Analysis of algorithms by the contraction method: additive and max-recursive sequences

Ralph Neininger[*]
Department of Mathematics
J. W. Goethe University
Robert-Mayer-Str. 10
60325 Frankfurt a. M.
Germany

Ludger Rüschendorf
Department of Mathematics
University of Freiburg
Eckerstr. 1
79104 Freiburg
Germany

September 30, 2003

Abstract

In the first part of this paper we give an introduction to the contraction method for the analysis of additive recursive sequences of divide and conquer type. Recently some general limit theorems have been obtained by this method based on a general transfer theorem. This allows to conclude from the recursive structure and the asymptotics of first moment(s) the limiting distribution. In the second part we extend the contraction method to max-recursive sequences. We obtain a general existence and uniqueness result for solutions of stochastic equations including maxima and sum terms. We finally derive a general limit theorem for max-recursive sequences of the divide and conquer type.

Keywords: Analysis of algorithms, parallel algorithms, limit laws, recurrence, probability metric, limit law for maxima.

1 Introduction to the contraction method

The analysis of algorithms is a rapidly expanding area of analysis. Since the introduction of the average case analysis in Knuth (1973) there have been developed several approaches to limit laws for various parameters of recursive algorithms, random trees and combinatorial structures. The contraction

[*]Research supported by an Emmy Noether Fellowship of the DFG.

method is a probabilistic technique of analysis with a broad range of applications which supplements the analytic techniques (generating functions) and the other probabilistic techniques like martingales or branching processes. The contraction technique was first introduced for the analysis of Quicksort in Rösler (1991) and further on developed independently in Rösler (1992) and Rachev and Rüschendorf (1995), as well as in Neininger and Rüschendorf (2003a, 2003b), see also the survey article Rösler and Rüschendorf (2001). It has been successfully applied to a broad range of algorithms (see Neininger (1999, 2001), Rösler (2001), Hwang and Neininger (2002), and Neininger and Rüschendorf (2003a)).

The idea of the contraction method is to reduce the analysis of an algorithm to the study of contraction properties of transformations associated to the algorithm, and then to use some variant of the Banach fixed point theorem. We explain some general aspects of this method at the example of the Quicksort algorithm.

Let L_n denote the number of comparisons of the Quicksort algorithm to sort n randomly permuted real numbers, see, e.g. Mahmoud (2000). Then

$$\ell_n = EL_n = 2n\log n + (2\gamma - 4)n + O(\ln n), \quad \gamma \text{ the Euler constant} \tag{1.1}$$

$$\text{and} \quad \sigma_n^2 = \mathrm{Var}(L_n) = \left(7 - \frac{2\pi^2}{3}\right) n^2 + O(n\ln n). \tag{1.2}$$

Régnier (1989) established that $Z_n = \frac{L_n - \ell_n}{n+1}$ is a L_2-bounded martingale and, therefore, a.s. convergence to some rv's Z holds:

$$Z_n \to Z \text{ a.s.} \tag{1.3}$$

In order to determine the distribution of Z it is useful to consider the recursive structure of L_n. By an obvious simple argument we have

$$L_n \stackrel{d}{=} L_{I_n} + \bar{L}_{n-1-I_n} + n - 1, \tag{1.4}$$

where $\bar{L}_k$ are independent copies of L_k, I_n is uniformly distributed on $\{0, \dots, n-1\}$ and independent of (L_k), (I_k). I_n is the size of the subgroup which is smaller than the first pivot element chosen by the Quicksort algorithm. After normalization, $Y_n = \frac{L_n - \ell_n}{n}$ satisfies the recursion

$$Y_n \stackrel{d}{=} \frac{I_n}{n} Y_{I_n} + \frac{n-1-I_n}{n} \bar{Y}_{n-1-I_n} + c_n(I_n), \quad \text{with } c_n(j) = \frac{n-1}{n} + \frac{1}{n}(\ell_j + \ell_{n-1-j} - \ell_n), \tag{1.5}$$

where $(\bar{Y}_n)$ is a distributional copy of (Y_n). With $c(x) := 2x\log x + 2(1-x)\log(1-x) + 1$ it is easy to see that: $\sup_{x\in(0,1]} |c_n([nx]) - c(x)| \le \frac{4}{n}\log n + O\left(\frac{1}{n}\right)$. Choosing w.l.g. some version of I_n such that

$$\frac{I_n}{n} \to \tau \text{ a.s.,} \tag{1.6}$$

where τ is uniformly distributed on $[0,1]$ we obtain from (1.3), (1.5) that the limit Y of Y_n exists a.s. and satisfies the *limit equation*:

$$Y \stackrel{d}{=} \tau Y + (1-\tau)\bar{Y} + c(\tau). \tag{1.7}$$

There exists exactly one solution of the limiting equation (1.7) in the class $\mathcal{M}_2(0)$ of probability measures on $\mathbb{R}$ with mean zero and with finite variance. To that purpose we define the transformation $T : \mathcal{M}_2(0) \to \mathcal{M}_2(0)$ by

$$T(P) = \mathcal{L}(\tau Y + (1-\tau)\bar{Y} + c(\tau)) \tag{1.8}$$

if $P = \mathcal{L}(Y)$. The operator T is closely related to the Quicksort algorithm. It is an asymptotic approximation of the recursion operator in (1.5). T is a contraction operator w.r.t. the minimal ℓ_2-metric defined for probability measures P, Q by

$$\ell_2(P,Q) = \inf\left\{\left(E(X-Y)^2\right)^{1/2}; X \stackrel{d}{=} P, Y \stackrel{d}{=} Q\right\}; \tag{1.9}$$

$$\ell_2(TP,TQ) \leq \sqrt{\tfrac{2}{3}}\,\ell_2(P,Q). \tag{1.10}$$

For the proof of (1.10) let $X_i \stackrel{d}{=} P$, $Y_i \stackrel{d}{=} Q$, $i = 1,2$ be i.i.d. copies of P, Q such that $E(X_i - Y_i)^2 = \ell_2^2(P,Q)$. Then

$$\begin{aligned} \ell_2^2(TP,TQ) &\leq E(\tau X_1 + (1-\tau)X_2 + c(\tau) - [\tau Y_1 + (1-\tau)Y_2 + c(\tau)])^2 \\ &= E[\tau^2(X_1-Y_1)^2 + (1-\tau)^2(X_2-Y_2)^2] \\ &= 2E\tau^2\ell_2^2(P,Q) = \tfrac{2}{3}\ell_2^2(P,Q). \end{aligned} \tag{1.11}$$

Thus by Banach's fixed point theorem the limiting equation (1.7) has a unique solution in $\mathcal{M}_2(0)$. The uniqueness of the solution of the limiting equation (1.7) implies that Y_n converges in distribution to Y

$$Y_n \xrightarrow{\mathcal{D}} Y, \tag{1.12}$$

where Y is the unique solution of the limiting fixed point equation (1.7) which is called the *Quicksort-distribution*.

The contraction method allows to extend this type of convergence argument to a general class of recursive algorithms. It simultaneously also allows to prove the essential convergence step (1.3) in the argument above without reference to a martingale argument as above. This is of considerable importance since a related martingale structure has been found only in few examples of recursive algorithms.

In section two of this paper we review some recent developments of the contraction method for additive recursive sequences of divide and conquer type. In the final part of the paper we develop some new tools which are basic for an extension of the contraction method to recursive sequences of divide and conquer type which are based on maxima (like parallel search algorithms). Similar as the additive recursive algorithms are 'relatives' of the classical central limit theorem for sums the max-based recursive algorithms can be considered as relatives of the classical central limit theorem for maxima.

2 Limit theorem for divide and conquer algorithms

In the recent paper Neininger and Rüschendorf (2003a) a general limit theorem has been derived for recursive algorithms and combinatorial structures by means of the contraction method. In comparison to the introductory example in section 1 the main progress in that paper is a general transfer theorem which allows to establish a limit law on the basis of the recursive structure and using the asymptotics of the first moment(s) of the sequence. Thus the strong information by the martingale structure can be replaced by the information on first moment(s). For a lot of examples of algorithms this information on moments is available by highly developed analytical methods.

A common type of univariate recursions (Y_n) of the divide and conquer type is of the following form:

$$Y_n \stackrel{d}{=} \sum_{r=1}^{K} Y^{(r)}_{I_r^{(n)}} + b_n, \quad n \geq n_0 \tag{2.1}$$

with $(Y_n^{(1)}), \ldots, (Y_n^{(K)})$, $(I^{(n)}, b_n)$ independent $Y_j^{(r)} \stackrel{d}{=} Y_j$, $P(I_r^{(n)} = n) \to 0$ and $\mathrm{Var}(Y_n) > 0$ for $n \geq n_1$. $I_r^{(n)}$ describe subgroup sizes of the divide and conquer algorithm and b_n is a toll function for the splitting into and merging of K smaller problems.

The analysis of the asymptotics of (Y_n) is based on the Zolotarev metric ζ_s on $\mathcal{M}$ the set of all probability measures on $\mathbb{R}^1$ defined by (see Zolotarev (1997))

$$\zeta_s(P, Q) = \sup_{f \in \mathcal{F}_s} |Ef(X) - Ef(Y)|, \tag{2.2}$$

where $\mathcal{L}(X) = P$, $\mathcal{L}(Y) = Q$, and $\mathcal{F}_s = \{f \in C^{(m)}(\mathbb{R}); \|f^{(m)}(x) - f^{(m)}(y)\| \leq |x-y|^\alpha\}$ with $s = m + \alpha, 0 < \alpha \leq 1, m \in \mathbb{N}_0$. Finiteness of $\zeta_s(\mathcal{L}(X), \mathcal{L}(Y))$ is guaranteed if X, Y have identical moments of orders $1, \ldots, m$ and finite absolute moments of order s. Since ζ_s is of main interest for $s \leq 3$, we introduce the following subspaces of $\mathcal{M}_s$ – the set of measures with finite s-th moments – to obtain finiteness of ζ_s. Define $\mathcal{M}_s(\mu)(\mathcal{M}_s(\mu, \sigma^2))$ for $1 < s \leq 2$ $(2 < s \leq 3)$ to be the elements in $\mathcal{M}_s$ with fixed first moment μ (resp. also fixed variance σ^2) and define $\mathcal{M}_s^*$ to be identical to $\mathcal{M}_s$ for $0 < s \leq 1$, to $\mathcal{M}_s(\mu)$ for $1 < s \leq 2$ and to $\mathcal{M}_s(\mu, \sigma^2)$ for $2 < s \leq 3$, where μ, σ^2 are fixed in the context.

An important property of ζ_s for the contraction method is that

$$\zeta_s(X+Z, Y+Z) \le \zeta_s(X,Y) \text{ and } \zeta_s(cX, cY) = |c|^s \zeta_s(X,Y) \tag{2.3}$$

for all Z independent of X, Y and $c \in \mathbb{R}\backslash\{0\}$, whenever these distances are finite. ζ_s convergence implies weak convergence.

For the limiting analysis of Y_n we need a stabilization condition for the recursive structure and a contraction condition for the limiting fixed-point equation; in more detail: Assume for functions $f, g : \mathbb{N}_0 \to \mathbb{R}_0^+$ with $g(n) > 0$ for $n \ge n_1$ we have the following *stabilization condition in* L_s

$$\begin{aligned} &\left(\frac{g(I_r^{(n)})}{g(n)}\right)^{1/2} \to A_r^*, \quad r = 1, \dots, K \quad \text{and} \\ &\frac{1}{g^{1/2}(n)}\left(b_n - f(n) + \sum_{r=1}^K f(I_r^{(n)})\right) \to b^*, \end{aligned} \tag{2.4}$$

as well as the *contraction condition*

$$E\sum_{r=1}^K |A_r^*|^s < 1. \tag{2.5}$$

Then the following limit theorem is obtained by the contraction method (see Neininger and Rüschendorf (2003a, Theorem 5.1)).

Theorem 2.1 *Let* (Y_n) *be* s*-integrable and satisfy the recursive equation* (2.1) *and let* f, g *satisfy the stabilization condition* (2.4) *and the contraction condition* (2.5) *for some* $0 < s \le 3$. *Furthermore, in case* $1 < s \le 3$ *assume the* moment convergence condition

$$\begin{aligned} EY_n &= f(n) + o(g^{1/2}(n)) \quad \text{if } 1 < s \le 2 \quad \text{and} \\ EY_n &= f(n) + o(g^{1/2}(n)), \mathrm{Var}(Y_n) = g(n) + o(g(n)) \quad \text{if } 2 < s \le 3. \end{aligned} \tag{2.6}$$

Then $\dfrac{Y_n - f(n)}{g^{1/2}(n)} \xrightarrow{\mathcal{D}} X$, *where* X *is the unique fixed-point of*

$$X \stackrel{d}{=} \sum_{r=1}^K A_r^* X^{(r)} + b^* \tag{2.7}$$

in $\mathcal{M}_s^*$ *(with* $\mu = 0, \sigma^2 = 1$*), where* $(A_1^*, \dots, A_K^*, b^*)$, $X^{(1)}, \dots, X^{(K)}$ *are independent,* $X^{(r)} \stackrel{d}{=} X$.

Remark 2.2 *a) For the proof of Theorem 2.1 one gets from the moment convergence condition* (2.6) *and the stabilization condition* (2.4) *the form of the limiting equation* (2.7). *The existence of a unique fixed point of*

(2.7) *follows from the contraction condition* (2.5) *by Banach's fixed point theorem. From the regularity properties of* ζ_s *in* (2.3) *we can argue that the contraction property in the limiting equation can be carried over to the recursive sequence.*

b) Note that in the case that the conditions are satisfied for $0 < s \leq 1$ *we do not need any information on the asymptotics of moments; for* $1 < s \leq 2$ *the asymptotics of the first moment is needed. The case* $0 < s \leq 1$ *arises for example for limit equations of the form*

$$X \stackrel{d}{=} \tfrac{1}{\sqrt{2}} X + \tfrac{1}{\sqrt{2}} \mathcal{N}(0,1), \tag{2.8}$$

with the standard normal distribution as unique solution, or of the form

$$W \stackrel{d}{=} UW + U, \tag{2.9}$$

U *uniformly distributed on* $(0,1)$*, with the Dickman distribution as unique solution. The Dickman distribution arises e.g. as a limit in the context of the Find algorithm. For normal limits, besides (2.8), the case* $2 < s \leq 3$ *is typical. Then typically the minimal* ℓ_p*-metrics (see (3.7)) cannot be used directly to derive normal limit laws directly.*

c) If the contraction method applies for $s = 1$ *then Theorem 2.1 applied with* ζ_1 *yields the asymptotics of the first order moment. If it applies for* $s = 2$*, then one needs asymptotics of the first moment and obtains the asymptotics of the second moment.*

d) A large class of examples for the application of Theorem 2.1 to the asymptotics of recursive algorithms has been established. Note that there are also several variants of this basic theorem (to the multivariate case, weighted recursions, random number of components, alternative contraction conditions, degenerative limits, ...). In particular one gets a classification of algorithms according to their contraction behavior. To get an impression of the range of application we give a list of established examples (without giving detailed references):

1) $\mathcal{M}_s, 0 < s \leq 1$**.** *Examples contain: FIND comparisons, number of exchange steps, Dickman, Multiple Quickselect, Bucket selection, Quicksort with error (number of inversions), leader election (flips), skip lists (size), ideals in forest poset, distances in random binary search trees (rBST), minimum spanning trees in rBST, random split tree (large toll).*

2) $\mathcal{M}_s(\mu), 1 < s \leq 2$**.** *Quicksort (comparisons, exchanges), internal path length in quad trees, m-ary search trees, median search trees and recursive trees, Wiener index in rBST and recursive trees. Yaglom's exponential limit law, random split trees for moderate toll.*

3) $\mathcal{M}_s(\mu, \sigma^2), 2 < s \leq 3$. *Quicksort (rec. calls), patterns in trees, size in m-ary search trees, size and path length in tries, digital search trees, and Patricia trees, merge sort (comparisons top-down version), vertices with outdegrees in recursive trees, random split trees with small toll.*

For these and related examples see Neininger and Rüschendorf (2003a), Hwang and Neininger (2002), and Neininger (1999, 2001).

e) In Neininger and Rüschendorf (2003a) it has been shown that one can derive from the convergence results in the Zolotarev metric several local and global limit theorems. It is also possible to obtain rate of convergence results. In Neininger and Rüschendorf (2002) it is shown that the convergence rate of the Quicksort algorithm is w.r.t. the Zolotarev metric ζ_3 *of the exact order* $\frac{\ln n}{n}$.

3 Contraction and fixed point properties with maxima

In this section we extend the analysis of algorithms defined via sums as in (2.1) to recursive algorithms including maximum and sum terms. The analysis in section 2 based on the Zolotarev metric ζ_s would go through in this case if one could find a metric μ_s which is not only regular of order s for sums as in (2.3) but also simultaneously for maxima too, i.e.

$$\begin{aligned} \mu_s(X \vee Z, Y \vee Z) &\leq \mu_s(X, Y) \quad \text{and} \\ \mu_s(cX, cY) &= |c|^s \mu_s(X, Y). \end{aligned} \tag{3.1}$$

It was however shown in Rachev and Rüschendorf (1992) that only trivial metrics may have this doubly ideal property.

For the central limit theorem for maxima the weighted Kolmogorov metric ϱ_s, defined by

$$\varrho_s(X, Y) = \sup_x |x|^s |F_X(x) - F_Y(x)| \tag{3.2}$$

is max-regular of order s for real rv's X, Y, i.e. it satisfies (3.1) and has been used for deriving limit theorems. But for recursions including also additive terms ϱ_s is not particular well suited (see Rachev and Rüschendorf (1995) and Cramer (1997)).

Limiting distributions of max-recursive sequences will typically be identified as unique solutions in some subclass of $\mathcal{M}$ of stochastic equations of the form:

$$X \stackrel{d}{=} \bigvee_{r=1}^{K} (A_r X_r + b_r), \tag{3.3}$$

where (X_r) are i.i.d. copies of X and $(A_r, b_r)_{1\le r\le K}$ are random coefficients independent of (X_r). The right hand side of (3.3) induces an operator $T : \mathcal{M} \to \mathcal{M}$ defined for $Q \in \mathcal{M}$ and $X \stackrel{d}{=} Q$ by

$$TQ = TX \stackrel{d}{=} \mathcal{L}\left(\bigvee_{r=1}^{K}(A_r X_r + b_r)\right). \tag{3.4}$$

If A_r, b_r have absolute s-th moments and $\mathcal{L}(X) \in \mathcal{M}_s$ then also TX has absolute s-th moments. So T can be considered as operator $\mathcal{M}_s \to \mathcal{M}_s$ in this case.

We next establish that the minimal ℓ_s-metric is well suited for the analysis of equations as in (3.3) although it is not doubly ideal of order s. We need the following simple lemma.

Lemma 3.1 *For all $a, b, c, d \in \mathbb{R}$ and $s > 0$ holds true:*

$$|a \vee b - c \vee d|^s \le |a - c|^s + |b - d|^s. \tag{3.5}$$

Proof: W.l.g. we assume that $b < a$, i.e., $a \vee b = a$ and $c < d$; otherwise we would have $c \vee d = c$ and so the left hand side of (3.5) is $|a - c|^s$ while the right hand side is $|a - c|^s + |b - d|^s$.

Furthermore, by symmetry we assume w.l.g. $a < d$. Then the left hand side of (3.5) is $|a - d|^s$. Noting that $|a - d|^s \le |b - d|^s$ since $b < a < d$ the result follows. $\square$

Define as usual the L_s-norm by

$$L_s(X, Y) = \begin{cases} (E|X - Y|^s)^{1/s}, & 1 \le s < \infty \\ E|X - Y|^s, & 0 < s < 1 \end{cases} \tag{3.6}$$

and the minimal L_s-metric ℓ_s by

$$\ell_s(P, Q) = \inf\{L_s(X, Y); X \stackrel{d}{=} P, Y \stackrel{d}{=} Q\}. \tag{3.7}$$

Then we obtain the following contraction property of T.

Proposition 3.2 *If (X_r) are i.i.d., $X_r \stackrel{d}{=} X_1$, (Y_r) are i.i.d., $Y_r \stackrel{d}{=} Y_1$ and A_r s-integrable, $1 \le r \le K$, then for $0 < s < \infty$*

a)
$$L_s\left(\bigvee_{r=1}^{K}(A_r X_r + b_r), \bigvee_{r=1}^{K}(A_r Y_r + b_r)\right) \le \left(E\sum_{r=1}^{K}|A_r|^s\right)^{1/s \wedge 1} L_s(X_1, Y_1). \tag{3.8}$$

b) For the operator T *defined in* (3.4) *holds*

$$\ell_s(TP, TQ) \le \left(E \sum_{r=1}^{K} |A_r|^s \right)^{1/s \wedge 1} \ell_s(P, Q). \tag{3.9}$$

Proof:

a) Consider the case $1 \le s$. Then from induction we get by Lemma 3.1

$$\begin{aligned} & L_s^s \left(\bigvee_{r=1}^{K} (A_r X_r + b_r), \bigvee_{r=1}^{K} (A_r Y_r + b_r) \right) \\ = \ & E \left| \bigvee_{r=1}^{K} (A_r X_r + b_r) - \bigvee_{r=1}^{K} (A_r Y_r + b_r) \right|^s \\ \le \ & \sum_{r=1}^{K} E|A_r(X_r - Y_r)|^s \\ = \ & \sum_{r=1}^{K} E|A_r|^s L_s^s(X_1, Y_1). \end{aligned}$$

The case $0 < s < 1$ is similar.

b) Choose (X_r) i.i.d., $X_r \stackrel{d}{=} P$ and (Y_r) i.i.d., $Y_r \stackrel{d}{=} Q$ such that $L_s(X_r, Y_r) = \ell_s(P, Q)$, $1 \le r \le K$. Then

$$\begin{aligned} \ell_s(TP, TQ) & \le L_s \left(\bigvee_{r=1}^{K} (A_r X_r + b_r), \bigvee_{r=1}^{K} (A_r X_r + b_r) \right) \\ & \le \left(\sum_{r=1}^{K} E|A_r|^s \right)^{1/s \wedge 1} L_s(X_1, Y_1) \\ & = \left(\sum_{r=1}^{K} E|A_r|^s \right)^{1/s \wedge 1} \ell_s(P, Q). \end{aligned}$$

□

Remark 3.3 *Note that inequality* (3.8) *holds more generally without any independence assumption on* (X_r, Y_r) *and thus may be used to analyze a more general class of stochastic equations. The* b_r *do not enter the contraction estimate in* (3.8), (3.9).

As a consequence we next obtain an existence and uniqueness result for the stochastic equation (3.3). For $\mu_0 \in \mathcal{M}$ define

$$\mathcal{M}_s(\mu_0) = \{\mu \in \mathcal{M}; \ell_s(\mu, \mu_0) < \infty\}, \tag{3.10}$$

the equivalence class of μ_0 w.r.t. ℓ_s. If $\mu_0 \in \mathcal{M}_s$, then $\mathcal{M}_s(\mu_0) = \mathcal{M}_s$.

Theorem 3.4 *Let for some $s > 0$ the coefficients A_r, b_r be s-integrable and $\mu_0 \in \mathcal{M}$ such that $\zeta = E\sum_{r=1}^K |A_r|^s < 1$ and $\ell_s(\mu_0, T\mu_0) < \infty$. Then the stochastic equation $X \stackrel{d}{=} \bigvee_{r=1}^K (A_r X_r + b_r)$ has a unique solution in $\mathcal{M}_s(\mu_0)$.*

Proof: Define for $n \geq 1$, $\mu_n = T\mu_{n-1} = T^n\mu_0$. Note that, by induction, $\ell_s(\mu_0, T\mu_0) < \infty$ implies $\ell_s(\mu_n, T\mu_{n+p}) < \infty$ for all $n \geq 0$, $p \geq 1$. Then by Proposition 3.2 using the triangle inequality for ℓ_s we obtain

$$\begin{aligned}\ell_s(\mu_n, \mu_{n+p}) &\leq \sum_{i=0}^{p-1} \ell_s(\mu_{n+i}, \mu_{n+i+1}) \\ &\leq \ell_s(\mu_0, \mu_1) \sum_{i=0}^{p-1} \zeta^{n+i} \leq \ell_s(\mu_0, \mu_1) \frac{\zeta^n}{1-\zeta} \\ &\to 0 \quad \text{as } \ell_s(\mu_0, \mu_1) < \infty.\end{aligned}$$

Therefore, (μ_n) is a Cauchy-sequence in the complete metric space $(\mathcal{M}_s(\mu_0), \ell_s)$. Any limiting point is a fixed point of T by Banach's fixed point theorem. For the uniqueness let $\mu, \nu \in \mathcal{M}_s(\mu_0)$ be fixed points of T. Then $\ell_s(T\mu, T\nu) \leq \zeta^{1/s \wedge 1} \ell_s(\mu, \nu)$ and thus $\ell_s(\mu, \nu) = 0$ and $\mu = \nu$. □

Remark 3.5 *a) Jagers and Rösler (2002) recently obtained a general existence result for equations of the form $X \stackrel{d}{=} \vee_r A_r X_r$ by relating them to solutions of the additive form $W \stackrel{d}{=} \sum_r A_r^\alpha W_r$. This additive equation has been well studied.*

b) If $\mu_0 \in \mathcal{M}_s$ then the condition $\ell_s(\mu_0, T\mu_0) < \infty$ is fulfilled. So under the contraction condition $\zeta < 1$ there exists a unique fixed point of T in $\mathcal{M}_s$. But there may be further fixed points not in $\mathcal{M}_s$ but in some $\mathcal{M}_s(\mu_0)$ without finite absolute moments of order s. So, for example the stochastic equation

$$X \stackrel{d}{=} \tfrac{1}{2}X_1 \vee \tfrac{1}{2}X_2$$

has the (trivial) solution $X = 0$ which is in $\mathcal{M}_s$. The contraction factor is $\zeta = (\frac{1}{2})^{s-1}$ w.r.t. ℓ_s which is smaller than 1 for any $s > 1$. The extreme value distribution with distribution function $F(x) = e^{-x^{-1}}, x \geq 0$ is a further (nontrivial) fixed point of this equation without finite first moment. In fact a basic result of extreme value theory says that any nondegenerate max-stable distribution is one of the three classical types of extreme value distributions (Gumbel, Weibull, Fréchet). Recall that a distribution function G is called max-stable if for all $n \in \mathbb{N}$ there exist $a_n > 0$, $b_n \in \mathbb{R}$ such that $G^n(\frac{x}{a_n} + b_n) = G(x), x \in \mathbb{R}$ i.e., a random variable $X \stackrel{d}{=} G$ satisfies the stochastic equations of the form

$$X \stackrel{d}{=} \bigvee_{r=1}^n (a_n X_r - b_n), \quad n \in \mathbb{N}. \tag{3.11}$$

This characterization yields uniqueness without any moment considerations but uses a system of stochastic equations instead of only one equation as above.

c) Central limit theorem. *As consequence of Propositions 3.2 and Theorem 3.4 one gets an easy proof of the central limit theorem for maxima. (For a general discussion of this topic see Zolotarev (1997) and Rachev (1991)).* *Let $F(x) = F_{Y_1}(x) = e^{-x^{-\alpha}}$, $x \geq 0$, be an extreme value distribution of first type and let (X_r) be an i.i.d. sequence with tail condition $\ell_s(X_1, Y_1) < \infty$ for some $s > \alpha$. Then for the maxima sequence $M_n := \max\{X_1, \ldots, X_n\}$ holds:*

$$\ell_s(n^{-1/\alpha} M_n, Y_1) \to 0. \tag{3.12}$$

For the proof note that Y_1 is a solution of the stochastic equation

$$Y_1 \stackrel{d}{=} n^{-1/\alpha} \bigvee_{r=1}^{n} Y_r. \tag{3.13}$$

This implies by Proposition 3.2

$$\begin{aligned} \ell_s(n^{-1/\alpha} M_n, Y_1) &= \ell_s\Big(n^{-1/\alpha} M_n, n^{-1/\alpha} \bigvee_{r=1}^{n} Y_r\Big) \\ &\leq (n \cdot n^{-s/\alpha})^{1/s \wedge 1} \ell_s(X_1, Y_1) \\ &= (n^{1-s/\alpha})^{1/s \wedge 1} \ell_s(X_1, Y_1) \ \to \ 0 \ \text{ as } s > \alpha. \end{aligned}$$

For $s \to \infty$ the rate approaches the optimal rate $n^{-1/\alpha}$.

d) Transformation of the fixed point equation. *The fixed point equation*

$$X \stackrel{d}{=} \bigvee_{r=1}^{K} (A_r X_r + b_r) \tag{3.14}$$

can be transformed in various ways. Let, e.g., $Y = \exp(\lambda X)$, then (3.14) *transforms to*

$$Y = \bigvee_{r=1}^{K} e^{\lambda b_r} Y_r^{A_r}, \tag{3.15}$$

in particular, for $A_r = 1, \lambda = 1$,

$$Y = \bigvee_{r=1}^{K} e^{b_r} Y_r. \tag{3.16}$$

For $Z = Y^{(\alpha)} = |X|^\alpha \operatorname{sgn}(X)$ *and* $W = \frac{1}{X^{(\alpha)}}$ (3.14) *transforms similarly to further equivalent forms, in particular in the case* $b_r = b$. *In this way all possible extreme value distributions can be reduced to the case of extreme value distributions of type 1 considered in Remark 3.5a (see also Zolotarev (1997)). Consider as example the stochastic equation:*

$$X \stackrel{d}{=} \bigvee_{r=1}^{2} (X_r - \ln 2). \tag{3.17}$$

This equation cannot be directly handled w.r.t. the ℓ_s*-metric. Using* $Y = \exp(X)$ *equation* (3.17) *transforms to*

$$Y \stackrel{d}{=} \tfrac{1}{2} Y_1 \vee \tfrac{1}{2} Y_2. \tag{3.18}$$

A solution is the extreme value distribution $F(x) = e^{-x^{-1}}, x \geq 0$. *The operator* T *corresponding to* (3.18) *has contraction factor* $\zeta = (\frac{1}{2})^{s-1}$ *with respect to* ℓ_s. *So for any* $s > 1$ F *is a unique fixed point in* $\mathcal{M}_s(F)$ *and the central limit theorem holds for* (Z_r) *with tail condition* $\ell_s(Z_r, Y) < \infty$, *i.e.,* $(1/n^{1/\alpha}) \vee_{r=1}^n Z_r \stackrel{d}{\to} Y$, *where* $Y \stackrel{d}{=} F$ *equivalently* $\vee_{r=1}^n W_r - (1/\alpha) \ln n \stackrel{d}{\to} X$, *where* X *is the corresponding solution of* (3.17), $Y = \exp(X)$ *and* $Z_r = \exp(W_r)$.

4 Max-recursive algorithms of divide and conquer type

We consider a general class of parameters of max-recursive algorithms of divide and conquer type:

$$Y_n \stackrel{d}{=} \bigvee_{r=1}^{K} \left(A_r(n) Y^{(r)}_{I_r^{(n)}} + b_r(n) \right), \quad n \geq n_0 \tag{4.1}$$

where $I_r^{(n)}$ are subgroup sizes, $b_r(n)$ random toll terms, $A_r(n)$ random weighting terms and $(Y_n^{(r)})$ are independent copies of (Y_n) independent also from $(A_r(n), b_r(n), I^{(n)})$.

With normalizing constants ℓ_n, σ_n let X_n denote the normalized sequence $X_n = \frac{Y_n - \ell_n}{\sigma_n}$. Then

$$\begin{aligned} X_n &= \bigvee_{r=1}^{K} \left(\frac{A_r(n) Y^{(r)}_{I_r^{(n)}}}{\sigma_n} + \frac{b_r(n)}{\sigma_n} \right) - \frac{\ell_n}{\sigma_n} \\ &= \bigvee_{r=1}^{K} \left(\left(A_r(n) \frac{\sigma_{I_r^{(n)}}}{\sigma_n} \right) X^{(r)}_{I_r^{(n)}} + \frac{1}{\sigma_n} \left(A_r(n) \ell_{I_r^{(n)}} + b_r(n) - \frac{\ell_n}{\sigma_n} \right) \right) \\ &= \bigvee_{r=1}^{K} \left(A_r^{(n)} X^{(r)}_{I_r^{(n)}} + b_r^{(n)} \right), \end{aligned} \tag{4.2}$$

where $b_r^{(n)} = \frac{1}{\sigma_n}(b_r(n) - \ell_n + A_r(n)\ell_{I_r^{(n)}})$ and $A_r^{(n)} = A_r(n)\frac{\sigma_{I_r^{(n)}}}{\sigma_n}$. Thus we obtain again the form (4.1) with modified coefficients.

As in section 2 we need a stabilization condition in L_s:

$$\left(A_1^{(n)}, \dots, A_K^{(n)}, b_1^{(n)}, \dots, b_K^{(n)}\right) \to (A_1^*, \dots, A_K^*, b_1^* \dots, b_K^*)\,. \tag{4.3}$$

Thus we obtain as limiting equation a stochastic equation of the form considered in section 3:

$$X \stackrel{d}{=} \bigvee_{r=1}^{K} (A_r^* X_r + b_r^*). \tag{4.4}$$

For existence and uniqueness of solutions of (4.4) we need the contraction condition:

$$E \sum_{r=1}^{K} |A_r^*|^s < 1. \tag{4.5}$$

For the application of the contraction method let T be the limiting operator,

$$TX \stackrel{d}{=} \bigvee_{r=1}^{K} (A_r^* X_r + b_r^*). \tag{4.6}$$

Then $\ell_s(X, TX) < \infty$ if X, A_r^*, b_r^* have finite absolute s-th moments, X a starting vector. More generally finiteness also holds under some tail equivalence conditions for X and the corresponding TX. Finally, to deal with the initial conditions we need the nondegeneracy condition: For any $\ell \in \mathbb{N}$ and $r = 1, \dots, K$ holds

$$E\left[1_{\{I_r^{(n)} \le \ell\} \cup \{I_r^{(n)} = n\}} |A_r^{(n)}|^s\right] \to 0. \tag{4.7}$$

Our main result gives a limit theorem for X_n.

Theorem 4.1 (Limit theorem for max-recursive sequences) *Let* (X_n) *be a max-recursive,* s*-integrable sequence as in* (4.1) *and assume the stabilization condition* (4.4)*, the contraction condition* (4.5)*, and the nondegeneracy condition* (4.7) *for some* $s > 0$*. Then* (X_n) *converges in distribution to a limit* X^**,* $\ell_s(X_n, X^*) \to 0$*.* X^* *is the unique solution of the limiting equation*

$$X^* \stackrel{d}{=} \bigvee_{r=1}^{K} (A_r^* X_r^* + b_r^*) \ \text{ in } \mathcal{M}_s\,. \tag{4.8}$$

Proof: By our assumption we have $E|A_r^*|^s, E|b_r^*|^s < \infty$ and so for any s-integrable X_0 holds $\ell_s(X_0, TX_0) < \infty$. Define the accompanying sequence

$$W_n := \bigvee_{r=1}^{K} \left(A_r^{(n)} X_r^* + b_r^{(n)}\right), \tag{4.9}$$

where $X_1^*, \ldots, X_K^*$ are i.i.d. copies of the solution X^* of the limiting equation, which exists and is unique by the contraction condition and Theorem 3.4. Then

$$\ell_s(X_n, X^*) \le \ell_s(X_n, W_n) + \ell_s(W_n, X^*). \tag{4.10}$$

From the stabilization condition we first show that

$$\ell_s(W_n, X^*) \to 0. \tag{4.11}$$

Subsequently, we assume $s \ge 1$. For the proof of (4.11) we use the stabilization condition (4.3)

$$\begin{aligned}
\ell_s(W_n, X^*) &= \ell_s\left(\bigvee_{r=1}^{K} \left(A_r^{(n)} X_r^* + b_r^{(n)}\right), \bigvee_{r=1}^{K} \left(A_r^* X_r^* + b_r^*\right)\right) \\
&\le \left(\sum_{r=1}^{K} L_s^s\left(A_r^{(n)} X_r^* + b_r^{(n)}, A_r^* X_r^* + b_r^*\right)\right)^{1/s} \\
&\le \left(\sum_{r=1}^{K} \left[L_s\left(A_r^{(n)} X_r^*, A_r^* X_r^*\right) + L_s\left(b_r^{(n)}, b_r^*\right)\right]^s\right)^{1/s} \\
&\le \left(\sum_{r=1}^{K} \left[L_s\left(A_r^{(n)}, A_r^*\right)\left(E|X^*|^s\right)^{1/s} + L_s(b_r^{(n)}, b_r^*)\right]^s\right)^{1/s} \\
&\to 0.
\end{aligned} \tag{4.12}$$

Next let Υ_n denote the joint distribution of $(A_1^{(n)}, \ldots, A_K^{(n)}, I^{(n)}, b_1^{(n)}, \ldots, b_K^{(n)})$ and let $(\alpha, j, \beta) = (\alpha_1, \ldots, \alpha_K, j_1, \ldots, j_K, \beta_1, \ldots, \beta_K)$. Then we obtain by a conditioning argument for $s \geq 1$

$$\begin{aligned}
\ell_s^s(X_n, W_n) &= \ell_s^s\left(\bigvee_{r=1}^K \left(A_r^{(n)} X_{I_r^{(n)}}^{(r)} + b_r^{(n)}\right), \bigvee_{r=1}^K \left(A_r^{(n)} X_r^* + b_r^{(n)}\right)\right) \qquad (4.13)\\
&\leq \int L_s^s\left(\bigvee_{r=1}^K \left(\alpha_r X_{j_r}^{(r)} + \beta_r\right), \bigvee_{r=1}^K \left(\alpha_r X_r^* + \beta_r\right)\right) d\Upsilon_n(\alpha, j, \beta)\\
&\leq \sum_{r=1}^K \int L_s^s\left(\alpha_r X_{j_r}^{(r)}, \alpha_r X_r^*\right) d\Upsilon_n(\alpha, j, \beta)\\
&= \sum_{r=1}^K \int |\alpha_r|^s \ell_s^s(X_{j_r}, X^*) d\Upsilon_n(\alpha, j, \beta)\\
&\leq p_n^s \ell_s^s(X_n, X^*) + \sum_{r=1}^K \int \mathbf{1}_{\{j_r < n\}} |\alpha_r|^s \ell_s^s(X_{j_r}, X^*) d\Upsilon_n(\alpha, j, \beta).
\end{aligned}$$

where $p_n = \left(E \sum_{r=1}^K \mathbf{1}_{\{I_r^{(n)} = n\}} |A_r^{(n)}|^s\right)^{1/s}$. With the inequality $(a+b)^{1/s} \leq a^{1/s} + b^{1/s}$ for all $a, b > 0$ and $s \geq 1$ we obtain with (4.10), (4.12) and (4.13)

$$\ell_s(X_n, X^*) \leq \frac{1}{1 - p_n}\left(\left(\sum_{r=1}^K E|A_r^{(n)}|^s\right)^{1/s} \max_{0 \leq j \leq n-1} \ell_s(X_j, X^*) + o(1)\right). \qquad (4.14)$$

Since, by (4.3), (4.5) and (4.7), we have $\left(\sum_{r=1}^K E|A_r^{(n)}|^s\right)^{1/s} \to \zeta < 1$ and $p_n \to 0$ as $n \to \infty$ it follows that the sequence $(\ell_s(X_n, X^*))_{n \geq 0}$ is bounded. Denote $\bar{\eta} := \sup_{n \geq 0} \ell_s(X_n, X^*)$ and $\eta := \limsup_{n \to \infty} \ell_s(X_n, X^*)$. Now we conclude that $\ell_s(X_n, X^*) \to 0$ as $n \to \infty$ by a standard argument. For all $\varepsilon > 0$ there is an $\ell \in \mathbb{N}$ such that $\ell_s(X_n, X^*) \leq \eta + \varepsilon$ for all $n \geq \ell$. Then with (4.13), (4.10), and (4.12) we obtain

$$\begin{aligned}
\ell_s(X_n, X^*) &\leq \frac{1}{1 - p_n}\left(\sum_{r=1}^K \int \mathbf{1}_{\{j_r \leq \ell\}} |\alpha_r|^s \ell_s^s(X_{j_r}, X^*) d\Upsilon_n(\alpha, j, \beta)\right.\\
&\qquad \left. + \sum_{r=1}^K \int \mathbf{1}_{\{j_r > \ell\}} |\alpha_r|^s \ell_s^s(X_{j_r}, X^*) d\Upsilon_n(\alpha, j, \beta) + o(1)\right)^{1/s}\\
&\leq \frac{1}{1 - p_n}\left((\bar{\eta})^s E \sum_{r=1}^K \left(\mathbf{1}_{\{I_r^{(n)} \leq \ell\}} |A_r^{(n)}|^s\right)\right.\\
&\qquad \left. + (\eta + \varepsilon)^s E \sum_{r=1}^K |A^{(n)}|^s + o(1)\right)^{1/s}.
\end{aligned}$$

With (4.7) and $n \to \infty$ we obtain

$$\eta \leq \zeta(\eta + \varepsilon) \qquad (4.15)$$

for all $\varepsilon > 0$. Since $\zeta < 1$ we obtain $\eta = 0$. The proof for $s < 1$ is similar. □

Remark 4.2 *Theorem 4.1 is restricted to the case of solutions of the limit equation in $\mathcal{M}_s$. In the existence and uniqueness result in Theorem 3.4 also solutions have been characterized without finite s-th moments. For several applications it is of interest to extend Theorem 4.1 to this more general case. This is to be considered in a separate paper.*

References

Cramer, M. (1997). Stochastic analysis of Merge-Sort algorithm. *Random Structures Algorithms 11*, 81–96.

Hwang, H.-K. and R. Neininger (2002). Phase change of limit laws in the quicksort recurrence under varying toll functions. *SIAM Journal on Computating 31*, 1687–1722.

Jagers, P. and U. Rösler (2002). Fixed points of max-recursive sequences. *Preprint*.

Knuth, D. E. (1973). *The Art of Computer Programming*, Volume 3: Sorting and Searching. Addison-Wesley Publishing Co., Reading.

Mahmoud, H. M. (2000). *Sorting*. Wiley-Interscience Series in Discrete Mathemstics and Optomization. Wiley-Interscience, New York.

Neininger, R. (1999). *Limit Laws for Random Recursive Structures and Algorithms*. Dissertation, University of Freiburg.

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

Neininger, R. and L. Rüschendorf (2002). Rates of convergence for Quicksort. *Journal of Algorithms 44*, 52–62.

Neininger, R. and L. Rüschendorf (2003a). A general limit theorem for recursive algorithms and combinatorial structures. To appear in: The Annals of Applied Probability.

Neininger, R. and L. Rüschendorf (2003b). On the contraction method with degenerate limit equation. To appear.

Rachev, S. T. (1991). *Probability Metrics and the Stability of Stochastic Models*. Wiley.

Rachev, S. T. and L. Rüschendorf (1992). Rate of convergene for sums and maxima and doubly ideal metrics. *Theory Prob. Appl. 37*, 276–289.

Rachev, S. T. and L. Rüschendorf (1995). Probability metrics and recursive algorithms. *Advances Applied Probability 27*, 770–799.

Régnier, M. (1989). A limiting distribution for quicksort. *RAIRO, Informatique Théoriqué et Appl. 33*, 335–343.

Rösler, U. (1991). A limit theorem for Quicksort. *RAIRO, Informatique Théoriqué et Appl. 25*, 85–100.

Rösler, U. (1992). A fixed point theorem for distribution. *Stochastic Processes Applications 42*, 195–214.

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

Rösler, U. and L. Rüschendorf (2001). The contraction method for recursive algorithms. *Algorithmica 29*, 3–33.

Zolotarev, V. M. (1997). *Modern Theory of Summation of Random Variables.* VSP, Utrecht.