On the Contraction Method with Reduced Independence Assumptions

Ralph Neininger ✉
Institute for Mathematics, Goethe University, 60054 Frankfurt a.M., Germany

Jasmin Straub ✉
Institute for Mathematics, Goethe University, 60054 Frankfurt a.M., Germany

──── **Abstract** ────

Recursive sequences of laws of random variables (and random vectors) are considered where an independence assumption which is usually made within the setting of the contraction method is dropped. This restricts the study to sequences which after normalization lead to asymptotic normality. We provide a general univariate central limit theorem which can directly be applied to problems from the analysis of algorithms and random recursive structures without further knowledge of the contraction method. Also multivariate central limit theorems are shown and bounds on rates of convergence are provided. Examples include some previously shown central limit analogues as well as new applications on Fibonacci matchings.

1 Introduction

Sequences $(Y_n)_{n \geq 0}$ of random vectors in $\mathbb{R}^d$, $d \in \mathbb{N}$, are considered which satisfy a distributional recursion

$$Y_n \stackrel{\mathrm{d}}{=} \sum_{r=1}^{K} A_r(n) Y_{I_r^{(n)}}^{(r)} + b_n, \qquad n \geq n_0, \tag{1}$$

where $\stackrel{\mathrm{d}}{=}$ denotes equality in distribution, $n_0 \in \mathbb{N}$, the coefficients $A_1(n), \ldots, A_K(n)$ are random $(d \times d)$-matrices and b_n is a d-dimensional random vector. Such recurrences often arise in the context of divide and conquer methods. Underlying such a recursion is a problem of size n that can be divided into K smaller subproblems of sizes $I_1^{(n)}, \ldots, I_K^{(n)}$, the toll term b_n measuring the "cost" of this division and the merger. Concerning the number of subproblems and the subproblem sizes, we will always make the following assumptions:

◇ The number K of subproblems is a fixed integer $K \geq 1$. However, extensions to K being random and depending on n are possible.

◇ The vector $I^{(n)} = (I_1^{(n)}, \ldots, I_K^{(n)})$ of the subproblem sizes is a random vector in $\{0, \ldots, n\}^K$.

Another integral part of this setting is the assumption that the subproblems are of the same nature as the original problem, or formally:

$$(Y_n^{(r)})_{n \geq 0} \stackrel{\mathrm{d}}{=} (Y_n)_{n \geq 0} \qquad \text{for } r = 1, \ldots, K. \tag{2}$$

Since this assumption guarantees the self-similarity between the initial structure and the parts into which the structure is decomposed, we will use the term *self-similarity condition* when referring to condition (2). Furthermore, we need some *conditional independence condition*

ensuring that given the subproblem sizes, the subproblems behave independently. To be more precise, within the contraction method, which is sketched below, usually it is assumed that

$$(A_1(n), \ldots, A_K(n), b_n, I^{(n)}), (Y_n^{(1)})_{n \geq 0}, \ldots, (Y_n^{(K)})_{n \geq 0} \quad \text{are independent.} \tag{3}$$

Note, however, that dependencies between the coefficients $A_1(n), \ldots, A_K(n), b_n$ and the subproblem sizes $I_1^{(n)}, \ldots, I_K^{(n)}$ are allowed.

Recurrences of the form (1) come up in various fields, see [11, 7] for many concrete examples ranging from complexity measures of recursive algorithms (e.g., the number of key comparisons used by Quicksort, Mergesort or Quickselect) to parameters of random trees (e.g., the size of tries and m-ary search trees, path lengths in digital search trees, (PATRICIA) tries and m-ary search trees or the number of leaves in quadtrees) to quantities of stochastic geometry (e.g., the number of maxima in right triangles). For all these examples, the contraction method can be used to derive limit laws, i.e., convergence in distribution of the normalized versions of the Y_n towards a limit distribution being either the normal distribution or some other distribution.

In recent years a couple of problems appeared which seemed to fall within the framework above, however, the conditional independence condition (3) was violated. Examples are central limit analogues for the complexity of Quicksort, the composition of cyclic (and other) urns and the number of leaves of random point quadtrees, see [6, 5, 4, 3]. Such applications, and new applications discussed below, can be covered under a weakened independence condition that

$$(A_1(n), \ldots, A_K(n), I^{(n)}), (Y_n^{(1)})_{n \geq 0}, \ldots, (Y_n^{(K)})_{n \geq 0} \quad \text{are independent.} \tag{4}$$

Note that, in contrast to the conditional independence condition (3), we allow dependencies between b_n and $(Y_n^{(1)})_{n \geq 0}, \ldots, (Y_n^{(K)})_{n \geq 0}$ here. Thus, condition (4) is slightly weaker than condition (3) and will be referred to as *partial conditional independence condition* in the following.

To observe a first implication of the partial conditional independence condition (4) on the setting of the contraction method we sketch the usual approach. We define the normalized sequence $(X_n)_{n \geq 0}$ by

$$X_n := C_n^{-1/2}(Y_n - M_n), \qquad n \geq 0, \tag{5}$$

where M_n is a d-dimensional vector and C_n a positive definite $(d \times d)$-matrix. Essentially, we choose M_n as the mean vector and C_n as the covariance matrix of Y_n if they exist (assuming $\mathrm{Cov}(Y_n)$ being positive definite for all sufficiently large n) or as the leading order terms in expansions of these moments as $n \to \infty$. The normalized quantities satisfy the following modified recursion:

$$X_n \overset{\mathrm{d}}{=} \sum_{r=1}^{K} A_r^{(n)} X_{I_r^{(n)}}^{(r)} + b^{(n)}, \qquad n \geq n_0, \tag{6}$$

with

$$A_r^{(n)} := C_n^{-1/2} A_r(n) C_{I_r^{(n)}}^{1/2}, \quad b^{(n)} := C_n^{-1/2} \left(b_n - M_n + \sum_{r=1}^{K} A_r(n) M_{I_r^{(n)}} \right) \tag{7}$$

and self-similarity and independence conditions as above. Then, limits

$$A_r^{(n)} \to A_r, \quad b^{(n)} \to b \tag{8}$$

are identified (in an appropriate sense) and a potential limit X of $(X_n)_{n \geq 0}$ in distribution is identified by satisfying the recursive distributional equation

$$X \overset{\mathrm{d}}{=} \sum_{r=1}^{K} A_r X_r + b, \tag{9}$$

where $X_1, \ldots, X_K$ are independent and identically distributed as X. Under the conditional independence condition (3) it is now justified to require on the right hand side of (9) that moreover $X_1, \ldots, X_K$ and $(A_1, \ldots, A_K, b)$ are independent. Hence, the distribution of the right hand side of (9) is then fully specified, see [7] for details. However, under the partial conditional independence condition (4) dependencies between b and the $X_1, \ldots, X_K$ have to be taken into account, where it is not clear in general how to define the right hand side of (9). For this reason, in the current extended abstract we restrict to the case $b = 0$, and require $X_1, \ldots, X_K$ and $(A_1, \ldots, A_K)$ to be independent regaining a well-defined right hand side of (9). The assumption $b = 0$ essentially restricts the framework to central limit theorems, although other limit laws, such a stable laws, are still possible as fixed-points, but not covered by our techniques.

Since most applications of the contraction method in the analysis of algorithms are for univariate Y_n, i.e., dimension $d = 1$, and with $A_r^{(n)} = 1$ for all $r = 1, \ldots, K$ and all $n \geq 0$ we provide a theorem convenient for applications. Note that the identification of the limits in (8) usually requires control on the expansions of moments related to the M_n and C_n. These expansions are already built into the subsequent theorem so that one does not need to have any knowledge about the underlying contraction method to apply it. For the reader's convenience we also recall the previous conditions (1), (2) and (4) for this frequently occurring univariate case: In dimension $d = 1$ consider the special case of (1) where Y_n satisfies the distributional recursion

$$Y_n \overset{\mathrm{d}}{=} \sum_{r=1}^{K} Y_{I_r^{(n)}}^{(r)} + b_n, \qquad n \geq n_0, \tag{10}$$

where $I^{(n)}, (Y_n^{(1)})_{n \geq 0}, \ldots, (Y_n^{(K)})_{n \geq 0}$ are independent (this is the partial conditional independence condition), $Y_j^{(r)}$ has the same distribution as Y_j for $j \geq 0$ and $r = 1, \ldots, K$, the subproblem sizes $I_r^{(n)}$ are in $\{0, 1, \ldots, n\}$ and satisfy $\mathbb{P}(I_r^{(n)} = n) \to 0$ as $n \to \infty$ and all appearing quantities are L_3-integrable. Then, we have the following theorem:

▶ **Theorem 1.** *Let $(Y_n)_{n \geq 0}$ be a sequence of random variables in $\mathbb{R}$ that satisfies recursion (10). Suppose that, for some positive functions f and g and as $n \to \infty$,*

$$\mathbb{E}[Y_n] = f(n) + \mathrm{o}(g^{1/2}(n)), \quad \mathrm{Var}(Y_n) = g(n) + \mathrm{o}(g(n)). \tag{11}$$

Further assume that for all $r = 1, \ldots, K$ and as $n \to \infty$, we have the L_3-convergences

$$\frac{g^{1/2}(I_r^{(n)})}{g^{1/2}(n)} \to A_r^*, \qquad \frac{1}{g^{1/2}(n)} \left(b_n - f(n) + \sum_{r=1}^{K} f(I_r^{(n)}) \right) \to 0 \tag{12}$$

with $(A_1^*)^2 + \cdots + (A_K^*)^2 = 1$ almost surely and $\mathbb{P}(\exists r : A_r^* = 1) < 1$. If the technical condition

$$\frac{\mathbb{P}(I_r^{(n)} < \ell)}{g^{3/2}(n)} \to 0 \qquad (n \to \infty) \tag{13}$$

is satisfied for any $\ell \in \mathbb{N}$ and $r = 1, \ldots, K$, then we have

$$\frac{Y_n - f(n)}{g^{1/2}(n)} \overset{\mathrm{d}}{\longrightarrow} \mathcal{N}(0, 1) \qquad (n \to \infty).$$

▶ **Remark 2.** Theorem 1 generalizes the central limit theorem in [7, Corollary 5.2], where the toll function b_n and $(Y_n^{(1)})_{n \geq 0}, \ldots, (Y_n^{(K)})_{n \geq 0}$ are additionally assumed to be independent.

In the current extended abstract we also consider the multivariate case of (1). We provide a multivariate central limit theorem under the partial conditional independence condition (4) and study bounds on the rates of convergence, see Theorem 3, in the Zolotarev metric ζ_3, which is the main tool underlying the proofs of the convergence theorems in this extended abstract, see Section 2. In Section 3 a sketch of the proof of Theorem 1 is given. Applications of Theorems 1 and 3 to the analysis of algorithms are given in Section 4.

2   A multivariate CLT including rates of convergence

In this section, we consider a sequence $(Y_n)_{n \geq 0}$ of d-dimensional random vectors satisfying the distributional recursion

$$Y_n \stackrel{\mathrm{d}}{=} \sum_{r=1}^{K} A_r(n) Y_{I_r^{(n)}}^{(r)} + b_n, \qquad n \geq n_0, \tag{14}$$

where $n_0 \in \mathbb{N}$, the coefficients $A_1(n), \ldots, A_K(n)$ are random $(d \times d)$-matrices, b_n is a d-dimensional random vector, $I^{(n)} = (I_1^{(n)}, \ldots, I_K^{(n)})$ is a random vector in $\{0, \ldots, n\}^K$ and all appearing quantities have finite third absolute moments. Furthermore, we assume that the self-similarity condition (2) and the partial conditional independence condition (4) are satisfied.

We assume that there exists some $n_1 \in \mathbb{N}_0$ such that the covariance matrix of Y_n is positive definite for $n \geq n_1$ and define the normalized sequence $(X_n)_{n \geq 0}$ by

$$X_n := C_n^{-1/2}(Y_n - M_n), \qquad n \geq 0, \tag{15}$$

where M_n is chosen as the mean vector of Y_n and C_n as the covariance matrix of Y_n for $n \geq n_1$ (and $C_n = \mathrm{Id}_d$ for $n < n_1$, where Id_d denotes the $d \times d$ identity matrix). The normalized quantities satisfy the following modified recursion:

$$X_n \stackrel{\mathrm{d}}{=} \sum_{r=1}^{K} A_r^{(n)} X_{I_r^{(n)}}^{(r)} + b^{(n)}, \qquad n \geq n_0, \tag{16}$$

with $A_r^{(n)}$ and $b^{(n)}$ given in (7) and self-similarity and independence relations as in (14).

To obtain a convergence result $X_n \to \mathcal{N}(0, \mathrm{Id}_d)$, we assume that the coefficients appearing in (7) converge appropriately, i.e., that there exist L_3-integrable random variables $A_1^*, \ldots, A_K^*$ such that, as $n \to \infty$,

$$(A_1^{(n)}, \ldots, A_K^{(n)}, b^{(n)}) \xrightarrow{L_3} (A_1^*, \ldots, A_K^*, 0), \tag{17}$$

with $A_1^*(A_1^*)^{\mathrm{T}} + \cdots + A_K^*(A_K^*)^{\mathrm{T}} = \mathrm{Id}_d$ almost surely. Then, from (16), we expect a limit X of X_n to satisfy the distributional fixed-point equation

$$X \stackrel{\mathrm{d}}{=} \sum_{r=1}^{K} A_r^* X^{(r)}, \tag{18}$$

where $(A_1^*, \ldots, A_K^*), X^{(1)}, \ldots, X^{(K)}$ are independent and $X^{(r)} \stackrel{\mathrm{d}}{=} X$ for $r = 1, \ldots, K$. Under the additional assumption $\sum_{r=1}^{K} \mathbb{E}[\|A_r^*\|_{\mathrm{op}}^3] < 1$, the multivariate standard normal distribution $\mathcal{N}(0, \mathrm{Id}_d)$ is the unique solution of equation (18) in the space $\mathcal{P}_3^d(0, \mathrm{Id}_d)$ of L_3-integrable probability distributions with mean vector 0 and covariance matrix Id_d, see, e.g., [7].

As a tool to derive convergence in distribution of (X_n) in (15) towards $\mathcal{N}(0, \mathrm{Id}_d)$ we use the Zolotarev metric ζ_3, which we only need and only define on the space $\mathcal{P}_3^d(0, \mathrm{Id}_d)$. For $\mathcal{L}(X), \mathcal{L}(Y) \in \mathcal{P}_3^d(0, \mathrm{Id}_d)$ we set

$$\zeta_3(X, Y) := \zeta_3(\mathcal{L}(X), \mathcal{L}(Y)) := \sup_{f \in \mathcal{F}_3} |\mathbb{E}[f(X) - f(Y)]| \tag{19}$$

where

$$\mathcal{F}_3 := \{f \in C^2(\mathbb{R}^d, \mathbb{R}) : \|f^{(2)}(x) - f^{(2)}(y)\| \le \|x - y\|\}, \tag{20}$$

with $C^2(\mathbb{R}^d, \mathbb{R})$ denoting the space of twice continuously differentiable functions from $\mathbb{R}^d$ to $\mathbb{R}$ and $f^{(2)}$ denoting the second derivative of $f \in C^2(\mathbb{R}^d, \mathbb{R})$.

In order to obtain such convergence, including a bound on the rate of convergence in ζ_3 we need a quantified version of the convergences in (17). In particular, we assume that

$$\left\| \sum_{r=1}^K \mathbb{E}[A_r^{(n)}(A_r^{(n)})^{\mathrm{T}}] - \mathrm{Id}_d \right\|_{\mathrm{op}} + \left\| \sum_{r=1}^K A_r^{(n)}(A_r^{(n)})^{\mathrm{T}} - \mathrm{Id}_d \right\|_{3/2}^{3/2} + \left\| b^{(n)} \right\|_3 = \mathrm{O}(R(n)) \tag{21}$$

for some monotonically decreasing sequence $R(n) > 0$ with $R(n) \to 0$. Furthermore, we assume the technical conditions

$$\left\| \mathbf{1}_{\{I_r^{(n)} < \ell\}} A_r^{(n)} \right\|_2^2 + \left\| \mathbf{1}_{\{I_r^{(n)} < \ell\}} A_r^{(n)} \right\|_3^3 = \mathrm{O}(R(n)), \quad n \to \infty, \tag{22}$$

for all $\ell \in \mathbb{N}$ and $r = 1, \ldots, K$ and

$$\left\| \mathbf{1}_{\{I_r^{(n)} = n\}} A_r^{(n)} \right\|_3 \to 0, \quad n \to \infty, \tag{23}$$

for all $r = 1, \ldots, K$. Under these assumptions, with the partial conditional independence condition (4) we have:

▶ **Theorem 3.** *Let $(X_n)_{n \ge 0}$ be given as in (15) with $(Y_n)_{n \ge 0}$ satisfying the distributional recurrence (14) with the self-similarity condition (2) and the partial conditional independence condition (4). Furthermore, assume that the coefficients $(A_1^{(n)}, \ldots, A_K^{(n)}, b^{(n)})$ defined in (7) converge to $(A_1^*, \ldots, A_K^*, 0)$ in the L_3 norm as $n \to \infty$ with $\sum A_r^*(A_r^*)^{\mathrm{T}} = \mathrm{Id}_d$ almost surely. If conditions (21), (22) and (23) are satisfied and if*

$$\limsup_{n \to \infty} \sum_{r=1}^K \mathbb{E}\left[\frac{R(I_r^{(n)})}{R(n)} \left\| A_r^{(n)} \right\|_{\mathrm{op}}^3 \right] < 1, \tag{24}$$

then we have, as $n \to \infty$,

$$\zeta_3(X_n, \mathcal{N}(0, \mathrm{Id}_d)) = \mathrm{O}(R(n)).$$

▶ **Remark 4.** Condition (21) in Theorem 3 should be compared to a similar result under the stronger conditional independence condition (3) which is stated in Theorem 3 of [8], see the corresponding condition (12) there where only $\|b^{(n)}\|_3^3 = \mathrm{O}(R(n))$ is required. In (21) in the present Theorem 3 no additional factor 3 in the exponent appears which is caused by the additional dependencies. However, it is not clear whether Theorem 3 is tight in this respect or if this factor 3 may be regained by some other argument.

▶ Remark 5. In some applications, we are only interested in showing (weak) convergence rather than estimating the rate of convergence. In this case, the formulation of Theorem 3 is more complex than necessary. More specifically, if no rates are needed, we can replace condition (24) by the assumption that

$$\sum_{r=1}^{K} \mathbb{E}[\|A_r^*\|_{\mathrm{op}}^3] < 1.$$

Furthermore, condition (21) can be dropped and instead of condition (22), it is enough to assume that

$$\left\|\mathbf{1}_{\{I_r^{(n)} < \ell\}} A_r^{(n)}\right\|_3 \to 0$$

as $n \to \infty$ for any $\ell \in \mathbb{N}$ and $r = 1, \ldots, K$. With these modified conditions, similar arguments as in the proof of Theorem 3 can be used to show that $\zeta_3(X_n, \mathcal{N}(0, \mathrm{Id}_d))$ converges to zero as $n \to \infty$ and hence, X_n converges in distribution to $\mathcal{N}(0, \mathrm{Id}_d)$.

A proof of Theorem 3 is given in the full paper version of this extended abstract.

3   Sketch of the proof of Theorem 1

We assume that all quantities are as in the formulation of Theorem 1.

Sketch of the proof of Theorem 1. Since we have $\mathrm{Var}(Y_n) = g(n) + \mathrm{o}(g(n))$ for some positive function g, we can find some constant $n_1 \in \mathbb{N}_0$ such that $\mathrm{Var}(Y_n)$ is positive for $n \geq n_1$. As before, we define the standardized quantities by

$$X_n := \frac{Y_n - \mu(n)}{\sigma(n)}, \qquad n \geq 0,$$

where $\mu(n) := \mathbb{E}[Y_n]$, $\sigma^2(n) := \mathrm{Var}(Y_n)$ for $n \geq n_1$ and $\sigma(n) = 1$ for $n < n_1$. The statement of the theorem follows directly from the asymptotic expansions in (11) and Slutsky's theorem if we show that the normalized quantities X_n converge in distribution to the standard normal distribution. Thus, it is sufficient to show that the Zolotarev distance $\Delta(n) := \zeta_3(X_n, \mathcal{N}(0,1))$ converges to zero as $n \to \infty$. The sequence $(X_n)_{n \geq 0}$ satisfies the modified recursion

$$X_n \stackrel{\mathrm{d}}{=} \sum_{r=1}^{K} A_r^{(n)} X_{I_r^{(n)}}^{(r)} + b^{(n)}, \qquad n \geq n_0, \tag{25}$$

with $I^{(n)}, (X_n^{(1)})_{n \geq 0}, \ldots, (X_n^{(K)})_{n \geq 0}$ independent, $X_j^{(r)}$ identically distributed as X_j for $j \geq 0$ and $r = 1, \ldots, K$ and

$$A_r^{(n)} = \frac{\sigma(I_r^{(n)})}{\sigma(n)}, \qquad b^{(n)} = \frac{1}{\sigma(n)} \left(b_n - \mu(n) + \sum_{r=1}^{K} \mu(I_r^{(n)})\right).$$

By conditions (11), (12) and (13), we have $A_r^{(n)} \to A_r^*$ and $b^{(n)} \to 0$ in L_3 for $r = 1, \ldots, K$. We define

$$Z_n := \sum_{r=1}^{K} A_r^{(n)} \tau_{I_r^{(n)}} N^{(r)},$$

where $I^{(n)}, N^{(1)}, \ldots, N^{(K)}$ are independent, the deterministic non-negative sequence $(\tau_n)_{n \geq 0}$ is defined by $\tau_n^2 = \operatorname{Var}(X_n)$ for $n \geq 0$ and $N^{(r)}$ has the standard normal distribution for $r = 1, \ldots, K$. Then, Z_n is centered and has variance

$$\operatorname{Var}(Z_n) = \sum_{r=1}^{K} \mathbb{E}\big[(A_r^{(n)})^2 \tau_{I_r^{(n)}}^2\big] = \sum_{r=1}^{K} \Big(\mathbb{E}\big[(A_r^{(n)})^2\big] + \mathbb{E}\big[\mathbf{1}_{\{I_r^{(n)} < n_1\}} (A_r^{(n)})^2 \big(\tau_{I_r^{(n)}}^2 - 1\big)\big] \Big).$$

We now observe that for any $r = 1, \ldots, K$, the latter summand in the above sum converges to zero by condition (13), since this condition and Jensen's inequality imply that $\mathbb{E}\big[\mathbf{1}_{\{I_r^{(n)} < n_1\}} (A_r^{(n)})^2\big] \to 0$. Together with the fact that $A_r^{(n)}$ converges in the L_3 norm (and thus also in L_2) to A_r^* with $(A_1^*)^2 + \cdots + (A_K^*)^2 = 1$ almost surely, we obtain that $\operatorname{Var}(Z_n)$ converges to 1. Hence, we can choose a deterministic sequence $(\kappa_n)_{n \geq 0}$ with $\kappa_n \to 0$ such that

$$Z_n^* := (1 + \kappa_n) Z_n$$

has variance 1 for $n \geq n_1$ (where n_1 may need to be enlarged). As a consequence, the distributions of X_n, Z_n^* and $\mathcal{N}(0,1)$ are all in $\mathcal{P}_3^1(0, \mathrm{Id}_1)$ for $n \geq n_1$ and we can apply the triangle inequality to obtain

$$\Delta(n) = \zeta_3(X_n, \mathcal{N}(0,1)) \leq \zeta_3(X_n, Z_n^*) + \zeta_3(Z_n^*, \mathcal{N}(0,1)), \qquad n \geq n_1.$$

With $Q_n := A_1^{(n)} X_{I_1^{(n)}}^{(1)} + \cdots + A_K^{(n)} X_{I_K^{(n)}}^{(K)}$ and Lemma 3.4 in [5] we find

$$\begin{aligned}
\zeta_3(X_n, Z_n^*) &\leq \zeta_3(Q_n, Z_n) + \|Q_n\|_3^2 \|b^{(n)}\|_3 + \frac{1}{2}\|Q_n\|_3 \|b^{(n)}\|_3^2 + \frac{1}{2}\|b^{(n)}\|_3^3 \\
&\quad + \Big(|\kappa_n| + \frac{1}{2}|\kappa_n|^2 + \frac{1}{2}|\kappa_n|^3\Big)\|Z_n\|_3^3 \\
&= \zeta_3(Q_n, Z_n) + o(1),
\end{aligned}$$

since $b^{(n)}$ converges to zero in the L_3 norm, κ_n converges to zero and $\|Z_n\|_3$ and $\|Q_n\|_3$ are bounded in n, the latter by Lemma 6 (note that we have $\sum \mathbb{E}[(A_r^*)^3] < 1$ by the assumptions $\sum (A_r^*)^2 = 1$ almost surely and $\mathbb{P}(\exists r : A_r^* = 1) < 1$ and that the technical condition (23) is satisfied since we assumed $\mathbb{P}(I_r^{(n)} = n) \to 0$ for $r = 1, \ldots, K$, thus Lemma 6 is applicable). Conditioning on $I^{(n)}$ implies that, for $n \geq n_1$,

$$\begin{aligned}
\zeta_3(Q_n, Z_n) &\leq \Big(\sum_{r=1}^{K} \mathbb{P}(I_r^{(n)} = n) \Big) \Delta(n) + \mathbb{E}\Big[\sum_{r=1}^{K} \mathbf{1}_{\{n_1 \leq I_r^{(n)} < n\}} (A_r^{(n)})^3 \Delta(I_r^{(n)}) \Big] \\
&\quad + \mathbb{E}\Big[\sum_{r=1}^{K} \mathbf{1}_{\{I_r^{(n)} < n_1\}} (A_r^{(n)})^3 \sup_{k < n_1} \zeta_3(X_k, \tau_k N^{(r)}) \Big] \\
&= o(1)\Delta(n) + \mathbb{E}\Big[\sum_{r=1}^{K} \mathbf{1}_{\{n_1 \leq I_r^{(n)} < n\}} (A_r^{(n)})^3 \Delta(I_r^{(n)}) \Big] + o(1),
\end{aligned}$$

where we used the assumption $\mathbb{P}(I_r^{(n)} = n) \to 0$ for $r = 1, \ldots, K$ and the technical condition (13) in the last step. Furthermore, we have $\zeta_3(Z_n^*, \mathcal{N}(0,1)) \to 0$. This can be seen by showing that $\ell_3(Z_n^*, \mathcal{N}(0,1)) \to 0$ and that $\|Z_n^*\|_3$ is bounded in n, where ℓ_3 denotes the minimal L_3-metric. Collecting all estimates, we obtain that

$$\Delta(n) \leq o(1)\Delta(n) + \mathbb{E}\Big[\sum_{r=1}^{K} \mathbf{1}_{\{n_1 \leq I_r^{(n)} < n\}} (A_r^{(n)})^3 \Delta(I_r^{(n)}) \Big] + o(1). \tag{26}$$

From this, the statement follows by a standard argument (see, e.g., [7, pp. 390–391]). ◄

▶ **Lemma 6.** *Let $(X_n)_{n \geq 0}$ be given as in (15) with $(Y_n)_{n \geq 0}$ satisfying the distributional recurrence (14). Furthermore, assume that the coefficients $(A_1^{(n)}, \ldots, A_K^{(n)}, b^{(n)})$ defined in (7) converge to $(A_1^*, \ldots, A_K^*, 0)$ in the L_3 norm as $n \to \infty$ with $\sum A_r^*(A_r^*)^{\mathrm{T}} = \mathrm{Id}_d$ almost surely and $\sum \mathbb{E}[\|A_r^*\|_{\mathrm{op}}^3] < 1$ and that the technical condition (23) is satisfied. Then we have, as $n \to \infty$, $\|X_n\|_3 = \mathrm{O}(1)$.*

The proof of Lemma 6 follows along the same lines as the proof of Lemma 2.3 in [6], generalizing to a more general setting with multivariate quantities and an arbitrary number K of subproblems here. However, in the sketch of the proof of Theorem 1 only the case $d = 1$ is needed. Further details can also be found in Straub [12].

4   Applications

As mentioned in the introduction, possible examples of distributional recurrences with dependent toll function, where our results apply can be found in [6, 5, 4, 3]. Since the central limit analogue for the complexity of Quicksort in [6] only contains a convergence result for the Zolotarev metric without a rate of convergence, we take up this example in section 4.1 and use Theorem 3 to rederive this central limit analogue and add a bound on the rate of convergence. Furthermore, in section 4.2, we present another application of Theorems 3 and 1 concerning recent work of Diaconis and Kolesnik [2] on Fibonacci permutations.

4.1   Refined Quicksort asymptotics

We consider the Quicksort algorithm where we set $K_0 = 0$ and denote by K_n, $n \geq 1$, the number of key comparisons needed by Quicksort to sort the list $(U_1, \ldots, U_n)$, where $(U_i)_{i \geq 1}$ is a sequence of independent and uniformly on the unit interval distributed random variables. With the normalization

$$C_n := \frac{K_n - \mathbb{E}[K_n]}{n + 1}, \qquad n \geq 0,$$

Régnier [9] showed for a suitable version that the sequence $(C_n)_{n \geq 0}$ is a martingale converging almost surely (and in L_p) to some non-degenerate limit C. Rösler [10] showed that C satisfies the distributional fixed-point equation

$$C \stackrel{\mathrm{d}}{=} U C^{(1)} + (1 - U) C^{(2)} + \varphi(U)$$

with U, $C^{(1)}$, $C^{(2)}$ independent, U uniform on the unit interval, $C^{(1)}$ and $C^{(2)}$ having the same distribution as C and $\varphi(u) := 2u \log u + 2(1 - u) \log(1 - u) + 1$ for $u \in [0, 1]$.

The aim of this section is to further quantify the almost sure convergence $C_n \to C$ by analyzing the residual $C_n - C$. Bindjeme and Fill [1, Theorem 1.4] found the L_2-norm of this residual explicitly which implies that

$$\left\| C_n - C \right\|_2^2 = 2 \frac{\log n}{n} + \mathrm{O}\left(\frac{1}{n}\right) \tag{27}$$

and in [6, Theorem 1.1] it is shown that

$$\sqrt{\frac{n}{2 \log n}} \, (C_n - C) \stackrel{\mathrm{d}}{\longrightarrow} \mathcal{N}(0, 1)$$

as $n \to \infty$. We now show that the application of Theorem 3 provides a rate of convergence in the Zolotarev metric ζ_3 for the latter convergence without much effort. For this, we need some of the results deduced in [1] and [6]. First of all, we use the notation $Y_n := C_n - C$, $n \geq 0$,

for the residuals. Note that we chose this notation, although differing from the notation used in [1] and [6], to guarantee that the notation is in accordance with the formulation of our theorems. We then observe that the residuals Y_n can be decomposed recursively. Equation (12) in [6] states a sample-pointwise recurrence for the error term Y_n (see also equation (2.6) in [1]), from which we obtain that Y_n satisfies the distributional recursion

$$Y_n \stackrel{\mathrm{d}}{=} \frac{I_n + 1}{n + 1} Y_{I_n}^{(1)} + \frac{n - I_n}{n + 1} Y_{n-1-I_n}^{(2)} + b_n, \qquad n \geq 1, \tag{28}$$

with I_n, $(Y_n^{(1)})_{n \geq 0}$, $(Y_n^{(2)})_{n \geq 0}$ independent, I_n uniformly distributed on $\{0, \ldots, n-1\}$, $Y_j^{(r)}$ distributed as Y_j for $j \geq 0$ and $r = 1, 2$ and some toll function b_n which is not independent of $(Y_n^{(1)})_{n \geq 0}$ and $(Y_n^{(2)})_{n \geq 0}$. Since the concrete representation of b_n is not needed in the following, we omit the details here and refer to [1] and [6]. Certainly, recurrence (28) is an instance of recursion (14) with $n_0 = 1$, $K = 2$, $I_1^{(n)} = I_n$, $I_2^{(n)} = n - 1 - I_n$ and $A_r(n) = (I_r^{(n)} + 1)/(n + 1)$ for $r = 1, 2$.

We denote the normalized residuals by

$$X_n := \frac{Y_n}{\sigma(n)}, \quad n \geq 0,$$

where $\sigma^2(n) := \mathrm{Var}(Y_n) > 0$ for $n \geq 0$. Note that $\sigma^2(n) = 2 \log n / n + \mathrm{O}(1/n)$ by (27) and the fact that both components C_n and C of Y_n are centered. The normalized quantities satisfy recursion (16) with the same parameters as above and

$$A_1^{(n)} = \frac{(I_n + 1)\,\sigma(I_n)}{(n+1)\,\sigma(n)}, \quad A_2^{(n)} = \frac{(n - I_n)\,\sigma(n - 1 - I_n)}{(n+1)\,\sigma(n)}, \quad b^{(n)} = \frac{b_n}{\sigma(n)}.$$

For these coefficients, we obtain the L_3-convergences (see [6])

$$A_1^{(n)} \to \sqrt{U_1} =: A_1^*, \quad A_2^{(n)} \to \sqrt{1 - U_1} =: A_2^*, \quad b^{(n)} \to 0.$$

Thus, we are in the situation of Section 2 and now check the conditions of Theorem 3 with $R(n) = \log^{-1/2}(n)$. First of all, Lemma 2.2 in [6] states that, as $n \to \infty$,

$$\|b^{(n)}\|_3 = \mathrm{O}\Big(\frac{1}{\sqrt{\log n}}\Big) = \mathrm{O}(R(n)).$$

The order of $\|(A_1^{(n)})^2 + (A_2^{(n)})^2 - 1\|_{3/2}$ can be bound, using [8, Lemma 7], by

$$\left\| \sum_{r=1}^{2} (A_r^{(n)})^2 - 1 \right\|_{3/2}$$
$$\leq \frac{1}{n \log n} \left\| I_n \log\Big(\frac{I_n}{n}\Big) + (n - 1 - I_n) \log\Big(\frac{n - 1 - I_n}{n}\Big) \right\|_{3/2} + \mathrm{O}\Big(\frac{1}{\log n}\Big)$$
$$= \mathrm{O}\Big(\frac{1}{\log n}\Big).$$

From this, we also obtain

$$\left| \sum_{r=1}^{2} \mathbb{E}[(A_r^{(n)})^2] - 1 \right| \leq \left\| \sum_{r=1}^{2} (A_r^{(n)})^2 - 1 \right\|_{1} \leq \left\| \sum_{r=1}^{2} (A_r^{(n)})^2 - 1 \right\|_{3/2} = \mathrm{O}\Big(\frac{1}{\log n}\Big).$$

Thus, condition (21) is satisfied. Since $A_1^{(n)}$ and $A_2^{(n)}$ are uniformly bounded and I_n is uniform on $\{0, \dots, n-1\}$, the technical conditions (22) and (23) are clearly satisfied. Furthermore, we can use arguments of [8, Lemma 7], the fact that I_n/n converges almost surely to U_1 and the dominated convergence theorem to show that

$$\limsup_{n \to \infty} \sum_{r=1}^{2} \mathbb{E}\left[\frac{R(I_r^{(n)})}{R(n)}\left(A_r^{(n)}\right)^3\right] = 2\,\mathbb{E}[U_1^{3/2}] = \frac{4}{5} < 1,$$

such that all assumptions of Theorem 3 are satisfied and we obtain the following result.

▶ **Theorem 7.** *For the number K_n of key comparisons used by Quicksort to sort the list $(U_1, \dots, U_n)$ with $(U_i)_{i \geq 1}$ independent and uniformly distributed on the unit interval and the almost sure limit C of $C_n = (K_n - \mathbb{E}[K_n])/(n+1)$, we have, as $n \to \infty$,*

$$\zeta_3\left(\frac{C_n - C}{\sqrt{\mathrm{Var}(C_n - C)}}, \mathcal{N}(0,1)\right) = \mathrm{O}\left(\frac{1}{\sqrt{\log n}}\right).$$

▶ **Remark 8.** In view of Remark 4 we are not sure whether the bound $\mathrm{O}(\log(n)^{-1/2})$ in Theorem 7 is tight or if $\mathrm{O}(\log(n)^{-3/2})$ may be the correct order.

4.2   Importance sampling for estimating the number of Fibonacci matchings

In this section we refer to the paper [2] by Diaconis and Kolesnik from which we adopt the notation and some of their results. The set $\mathfrak{F}_n$ of *Fibonacci matchings of size n* is defined by

$$\mathfrak{F}_n = \{\pi \in \mathcal{S}_n : |\pi(i) - i| \leq 1 \text{ for } 1 \leq i \leq n\},$$

where $\mathcal{S}_n$ denotes the set of permutations of $\{1, \dots, n\}$ (note that Diaconis and Kolesnik use the notation $\mathcal{F}_{n,1}$ instead of $\mathfrak{F}_n$). The set $\mathfrak{F}_4$ of Fibonacci matchings of size $n = 4$ is displayed in Figure 1.
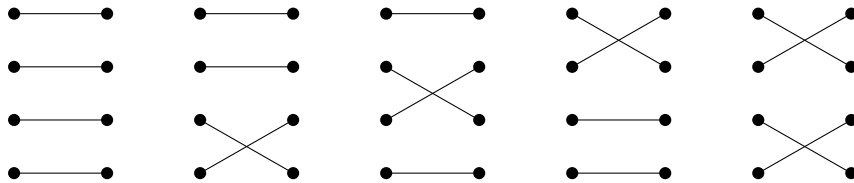
■ **Figure 1** The 5 Fibonacci matchings of size $n = 4$.

Note that the cardinality of the set $\mathfrak{F}_n$ is easily computed by considering whether $\pi(1) = 1$ or $\pi(1) = 2$ and coincides with the $(n+1)$-th Fibonacci number (which explains the name). Although the number of Fibonacci matchings is known, Diaconis and Kolesnik [2] present different importance sampling algorithms for estimating the size of $\mathfrak{F}_n$. These algorithms in each step match the current index with an index chosen uniformly at random among the remaining allowed indices. To be more precise, Diaconis and Kolesnik present three such algorithms differing in the order the indices are matched (see [2] for details):

- The *random algorithm A_r* matches the indices in uniformly random order,
- the *fixed-order algorithm A_f* matches them in fixed order from top to bottom and
- the *greedy algorithm A_g* matches them in a certain greedy order. More precisely, A_g always matches the smallest unmatched index among those indices with the maximal number of remaining choices. This means that algorithm A_g always starts by matching

index 2 uniformly at random with one of 1, 2, 3. If $\pi(2) \in \{1, 2\}$ (i.e., either $\pi(2) = 1$ and consequently $\pi(1) = 2$ or vice versa), then the next index to be matched is index 4 (uniformly among 3, 4, 5), since this is the smallest index with 3 remaining choices. Otherwise, i.e., if $\pi(2) = 3$, then the assignments $\pi(1) = 1$ and $\pi(3) = 2$ are forced and the next index to be matched is index 5 (uniformly among 4, 5, 6).

We summarize some of the results given in [2]: For any Fibonacci matching $\pi \in \mathfrak{F}_n$, we denote by $P_{\mathrm{r}}(\pi)$, $P_{\mathrm{f}}(\pi)$ and $P_{\mathrm{g}}(\pi)$ the probability of π under the algorithm A_{r}, A_{f} and A_{g}, respectively. For Π_n chosen uniformly at random from the set $\mathfrak{F}_n$ of Fibonacci matchings, we have

$$\mathbb{E}[-\log(P_\chi(\Pi_n))] = \mu_\chi n + \mathrm{O}(1), \qquad \mathrm{Var}(-\log(P_\chi(\Pi_n))) = \sigma_\chi^2 n + \mathrm{O}(1), \tag{29}$$

for $\chi = \mathrm{r}, \mathrm{f}, \mathrm{g}$, where $\mu_\chi \in (0.49, 0.51)$ and $\sigma_\chi^2 > 0$ can be computed exactly (we refer to [2] for the concrete values). Furthermore, for $\chi = \mathrm{r}, \mathrm{f}, \mathrm{g}$ and as $n \to \infty$, we have the central limit theorem [2, Theorem 1.1]

$$\frac{-\log(P_\chi(\Pi_n)) - \mu_\chi n}{\sigma_\chi \sqrt{n}} \xrightarrow{\mathrm{d}} \mathcal{N}(0, 1). \tag{30}$$

While this statement can be obtained with [7, Corollary 5.2] for the first two algorithms (see [2, Theorems 3.2 and 3.4]), there is no obvious way of applying [7, Corollary 5.2] in the greedy case due to arising dependencies. Instead, Diaconis and Kolesnik use arguments from renewal theory to show (30) for the greedy algorithm A_{g}. Using the results of the current extended abstract, one can also handle these additional dependencies arising in the greedy case. For this reason we focus on algorithm A_{g} from now on and define the random variable

$$Y_n = -\log(P_{\mathrm{g}}(\Pi_n)),$$

where, as before, Π_n is uniformly distributed on $\mathfrak{F}_n$ and $P_{\mathrm{g}}(\pi)$ denotes the probability of π under the greedy algorithm A_{g} for any $\pi \in \mathfrak{F}_n$. Recall that algorithm A_{g} always starts by matching index 2 uniformly at random with one of 1, 2, 3. Consequently, for a fixed Fibonacci permutation π, the probability that index 2 is matched correctly with $\pi(2)$ by A_{g} equals $1/3$. Depending on the value of $\pi(2)$, the resulting number of indices that are neither matched nor forced is $n - 2$ or $n - 3$ afterwards. Thus, we obtain that

$$Y_n \stackrel{\mathrm{d}}{=} Y_{I_1^{(n)}} + \log 3,$$

where $I_1^{(n)}$ takes the values $n - 2$ and $n - 3$ with probabilities $2|\mathfrak{F}_{n-2}|/|\mathfrak{F}_n|$ and $|\mathfrak{F}_{n-3}|/|\mathfrak{F}_n|$, respectively, and is independent of $(Y_j)_{j \geq 0}$. However, using this recursion, our theorems do not apply, since there is only one subproblem of almost the same size as the original problem (i.e., $A_1^* = 1$).

Instead, to obtain a recursion to which our framework applies, we now divide the permutation at the middle, more precisely at index $k_n = \lfloor n/2 \rfloor$, instead of dividing it at the top. Now, consider whether $\pi(k_n) = k_n - 1$, $\pi(k_n) = k_n$ or $\pi(k_n) = k_n + 1$. In the first case, the resulting subproblem sizes are $k_n - 2$ and $n - k_n$, whereas they are $k_n - 1$ and $n - k_n$ in the second case and $k_n - 1$ and $n - k_n - 1$ in the third case. Hence, we obtain the recursive decomposition

$$Y_n \stackrel{\mathrm{d}}{=} Y_{I_1^{(n)}}^{(1)} + Y_{I_2^{(n)}}^{(2)} + b_n, \tag{31}$$

where the vector $I^{(n)} = (I_1^{(n)}, I_2^{(n)})$ contains the subproblem sizes and has distribution

$$\mathbb{P}(I_1^{(n)} = i_1, I_2^{(n)} = i_2) = \frac{1}{|\mathfrak{F}_n|} \begin{cases} |\mathfrak{F}_{k_n-2}| \cdot |\mathfrak{F}_{n-k_n}|, & i_1 = k_n - 2, i_2 = n - k_n, \\ |\mathfrak{F}_{k_n-1}| \cdot |\mathfrak{F}_{n-k_n}|, & i_1 = k_n - 1, i_2 = n - k_n, \\ |\mathfrak{F}_{k_n-1}| \cdot |\mathfrak{F}_{n-k_n-1}|, & i_1 = k_n - 1, i_2 = n - k_n - 1, \end{cases}$$

$I^{(n)}$, $(Y_j^{(1)})_{j \geq 0}$ and $(Y_j^{(2)})_{j \geq 0}$ are independent, $Y_j^{(r)}$ has the same distribution as Y_j for $j \geq 0$ and $r = 1, 2$ and b_n is a random variable taking values between 0 and $\log(9/2)$. However, note that b_n is not independent of $(Y_j^{(1)})_{j \geq 0}$ and $(Y_j^{(2)})_{j \geq 0}$, which is the reason why Corollary 5.2 of [7] does not apply. However, we can use Theorem 1 instead: The sequence $(Y_n)_{n \geq 0}$ satisfies recursion (10) as well as condition (11) with $f(n) = \mu_{\mathrm{g}} n$ and $g(n) = \sigma_{\mathrm{g}}^2 n$. Furthermore, conditions (12) and (13) are obviously satisfied with $A_1^* = A_2^* = 1/\sqrt{2}$. Thus, Theorem 1 implies

$$\frac{Y_n - \mu_{\mathrm{g}} n}{\sigma_{\mathrm{g}} \sqrt{n}} \xrightarrow{\mathrm{d}} \mathcal{N}(0, 1),$$

giving another proof of (30) for the greedy case. Note that we can also apply Theorem 3 to derive a bound on the rate of convergence in the limit theorem (30) in the Zolotarev ζ_3 metric. While in the following Theorem 9 we cover algorithm A_{g}, corresponding results can also be derived for the other two algorithms via [8, Theorem 3].

▶ **Theorem 9.** *Let $P(\pi)$ be the probability of π under the random algorithm A_{g} for any Fibonacci permutation π. Further set $Y_n = -\log(P(\Pi_n))$, where Π_n is uniformly distributed on the set $\mathfrak{F}_n$ of Fibonacci matchings of length n. Then, for any $\varepsilon > 0$, as $n \to \infty$, we have*

$$\zeta_3 \left(\frac{Y_n - \mathbb{E}[Y_n]}{\sqrt{\mathrm{Var}(Y_n)}}, \mathcal{N}(0, 1) \right) = \mathrm{O}(n^{-1/2+\varepsilon}).$$

The proof of Theorem 9 follows easily from Theorem 3 using recurrence (31), $R(n) = n^{-1/2+\varepsilon}$ and noting that

$$\left\| \sum_{r=1}^{2} (A_r^{(n)})^2 - 1 \right\|_{3/2}^{3/2} + \left\| b^{(n)} \right\|_3 = \mathrm{O}(n^{-1/2}).$$

---- **References** ----

1   Patrick Bindjeme and James Allen Fill. Exact L^2-distance from the limit for QuickSort key comparisons (extended abstract). In *23rd Intern. Meeting on Probabilistic, Combinatorial, and Asymptotic Methods for the Analysis of Algorithms (AofA'12)*, Discrete Math. Theor. Comput. Sci. Proc., AQ, pages 339–348. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2012.

2   Persi Diaconis and Brett Kolesnik. Randomized sequential importance sampling for estimating the number of perfect matchings in bipartite graphs. *Adv. in Appl. Math.*, 131:Article 102247, 2021. `doi:10.1016/j.aam.2021.102247`.

3   Michael Fuchs, Noela S. Müller, and Henning Sulzbach. Refined asymptotics for the number of leaves of random point quadtrees. In *29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms*, volume 110 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 23, 16. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2018.

4   Noela Müller. *Central limit theorems for multicolour urn models.* PhD thesis, Goethe University Frankfurt a.M., 2017. `urn:nbn:de:hebis:30:3-453587`.

5   Noela Müller and Ralph Neininger. Refined asymptotics for the composition of cyclic urns. *Electron. J. Probab.*, 23:Paper no. 117, 20, 2018. `doi:10.1214/18-EJP243`.

6   Ralph Neininger. Refined Quicksort asymptotics. *Random Structures Algorithms*, 46(2):346–361, 2015. `doi:10.1002/rsa.20497`.

7   Ralph Neininger and Ludger Rüschendorf. A general limit theorem for recursive algorithms and combinatorial structures. *Ann. Appl. Probab.*, 14(1):378–418, 2004. `doi:10.1214/aoap/1075828056`.

8   Ralph Neininger and Jasmin Straub. Convergence rates in the probabilistic analysis of algorithms. In *31st International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms*, volume 159 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 22, 13. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2020.

9   Mireille Régnier. A limiting distribution for quicksort. *RAIRO Inform. Théor. Appl.*, 23(3):335–343, 1989. `doi:10.1051/ita/1989230303351`.

10   Uwe Rösler. A limit theorem for "Quicksort". *RAIRO Inform. Théor. Appl.*, 25(1):85–100, 1991. `doi:10.1051/ita/1991250100851`.

11   Uwe Rösler and Ludger Rüschendorf. The contraction method for recursive algorithms. *Algorithmica*, 29(1-2):3–33, 2001. Average-case analysis of algorithms (Princeton, NJ, 1998). `doi:10.1007/BF02679611`.

12   Jasmin Straub. *On Rates of Convergence in the Probabilistic Analysis of Algorithms*. PhD thesis, Goethe University Frankfurt a.M., 2021. `doi:10.21248/gups.62402`.