(n)}))\bigg|^3 dP \to 0.$$

By Chernoff's bound we have $\mathbb{P}(B) \le 2\exp(-\sqrt{n})$. With $m_2 := \min_{x\in[0,1]} \varpi_2(x) > 0$ we obtain

$$\begin{aligned}\int_B \bigg|\frac{I_r^{(n)}(\varpi_1(\log_2 I_r^{(n)}) - \varpi_1(\log_2 n))}{g^{1/2}(n)}\bigg|^3 dP\\ \le \bigg(\frac{2\|\varpi_1\|_\infty}{m_2^{1/2}}\bigg)^3 n^{3/2} 2\exp(-\sqrt{n}) \to 0\end{aligned}$$

for $r = 1, 2$, which implies

$$\int_B \bigg|\frac{1}{g^{1/2}(n)}(1 - f(n) + f(I_1^{(n)}) + f(I_2^{(n)}))\bigg|^3 dP \to 0.$$

Thus together we obtain (47).

For (46) note that with $g(n) := n\varpi_2(\log_2 n)$ we have

$$\frac{g(I_r^{(n)})}{g(n)} = \frac{I_r^{(n)}}{n}\frac{\varpi_2(\log_2 I_r^{(n)})}{\varpi_2(\log_2 n)}.$$

By the strong law of large numbers (SLLN) and and dominated convergence, $I_r^{(n)}/n \to 1/2$ in any L_p. Furthermore

$$\begin{aligned}\bigg|\frac{\varpi_2(\log_2 I_r^{(n)})}{\varpi_2(\log_2 n)} - 1\bigg| &= \bigg|\frac{\varpi_2(\log_2(n) + 1 + \log_2(I_r^{(n)}/n)) - \varpi_2(\log_2 n)}{\varpi_2(\log_2 n)}\bigg|\\ &\le \frac{1}{m_2}\bigg|\varpi_2\bigg(\log_2(n) + 1 + \log_2\bigg(\frac{I_r^{(n)}}{n}\bigg)\bigg) - \varpi_2(\log_2 n)\bigg|.\end{aligned}$$

By the SLLN and the continuity of ϖ_2 this tends to zero almost surely. By dominated convergence, the convergence holds in any L_p. Together this implies

$$\frac{g(I_r^{(n)})}{g(n)} \to \frac{1}{2}, \qquad \text{in } L_s$$

for any $s > 0$. $\square$

REMARK. Note that the properties of ϖ_1 and ϖ_2 that are needed are the periodicity, the lower positive bound for ϖ_2, that ϖ_1 is continuously differentiable and that ϖ_2 is continuous. This seems to be of some generality as the following examples on path lengths in digital structures show. However, there are also examples where the differentiability of ϖ_1 fails to hold and our method still may be applied as shown below for the top-down mergesort.

Path lengths in digital structures. The fundamental search trees based on bit comparisons are the digital search tree, the trie and the Patricia trie; see [57]. The cost to build up these trees from n data is measured by their path lengths Y_n, that is, the internal path length for digital search trees and the external path lengths for tries and Patricia tries. In the symmetric Bernoulli model we have $Y_0 = 0$ and

$$Y_n \stackrel{\mathcal{D}}{=} Y^{(1)}_{I_1^{(n)}} + Y^{(2)}_{I_2^{(n)}} + n - \tilde{b}_n, \qquad n \ge 1,$$

where $I_1^{(n)} \sim B(n-1, 1/2)$, $I_1^{(n)} + I_2^{(n)} = n - 1$ and $\tilde{b}_n = 1$ for digital search trees, and for the other two structures, $I_1^{(n)} \sim B(n, 1/2)$, $I_1^{(n)} + I_2^{(n)} = n$ and $\tilde{b}_n = 0$ for the trie and $\tilde{b}_n = \mathbf{1}_{\{I_1^{(n)} \in \{0,n\}\}} n$ for Patricia tries. The small disturbance in the recursion is reflected by their similar moments,

$$\mathbb{E} Y_n = n \log_2 n + n\varpi_3(\log_2 n) + O(\log n),$$
$$\operatorname{Var} Y_n = n\varpi_4(\log_2 n) + O(\log^2 n),$$

where ϖ_r are periodic functions (with period 1) that vary from one of the structures to the other; see [28–31]. It is known that ϖ_4 is continuous and positive in each case. For ϖ_3 we have the representations

$$\varpi_3(x) = C + \frac{1}{\ln 2} \sum_{k \in \mathbb{Z} \setminus \{0\}} \Gamma(-\omega_k) e^{2k\pi i x}, \qquad x \in \mathbb{R},$$

for the digital search tree and

$$\varpi_3(x) = C + \frac{1}{\ln 2} \sum_{k \in \mathbb{Z} \setminus \{0\}} \omega_k \Gamma(-\omega_k) e^{2k\pi i x}, \qquad x \in \mathbb{R},$$

for the trie and Patricia trie, where the constant C varies from case to case and

$$\omega_k = 1 + \frac{2k\pi i}{\ln 2}, \qquad k \in \mathbb{Z}.$$

Note that for Fourier series $h(x)=\sum_{k\in\mathbb{Z}}a_k e^{2\pi ikx}$ the condition $|a_k|=O(k^{-(m+2)})$ for $|k|\to\infty$ implies that h is m times continuously differentiable. Since $|\Gamma(-\omega_k)|$ decays exponentially for $|k|\to\infty$, we obtain that ϖ_3 is infinite differentiable in all cases. Only the C^1 property is needed. This implies asymptotic normality, which for digital search trees and tries was proven in [25] and [27], respectively; see also [54]:

COROLLARY 5.8. *The normalized internal* (*resp*., *external*) *path lengths* $(Y_n-\mathbb{E}Y_n)/\sqrt{\mathrm{Var}(Y_n)}$ *of digital search trees*, *tries and Patricia tries are asymptotically normal in the symmetric Bernoulli model*.

PROOF. We apply Corollary 5.2. Condition (46) follows from the continuity of ϖ_4 as in the proof of Corollary 5.7. For (47) we have

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\bigl(n-\tilde{b}_n-n\log_2 n-n\varpi_3(\log_2 n)+I_1^{(n)}\log_2 I_1^{(n)}\\
&\qquad+I_1^{(n)}\varpi_3(\log_2 I_1^{(n)})+I_2^{(n)}\log_2 I_2^{(n)}+I_2^{(n)}\varpi_3(\log_2 I_2^{(n)})\bigr)\\
&\quad=-\frac{\tilde{b}_n}{\sqrt{n}}+\frac{1}{\sqrt{n}}\bigl(n+I_1^{(n)}\log_2 I_1^{(n)}+I_2^{(n)}\log_2 I_2^{(n)}-n\log_2 n\bigr) &\text{(61)}\\
&\qquad+\frac{1}{\sqrt{n}}\bigl(I_1^{(n)}\varpi_3(\log_2 I_1^{(n)})+I_2^{(n)}\varpi_3(\log_2 I_2^{(n)})-n\varpi_3(\log_2 n)\bigr). &\text{(62)}
\end{aligned}$$

Now, $\tilde{b}_n/\sqrt{n}$ tends to zero in L_p for any $p>0$ in all three cases. The summand in (62) is essentially the term (57) estimated in the proof of Corollary 5.7. The second summand in (61) can also be seen to tend to zero in L_3 by the estimates of Corollary 5.7: Applying (53) leads to (58) with $\Phi_{n,1}=\Phi_{n,2}=1$ there; thus we can conclude as in Corollary 5.7. □

For related recursions that arise in the analysis of the size and path length of bucket digital search trees, see [19].

Mergesort. The number of key comparisons Y_n of top-down mergesort (for definition and mathematical analysis, see [15] and [20]), applied to a list of n randomly permuted items, satisfies $Y_0=0$ and

$$Y_n\stackrel{\mathcal{D}}{=}Y^{(1)}_{I_1^{(n)}}+Y^{(2)}_{I_2^{(n)}}+n-S_n,\qquad n\ge 1,$$

where $I_1^{(n)}=\lceil n/2\rceil$, $I_2^{(n)}=n-I_1^{(n)}$ and S_n is a random variate that is independent of $(Y_n^{(1)})$ and $(Y_n^{(2)})$; see [31], Section 5.2.4. In particular, we have $\mathbb{E}S_n^3=O(1)$. Flajolet and Golin [15] proved

$$\begin{aligned}
\mathbb{E}Y_n&=n\log_2 n+n\varpi_5(\log_2 n)+O(1), &\text{(63)}\\
\mathrm{Var}(Y_n)&=n\varpi_6(\log_2 n)+o(n),
\end{aligned}$$

where ϖ_5 and ϖ_6 are continuous functions with period 1, ϖ_6 is positive and ϖ_5 is not differentiable. In particular,

$$\varpi_5(u) = C + \frac{1}{\ln 2} \sum_{k \in \mathbb{Z} \setminus \{0\}} \frac{1 + \Psi(\chi_k)}{\chi_k(\chi_k + 1)} e^{2k\pi i u}, \qquad u \in \mathbb{R}, \tag{64}$$

where $C \in \mathbb{R}$ is a constant, Ψ is a complex function of $O(1)$ on the imaginary line $\Re(s) = 0$ and

$$\chi_k = \frac{2\pi i k}{\ln 2}, \qquad k \in \mathbb{Z}.$$

Moreover, Flajolet and Golin showed that $(Y_n - \mathbb{E}Y_n)/\sqrt{\operatorname{Var}(Y_n)}$ satisfies a central limit theorem by applying Lyapunov's condition. Hwang [20] found a local limit theorem and large deviations including rates of convergence and, in [21], gave full (exact) asymptotic expansions for the mean and variance. Cramer [9] obtained the (central) limit law by applying the contraction method. For methodological reasons, we rederive this limit law on the basis of the expansions (63) including the representation for ϖ_5 directly from Corollary 5.2: We have

$$f(n) = n \log_2(n) + n\varpi_5(\log_2 n), \qquad g(n) = n\varpi_6(\log_2 n).$$

Thus, by $I_1^{(n)} = \lceil n/2 \rceil$ and the continuity of ϖ_6, we have deterministically

$$\left(\frac{g(I_1^{(n)})}{g(n)}\right)^{1/2} \to \frac{1}{\sqrt{2}}, \qquad \left(\frac{g(I_2^{(n)})}{g(n)}\right)^{1/2} \to \frac{1}{\sqrt{2}}.$$

Since $\mathbb{E}S_n^3 = O(1)$ we obtain, in L_3,

$$\begin{aligned} &\frac{1}{\sqrt{n}}\bigl(f(I_1^{(n)}) + f(I_2^{(n)}) - f(n) + n - S_n\bigr) \\ &\quad = \frac{1}{\sqrt{n}}\Biggl(\left\lceil\frac{n}{2}\right\rceil \log_2\left\lceil\frac{n}{2}\right\rceil + \left(n - \left\lceil\frac{n}{2}\right\rceil\right) \log_2\left(n - \left\lceil\frac{n}{2}\right\rceil\right) \\ &\qquad\qquad - n\log_2(n) + \left\lceil\frac{n}{2}\right\rceil \varpi_5\left(\log_2\left\lceil\frac{n}{2}\right\rceil\right) \\ &\qquad\qquad + \left(n - \left\lceil\frac{n}{2}\right\rceil\right)\varpi_5\left(\log_2\left(n - \left\lceil\frac{n}{2}\right\rceil\right)\right) - n\varpi_5(\log_2 n) + n - S_n\Biggr) \\ &\quad \to 0, \end{aligned} \tag{65}$$

where the only nontrivial contribution comes from the ϖ_5 terms in (65) with odd n, say $n = 2m + 1$. The asymptotic cancellation of these terms follows from

$$\bigl|\varpi_5(\log_2(m+1)) - \varpi_5(\log_2(m+1/2))\bigr| = o(m^{-1/2}).$$

For this note that

$$\left|\exp(2k\pi i \log_2(m+1)) - \exp\left(2k\pi i \log_2\left(m+\frac{1}{2}\right)\right)\right|$$
$$\leq \min\left\{2, \frac{\pi k}{(m+1/2)\ln(2)}\right\}, \qquad k \in \mathbb{Z}, m \in \mathbb{N},$$

and that

$$\left|\frac{1+\Psi(\chi_k)}{\chi_k(\chi_k+1)}\right| \leq \frac{c}{k^2}, \qquad k \in \mathbb{Z}\setminus\{0\},$$

with a constant $c > 0$. Now we split the range of summation in (64) into the ranges $|k| \leq m$ and $|k| > m$ and note that $\sum_{k>m} k^{-2} = O(1/m)$. This yields

$$\left|\varpi_5(\log_2(m+1)) - \varpi_5\left(\log_2\left(m+\frac{1}{2}\right)\right)\right| = O\left(\frac{H_m}{m}\right).$$

COROLLARY 5.9. *The normalized number of key comparisons* $(Y_n - \mathbb{E}Y_n)/\sqrt{\mathrm{Var}(Y_n)}$ *of top-down mergesort applied to a randomly permuted number of items is asymptotically normal.*

For other variants of mergesort and a limit law for the queue mergesort, see [6] and the references therein. Note that the limit law for queue mergesort in [6] cannot be obtained by Corollary 5.2 since the corresponding prefactors $(g(I_r^{(n)})/g(n))^{1/2}$ do not converge, $r = 1, 2$. However, the extension of our approach in Section 5.2 seems to be promising.

5.3.3. *Other orders for mean and variance.*

Maxima in right triangles. We consider the number Y_n of maxima of n independent, uniform samples in a right triangle in $\mathbb{R}^2$ with vertices $(0,0)$, $(1,0)$ and $(0,1)$; see [3] and [4]. According to Proposition 1 in [3], this number satisfies the recursion $Y_0 = 0$ and

$$Y_n \stackrel{\mathcal{D}}{=} Y^{(1)}_{I_1^{(n)}} + Y^{(2)}_{I_2^{(n)}} + 1,$$

where the indices $I_1^{(n)}$ and $I_2^{(n)}$ are given as the first two components of a mixed trinomial distribution as follows: Let (U_n, V_n) denote the point that maximizes the sum of the components in the sample of the n points. Then $I^{(n)} = (I_1^{(n)}, I_2^{(n)}, I_3^{(n)})$ conditioned on $(U_n, V_n) = (u, v)$ is multinomially distributed:

$$\mathbb{P}^{I^{(n)}|(U_n,V_n)=(u,v)} = M\left(n-1, \frac{u^2}{(u+v)^2}, \frac{v^2}{(u+v)^2}, \frac{2uv}{(u+v)^2}\right).$$

The mean and variance satisfy (see [3])

$$\mathbb{E}Y_n = \sqrt{\pi}\sqrt{n} + O(1),$$
$$\mathrm{Var}(Y_n) = \sigma^2\sqrt{n} + O(1),$$

with some $\sigma^2 > 0$. Thus we have the orders $f(n) = \sqrt{\pi}\sqrt{n}$ and $g(n) = \sigma^2\sqrt{n}$. Moreover, $I^{(n)}/n \to (U^2, (1-U)^2, 2U)$, where U is unif[0, 1] distributed. So we obtain $I_1^{(n)}/n \to U^{1/2}$ and $I_2^{(n)}/n \to (1-U)^{1/2}$ in any L_p. In [4] by the method of moments, a central limit law for Y_n was derived. We can give an easy approach to asymptotic normality based on Corollary 5.2:

COROLLARY 5.10. *Let Y_n be the number of maxima of n independent uniform samples in a right triangle as described above. Then*

$$\frac{Y_n - \mathbb{E}Y_n}{\sqrt{\mathrm{Var}(Y_n)}} \stackrel{\mathcal{L}}{\to} \mathcal{N}(0,1).$$

PROOF. It remains to check (47). We have

$$\begin{aligned}
&\frac{1}{n^{1/4}}\left|(I_1^{(n)})^{1/2} + (I_2^{(n)})^{1/2} - n^{1/2}\right| \\
&\qquad \le n^{1/4}\left|\left(\frac{I_1^{(n)}}{n-1}\right)^{1/2} - \frac{U_n}{U_n+V_n}\right| + n^{1/4}\left|\left(\frac{I_2^{(n)}}{n-1}\right)^{1/2} - \frac{V_n}{U_n+V_n}\right| \\
&\qquad\quad + O(n^{-3/4}).
\end{aligned}$$

It is sufficient to bound one of these two nontrivial summands. We denote

$$A := \left\{\left(\frac{I_1^{(n)}}{n-1}\right)^{1/2} + \frac{U_n}{U_n+V_n} \ge n^{-3/14}\right\},$$
$$B := \left\{\left(\frac{I_1^{(n)}}{n-1}\right)^{1/2} + \frac{U_n}{U_n+V_n} < n^{-3/14}\right\}.$$

On A it holds that

$$\left|\left(\frac{I_1^{(n)}}{n-1}\right)^{1/2} - \frac{U_n}{U_n+V_n}\right|^3 \le n^{9/14}\left|\frac{I_1^{(n)}}{n-1} - \frac{U_n^2}{(U_n+V_n)^2}\right|^3.$$

Thus denoting by Υ_n the distribution of $U_n^2/(U_n+V_n)^2$ and denoting by $B_{n,u}$ a $B(n-1,u)$ distributed random variable, we obtain

$$\begin{aligned}
\int_A \left|\left(\frac{I_1^{(n)}}{n-1}\right)^{1/2} - \frac{U_n}{U_n+V_n}\right|^3 dP &\le n^{9/14}\int\left|\frac{B_{n,u}}{n-1} - u\right|^3 d\Upsilon_n(u) \\
&\le n^{9/14}Cn^{-3/2} = Cn^{-6/7},
\end{aligned}$$

AAP ims v.2003/08/05 Prn:1/09/2003; 11:54 F:aap165.tex; (Aiste) p. 34

since the third absolute central moment of $B_{n,u}$ is bounded by $Cn^{3/2}$ with a constant $C > 0$ uniformly in $p \in [0, 1]$, followed by Chernoff's bound. On B we use the trivial bound

$$\int_B \left| \left(\frac{I_1^{(n)}}{n-1} \right)^{1/2} - \frac{U_n}{U_n + V_n} \right|^3 d\mathbb{P} \leq \mathbb{P}(B) n^{-9/14}.$$

Finally we obtain, with Υ_n' denoting the distribution of $U_n + V_n$,

$$\begin{aligned}
\mathbb{P}(B) &= \mathbb{P}\left(\frac{U_n}{U_n + V_n} \leq n^{-3/14} \right) \\
&= \mathbb{P}\left(\frac{U_n}{U_n + V_n} \leq n^{-3/14}, 0 \leq U_n + V_n \leq \frac{1}{2} \right) \\
&\quad + \mathbb{P}\left(\frac{U_n}{U_n + V_n} \leq n^{-3/14}, \frac{1}{2} \leq U_n + V_n \leq 1 \right) \\
&\leq \left(\frac{1}{4} \right)^n + \mathbb{P}\left(U_n \leq n^{-3/14}, U_n + V_n \geq \frac{1}{2} \right) \\
&= \left(\frac{1}{4} \right)^n + \int_{1/2}^1 \frac{\sqrt{2} n^{-3/14}}{\sqrt{2u}} \, d\Upsilon_n'(u) \\
&= O(n^{-3/14}).
\end{aligned}$$

Thus putting everything altogether, we obtain

$$\begin{aligned}
&\left\| \frac{1}{g^{1/2}(n)} \bigl(f(n) - f(I_1^{(n)}) - f(I_2^{(n)}) \bigr) \right\|_3 \\
&\qquad = O\bigl((n^{3/4} n^{-6/7})^{1/3} \bigr) = O(n^{-1/28}).
\end{aligned} \tag{66}$$

The assertion follows by Corollary 5.2. □

Note that estimates for the term (66) are also required in the moments method approach in [3]; compare the (different) estimates of $V_r(n)$ there on pages 14 and 15.

5.4. *Multivariate central limit laws.* The generalization of Corollary 5.2 to higher dimensions is straightforward and is omitted here. Applications of such an extension cover, for example, limit laws for recursions as in [33]. Instead, we extend our approach to multivariate recursions where the number K of copies of the parameter on the right side may be random, depending on n, and even satisfy $K = K_n \to \infty$ for $n \to \infty$. Clearly, in such a situation the heuristic that a stabilization of the coefficients as in (24) plus a contraction condition as in (25) leads to a fixed-point equation and a related convergence theorem as in

Theorem 4.1 no longer holds. However, for normal convergence problems, we use the fact that the multivariate standard normal distribution satisfies *all* the equations

$$X \stackrel{\mathcal{D}}{=} \sum_{r=1}^{K_n} A_r X^{(r)} \tag{67}$$

with $(K_n, A_1, A_2, \dots), X^{(1)}, X^{(2)}, \dots$ independent, $X \sim X^{(r)}$ for all $r \geq 1$ and $\sum_{r=1}^{K_n} A_r A_r^t = \mathrm{Id}_d$. For problems with $K_n \to \infty$, we typically have A_r tending to zero for $n \to \infty$ for every fixed r, so that we have no meaningful limiting equation at hand. Nevertheless, our arguments may apply, since we are able to compare with a normal distribution for every n by means of an (in n changing) equation of the type (67).

5.4.1. *Random recursive trees.* The recursive tree of order n is a rooted tree on n vertices labeled 1 to n such that for each $k = 2, \dots, n$, the labels of the vertices on the path from the root to the node labeled k form an increasing sequence. For a random recursive tree of order n, we assume the uniform model on the space of all recursive trees of order n, where all $(n-1)!$ trees are equally likely. It is well known that a random recursive tree may be obtained by successively adjoining children to the existing tree, where for the nth vertex the parent node is chosen uniformly from the vertices labeled $1, \dots, n-1$. For a survey of recursive trees, we refer to [56]. Mahmoud and Smythe [40] studied the joint distribution of $Y_n = (B_n, R_n, G_n)$, where B_n, R_n and G_n are the number of vertices in the tree with out-degree 0, 1 and 2, respectively. Based on a formulation as a generalized Pólya–Eggenberger urn model, they derived the mean $\mathbb{E} Y_n = n(1/2, 1/4, 1/8) + O(1)$ and the covariance matrix $\mathrm{Cov}(Y_n) = n\Sigma_0 + O(1)$, where Σ_0 is explicitly given in [40], Theorem 4.1. By an application of a martingale central limit theorem, the asymptotic trivariate normality is shown.

Here, we offer a recursive approach to parameters of random recursive trees based on a generalization of our previous settings, allowing the number K of summands on the right-hand side of (1) to be random, depending on n, and tending to infinity for $n \to \infty$.

There are several possibilities for a recursive decomposition of a parameter of a random recursive tree: First of all we may decompose by counting the parameters of all the K_n subtrees separately and calculate from these the parameter of the whole tree. This is the line we follow in the subsequent analysis. Here the random number K_n of subtrees of the root has a representation as a sum of independent Bernoulli random variables. Second, we could also subdivide into, for example, the leftist subtree of the root and the rest of the tree (including the root) as was done in [13] for the analysis of the internal path length in random recursive trees. However, the example of Mahmoud and Smythe [40] will, in this decomposition, not be covered by our present approach, since a dependence between b_n and $(Y_n^{(2)})$ is present which is forbidden in our setup. Third, we may first use a bijection between recursive trees and binary search trees, the "oldestchildnextsibling" (see [1]), transpose the parameter under consideration into the binary search tree setting and use the recursive structure of the binary search tree. However, parameters of a simple form for recursive trees may have more complex counterparts in the binary search tree setup.

The vector Y_n defined above satisfies the recursions $Y_0 = (0, 0, 0)$, $Y_1 = (1, 0, 0)$ and

$$Y_n \stackrel{\mathcal{D}}{=} \sum_{r=1}^{K_n} Y^{(r)}_{I_r^{(n)}} + b_n, \qquad n \geq n_0, \tag{68}$$

with $n_0 = 2$ and $b_n = (\mathbf{1}_{\{K_n=0\}}, \mathbf{1}_{\{K_n=1\}}, \mathbf{1}_{\{K_n=2\}})$, where $(K_n, I^{(n)}, b_n)$, $(Y_n^{(1)})$, $(Y_n^{(2)}), \dots$ are independent with $Y_j^{(r)} \sim Y_j$ for all $j \geq 0$, $r \geq 1$. Here, we use that given the number K_n and the cardinalities $I_1^{(n)}, \dots, I_{K_n}^{(n)}$ of the subtrees of the root, these subtrees are random recursive trees of orders $I_1^{(n)}, \dots, I_{K_n}^{(n)}$, respectively, and are independent of each other.

More generally, we assume that (Y_n) is a sequence of random d-dimensional vectors satisfying (68), where b_n is "small"; more precisely, $b_n/\sqrt{n} \to 0$ in L_3, and mean and (co)variances of Y_n are linear, that is,

$$\mathbb{E}Y_n = n\mu + o\big(\sqrt{n}\big), \qquad \mathrm{Cov}(Y_n) = n\Sigma + o(n), \tag{69}$$

with components $\mu_l \neq 0$ for $l = 1, \dots, d$ and Σ being positive definite. We assume $\mathrm{Cov}(Y_n)$ to be positive definite for some $n \geq n_1$ and scale as in (22) with $2 < s \leq 3$, $X_n := C_n^{-1/2}(Y_n - M_n)$. This leads to the recursion

$$X_n \stackrel{\mathcal{D}}{=} \sum_{r=1}^{K_n} A_r^{(n)} X^{(r)}_{I_r^{(n)}} + b^{(n)}, \qquad n \geq n_1, \tag{70}$$

with

$$A_r^{(n)} = \left(\frac{I_r^{(n)}}{n}\right)^{1/2} \mathrm{Id}_d + o(1), \qquad b^{(n)} = o(1), \tag{71}$$

where the $o(1)$ terms are converging uniformly. A substitute for the contraction condition (25) is given by the next lemma:

LEMMA 5.11. *Let K_n be the out-degree of the root of a random recursive tree of order n and let $I_1^{(n)}, \dots, I_{K_n}^{(n)}$ be the cardinalities of the subtrees of the root. Then, for all $s > 2$,*

$$\limsup_{n\to\infty} \mathbb{E} \sum_{r=1}^{K_n} \left(\frac{I_r^{(n)}}{n}\right)^{s/2} \leq \frac{6+s}{4+2s} < 1$$

holds.

PROOF. By Theorem 5.1 in [39] we have, in particular, $I_1^{(n)}/n \to U \sim$ unif[0, 1] almost surely. Hence it follows that

$$\begin{aligned}
\mathbb{E}\sum_{r=1}^{K_n}\left(\frac{I_r^{(n)}}{n}\right)^{s/2} &\leq \mathbb{E}\left[\left(\frac{I_1^{(n)}}{n}\right)^{s/2} + \sum_{r=2}^{K_n}\frac{I_r^{(n)}}{n}\right] \\
&= \mathbb{E}\left[\left(\frac{I_1^{(n)}}{n}\right)^{s/2} + \frac{n-1-I_1^{(n)}}{n}\right] \\
&\to \frac{6+s}{4+2s}, \qquad n\to\infty.
\end{aligned} \tag{72}$$

The assertion follows. $\square$

COROLLARY 5.12. *Let* (Y_n) *be a sequence of vectors of parameters of a random recursive tree in* L_3 *satisfying* (68) *and* (69). *Then, for all* $2 < s \leq 3$,

$$\zeta_s\big(\mathrm{Cov}(Y_n)^{-1/2}(Y_n - \mathbb{E}Y_n), \mathcal{N}(0, \mathrm{Id}_d)\big) \to 0, \qquad n\to\infty,$$

holds.

PROOF. For an estimate as in the proof of Theorem 4.1, we introduce the accompanying sequence

$$Q_n \stackrel{\mathcal{D}}{=} \sum_{r=1}^{K_n} A_r^{(n)}\Big(\mathbf{1}_{\{I_r^{(n)}<n_1\}}X^{(r)}_{I_r^{(n)}} + \mathbf{1}_{\{I_r^{(n)}\geq n_1\}}N^{(r)}\Big) + b^{(n)}, \qquad n\geq 2,$$

where $(A_1^{(n)},\dots,A_{K_n}^{(n)}, b^{(n)}, I^{(n)}, K_n), N^{(1)}, N^{(2)},\dots,(X_n^{(1)}),(X_n^{(2)}),\dots$ are independent with $N^{(r)}\sim\mathcal{N}(0,\mathrm{Id}_d)$, $X_j^{(r)}\sim X_j$ for $r\geq 1$, $0\leq j<n_1-1$. Then we have

$$\zeta_s\big(X_n, \mathcal{N}(0,\mathrm{Id}_d)\big) \leq \zeta_s(X_n, Q_n) + \zeta_s\big(Q_n, \mathcal{N}(0,\mathrm{Id}_d)\big).$$

As in the proof of Theorem 4.1, we obtain [cf. (33)]

$$\zeta_s(X_n, Q_n) \leq \int\!\!\int \sum_{r=1}^{k} \mathbf{1}_{\{j_r\geq n_1\}}\|\alpha_r\|_{\mathrm{op}}^s \zeta_s(X_{j_r}, X)\, d\Upsilon_n^{(k)}(\alpha, j)\, d\mathbb{P}^{K_n}(k), \tag{73}$$

where $\Upsilon_n^{(k)}$ is the joint distribution of $(A_1^{(n)},\dots,A_{K_m}^{(n)}, I^{(n)})$ given $K_n = k$ and $\alpha = (\alpha_1,\dots,\alpha_k)$, $j = (j_1,\dots,j_k)$. This implies

$$\zeta_s\big(X_n, \mathcal{N}(0,\mathrm{Id}_d)\big) \leq \left(\mathbb{E}\sum_{r=1}^{K_n}\|A_r^{(n)}\|_{\mathrm{op}}^s\right) \sup_{n_1\leq j\leq n-1}\zeta_s\big(X_j, \mathcal{N}(0,\mathrm{Id}_d)\big) + o(1). \tag{74}$$

By (71) and Lemma 5.11 the prefactor is less than or equal to $(6+s)/(4+2s) < 1$

AAP ims v.2003/08/05 Prn:1/09/2003; 11:54 F:aap165.tex; (Aiste) p. 38

for n sufficiently large, which gives that $(\zeta(X_n, \mathcal{N}(0, \mathrm{Id}_d)))$ is bounded. Then with $\xi := \limsup \zeta(X_n, \mathcal{N}(0, \mathrm{Id}_d))$, we obtain from (73), as in the proof of Theorem 4.1,

$$\xi \le \frac{6+s}{4+2s}(\xi + \varepsilon)$$

for all $\varepsilon > 0$; thus $\xi = 0$. $\square$

As a consequence we obtain, in particular, the asymptotic trivariate normality result of Mahmoud and Smythe [40].

5.4.2. *Random plane-oriented recursive trees.* Plane-oriented recursive trees are recursive trees with ordered sets of descendents; a random plane-oriented recursive tree of order n is chosen with equal probability from the space of all such trees, with n nodes. As for random recursive trees, there is a probabilistic growth rule available to build up the plane-oriented counterpart; for details and definitions, see [56].

Parameters of random recursive trees which admit a representation (68) have counterparts for plane-oriented recursive trees, where (68) is still valid, but the distribution of $(K_n, I^{(n)})$ has to be adjusted. These two trees behave quite differently. For example, the growth order of the number of subtrees K_n of the root, on the average, changes from being logarithmic to $\sqrt{n}$ when switching to the plane-oriented version. A convergence scheme similar to Corollary 5.12 can be built upon the following contraction lemma, which corresponds to Lemma 5.11:

LEMMA 5.13. *Let K_n be the out-degree of the root of a random plane-oriented recursive tree of order n and let $I_1^{(n)}, \dots, I_{K_n}^{(n)}$ be the cardinalities of the subtrees of the root. Then, for all $s > 2$,*

$$\limsup_{n\to\infty} \mathbb{E} \sum_{r=1}^{K_n} \left(\frac{I_r^{(n)}}{n}\right)^{s/2} \le \frac{5+2s}{3+3s} < 1$$

holds.

PROOF. We enumerate the subtrees of the root such that the first subtree in our enumeration is the one with root labeled 2. Then by Theorem 5 in [42], $I^{(n)}/n \to V$ holds almost surely, where V has the beta$(1/2, 1)$ distribution. In particular, $\mathbb{E}V^{s/2} = 1/(s+1)$ for $s > 0$. An estimate similar to (72) leads to the assertion. $\square$

From this contraction property we obtain that our Corollary 5.12 is valid also for the plane-oriented version of the recursive tree.

COROLLARY 5.14. *Let* (Y_n) *denote the vector of parameters of plane-oriented recursive trees in* L_3 *that satisfy* (68) *and* (69). *Then*

$$\zeta_s\big(\mathrm{Cov}(Y_n)^{-1/2}(Y_n - EY_n), \mathcal{N}(0, \mathrm{Id}_d)\big) \to 0, \qquad n \to \infty.$$

For possible applications, see Section 8 in [42].

Acknowledgment. We thank Hsien-Kuei Hwang for pointing out some of the applications in Section 5. The referee's comments are very much appreciated.

REFERENCES

[1] AHO, A. V., HOPCROFT, J. E. and ULLMAN, J. D. (1983). *Data Structures and Algorithms*. Addison–Wesley, Reading, MA.

[2] BAEZA-YATES, R. A. (1987). Some average measures in m-ary search trees. *Inform. Process. Lett.* **25** 375–381.

[3] BAI, Z.-D., HWANG, H.-K., LIANG, W.-Q. and TSAI, T.-H. (2001). Limit theorems for the number of maxima in random samples from planar regions. *Electronic J. Probab.* **6**.

[4] BAI, Z.-D., HWANG, H.-K. and TSAI, T.-H. (2002). Berry–Esseen bounds for the number of maxima in planar regions. Preprint.

[5] BURTON, R. and RÖSLER, U. (1995). An L_2 convergence theorem for random affine mappings. *J. Appl. Probab.* **32** 183–192.

[6] CHEN, W.-M., HWANG, H.-K. and CHEN, G.-H. (1999). The cost distribution of queue-mergesort, optimal mergesorts, and power-of-2 rules. *J. Algorithms* **30** 423–448.

[7] CHERN, H.-H. and HWANG, H.-K. (2001). Phase changes in random m-ary search trees and generalized quicksort. *Random Structures Algorithms* **19** 316–358.

[8] CHERN, H.-H., HWANG,H.-K. and TSAI, T.-H. (2002). An asymptotic theory for Cauchy–Euler differential equations with applications to the analysis of algorithms. *J. Algorithms* **44** 177–225.

[9] CRAMER, M. (1997). Stochastic analysis of the merge–sort algorithm. *Random Structures Algorithms* **11** 81–96.

[10] CRAMER, M. and RÜSCHENDORF, L. (1995). Analysis of recursive algorithms by the contraction method. In *Athens Conference on Applied Probability and Time Series Analysis* 18–33. Springer, New York.

[11] DEVROYE, L. (1991). Limit laws for local counters in random binary search trees. *Random Structures Algorithms* **2** 303–315.

[12] DEVROYE, L. (2003). Limit laws for sums of functions of subtrees of random binary search trees. *SIAM J. Comput.* To appear.

[13] DOBROW, R. P. and FILL, J. A. (1999). Total path length for random recursive trees. *Combin. Probab. Comput.* **8** 317–333.

[14] FILL, J. A. (1996). On the distribution of binary search trees under the random permutation model. *Random Structures Algorithms* **8** 1–25.

[15] FLAJOLET, P. and GOLIN, M. (1994). Mellin transforms and asymptotics. The mergesort recurrence. *Acta Inform.* **31** 673–696.

[16] FLAJOLET, P., GOURDON, X. and MARTÍNEZ, C. (1997). Patterns in random binary search trees. *Random Structures Algorithms* **11** 223–244.

[17] GEIGER, J. (2000). A new proof of Yaglom's exponential limit law. In *Mathematics and Computer Science* 245–249. Birkhäuser, Basel.

[18] GRÜBEL, R. and RÖSLER, U. (1996). Asymptotic distribution theory for Hoare's selection algorithm. *Adv. in Appl. Probab.* **28** 252–269.

[19] HUBALEK, F., HWANG, H.-K., LEW, W., MAHMOUD, H. M. and PRODINGER, H. (2002). A multivariate view of random bucket digital search trees. *J. Algorithms* **44** 121–158.

[20] HWANG, H.-K. (1996). Asymptotic expansions of the mergesort recurrences. *Acta Inform.* **35** 911–919.

[21] HWANG, H.-K. (1998). Limit theorems for mergesort. *Random Structures Algorithms* **8** 319–336.

[22] HWANG, H.-K. (2001). Lectures on asymptotic analysis given at McGill Univ., Montreal.

[23] HWANG, H.-K. and NEININGER, R. (2002). Phase change of limit laws in the quicksort recurrence under varying toll functions. *SIAM J. Comput.* **31** 1687–1722.

[24] HWANG, H.-K. and TSAI, T.-H. (2002). Quickselect and Dickman function. *Combin. Probab. Comput.* **11** 353–371.

[25] JACQUET, P. and RÉGNIER, M. (1988). Normal limiting distribution of the size and the external path length of tries. Technical Report RR-0827, INRIA-Rocquencourt.

[26] JACQUET, P. and RÉGNIER, M. (1988). Normal limiting distribution of the size of tries. In *Performance'87* 209–223. North-Holland, Amsterdam.

[27] JACQUET, P. and SZPANKOWSKI, W. (1995). Asymptotic behavior of the Lempel–Ziv parsing scheme and digital search trees. *Theoret. Comput. Sci.* **144** 161–197.

[28] KIRSCHENHOFER, P., PRODINGER, H. and SZPANKOWSKI, W. (1989). On the balance property of Patricia tries: External path length viewpoint. *Theoret. Comput. Sci.* **68** 1–17.

[29] KIRSCHENHOFER, P., PRODINGER, H. and SZPANKOWSKI, W. (1989). On the variance of the external path length in a symmetric digital trie. *Discrete Appl. Math.* **25** 129–143.

[30] KIRSCHENHOFER, P., PRODINGER, H. and SZPANKOWSKI, W. (1994). Digital search trees again revisited: The internal path length perspective. *SIAM J. Comput.* **23** 598–616.

[31] KNUTH, D. E. (1973). *The Art of Computer Programming* **3**. Addison–Wesley, Reading, MA.

[32] KODAJ, B. and MÓRI, T. F. (1997). On the number of comparisons in Hoare's algorithm "FIND." *Studia Sci. Math. Hungar.* **33** 185–207.

[33] LEW, W. and MAHMOUD, H. M. (1994). The joint distribution of elastic buckets in multiway search trees. *SIAM J. Comput.* **23** 1050–1074.

[34] MAHMOUD, H. M. (1992). *Evolution of Random Search Trees*. Wiley, New York.

[35] MAHMOUD, H. M. (1995). The joint distribution of the three types of nodes in uniform binary trees. *Algorithmica* **13** 313–323.

[36] MAHMOUD, H. M. (2000). *Sorting. A Distribution Theory*. Wiley, New York.

[37] MAHMOUD, H. M. and PITTEL, B. (1988). Asymptotic joint normality of outdegrees of nodes in random recursive trees. *Random Structures Algorithms* **3** 255–266.

[38] MAHMOUD, H. M. and PITTEL, B. (1989). Analysis of the space of search trees under the random insertion algorithm. *Journal of Algorithms* **10** 52–75.

[39] MAHMOUD, H. M. and SMYTHE, R. T. (1991). On the distribution of leaves in rooted subtrees of recursive trees. *Ann. Appl. Probab.* **1** 406–418.

[40] MAHMOUD, H. M. and SMYTHE, R. T. (1992). Asymptotic joint normality of outdegrees of nodes in random recursive trees. *Random Structures Algorithms* **3** 255–266.

[41] MAHMOUD, H. M., MODARRES, R. and SMYTHE, R. T. (1995). Analysis of quickselect: An algorithm for order statistics. *RAIRO Inform. Théor. Appl.* **29** 255–276.

[42] MAHMOUD, H. M., SMYTHE, R. T. and SZYMAŃSKI, J. (1993). On the structure of random plane-oriented recursive trees and their branches. *Random Structures Algorithms* **4** 151–176.

[43] NEININGER, R. (2001). On a multivariate contraction method for random recursive structures with applications to Quicksort. *Random Structures Algorithms* **19** 498–524.

[44] NEININGER, R. and RÜSCHENDORF, L. (1999). On the internal path length of d-dimensional quad trees. *Random Structures Algorithms* **15** 25–41.

AAP ims v.2003/08/05 Prn:1/09/2003; 11:54 F:aap165.tex; (Aiste) p. 41

[45] NEININGER, R. and RÜSCHENDORF, L. (2002). Rates of convergence for Quicksort. *J. Algorithms* **44** 52–62.

[46] PITTEL, B. (1999). Normal convergence problem? Two moments and a recurrence may be the clues. *Ann. Appl. Probab.* **9** 1260–1302.

[47] RACHEV, S. T. (1991). *Probability Metrics and the Stability of Stochastic Models*. Wiley, New York.

[48] RACHEV, S. T. and RÜSCHENDORF, L. (1994). On the rate of convergence in the CLT with respect to the Kantorovich metric. In *Probability in Banach Spaces* **9** 193–207. Birkhäuser, Boston.

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

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

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

[52] RÖSLER, U. (2001). On the analysis of stochastic divide and conquer algorithms. *Algorithmica* **29** 238–261.

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

[54] SCHACHINGER, W. (2000). Asymptotic normality of recursive algorithms via martingale difference arrays. Preprint.

[55] SENATOV, V. V. (1980). Some uniform estimates of the convergence rate in the multidimensional central limit theorem. *Theor. Probab. Appl.* **25** 745–759.

[56] SMYTHE, R. T. and MAHMOUD, H. M. (1994). A survey of recursive trees. *Teor. Īmovīr. ta Mat. Statist.* **51** 1–29.

[57] SZPANKOWSKI, W. (2001). *Average Case Analysis of Algorithms on Sequences*. Wiley, New York.

[58] TENENBAUM, G. (1995). *Introduction to Analytic and Probabilistic Number Theory*. (C. B. Thomas, transl.). Cambridge Univ. Press.

[59] ZOLOTAREV, V. M. (1976). Approximation of the distributions of sums of independent random variables with values in infinite-dimensional spaces. *Theor. Probab. Appl.* **21** 721–737.

[60] ZOLOTAREV, V. M. (1977). Ideal metrics in the problem of approximating distributions of sums of independent random variables. *Theory Probab. Appl.* **22** 433–449.

SCHOOL OF COMPUTER SCIENCE
MCGILL UNIVERSITY
3480 UNIVERSITY STREET
MONTREAL
CANADA H3A 2K6
E-MAIL: neiningr@jeff.cs.mcgill.ca

INSTITUT FÜR MATHEMATISCHE STOCHASTIK
UNIVERSITÄT FREIBURG
ECKERSTRASSE 1
79104 FREIBURG
GERMANY
E-MAIL: ruschen@stochastik.uni-freiburg.de