Density Approximation and Exact Simulation of Random Variables that are Solutions of Fixed-Point Equations

Luc Devroye[1] and Ralph Neininger[2]
School of Computer Science
McGill University
3480 University Street
Montreal, H3A 2K6
Canada

February 8, 2002

Abstract

An algorithm is developed for the exact simulation from distributions that are defined as fixed-points of maps between spaces of probability measures. The fixed-points of the class of maps under consideration include examples of limit distributions of random variables studied in the probabilistic analysis of algorithms. Approximating sequences for the densities of the fixed-points with explicit error bounds are constructed. The sampling algorithm relies on a modified rejection method.

AMS subject classifications. Primary: 65C10; secondary: 65C05, 68U20, 11K45.

Key words. Random variate generation; Fixed-point equation; Perfect simulation; Rejection method; Monte Carlo method.

1 Introduction

Let $\mathcal{L}(X)$ be the distribution of a random variable X that satisfies a distributional fixed-point equation of the form

$$X \sim \sum_{r=1}^{K} A_r X^{(r)} + b, \tag{1}$$

where the symbol $\sim$ denotes equality in distribution, $X^{(1)}, \dots, X^{(K)}, (A_1, \dots, A_K, b)$ are independent with $\mathcal{L}(X^{(r)}) = \mathcal{L}(X)$ for all r and given random variables $A_1, \dots, A_K, b$, and $K \geq 1$ is a fixed integer. In such a case we call $\mathcal{L}(X)$ or X a fixed-point of (1). Under various assumptions on $(A_1, \dots, A_K, b)$ and X it is known that such a fixed-point $\mathcal{L}(X)$ is unique, see (2) below.

[1]Research of both authors supported by NSERC grant A3450.
[2]Research supported by the Deutsche Forschungsgemeinschaft.

For a subclass of fixed-point equations of the form (1) which is particularly important in theoretical computer science we establish the existence of densities of the fixed-points, give algorithmically computable approximating sequences for these densities, and establish explicit error bounds for the approximation. We show that this can, in principle, be turned into an algorithm for the perfect simulation from the fixed-point distribution when we use the rejection method. The algorithm takes with probability one a finite time, but is not powerful enough to yield a practical simulation method in general. Our work should be considered more as a theoretical contribution, establishing the existence of an exact algorithm that can be designed based on the form of the fixed-point equation only.

Distributions appearing as fixed-points of equations as (1) appear in many different applied and pure areas of probability theory. The case $K = 1$ plays an important role in financial modelling, insurance mathematics, and hydrology, when the fixed-point equation $X \sim AX + b$ may characterize the stationary distribution of generalized autoregressive processes such as ARMA, ARCH or GARCH, used in modelling a stationary time series. Usually conditions for the existence of such stationary distributions are of interest and much effort is made to estimate the tails of these distributions. See Takás [41], Kesten [24], Vervaat [43], Bougerol and Picard [2], Goldie and Grübel [15], de Bruijn [7], Goldie and Maller [16], and Embrechts and Goldie [9], Embrechts, Klüpelberg and Mikosch [10, section 8.4].

Interestingly, the same equations $X \sim AX + b$ appear as well in theoretical computer science as the limit distributions of cost measures of one-sided divide and conquer algorithms, e.g., Hoare's selection algorithm. Here, the fixed-point property appears in many recursive algorithms. One of these distributions satisfying $X \sim UX + 1$ with U uniform $[0, 1]$ is the Dickman distribution, which has been studied in number theory, see Mahmoud, Moddarres, and Smythe [28], Grübel na Rösler [17], and Hwang and Tsai [23].

The case of fixed-point equations (1) with $K \geq 2$ usually appears in problems with a branching nature like branching processes, random fractals, and recursive algorithms. When a recursive algorithm divides the problem into $K \geq 2$ parts to recurse on them, the general case of equation (1) may characterize the limit distribution $\mathcal{L}(X)$ of an associate parameter. We give many examples in this area below, the most important being the limit distribution of the running time of the quicksort algorithm (see Figure 1 for the corresponding equation).

Approximate generation of X is possible by iterating (1) sufficiently often. It is easy to see that an infinite number of repetitions leads to an infinite complete K-ary tree, as at each step, each $X^{(r)}$ on the right-hand-side of (1) must be replaced. Breaking that tree off leads to an approximation. While this is a valid approach, we are asking the more fundamental question of how to simulate the fixed-point random variable X exactly.

This problem is virtually unsolved, an exception being Devroye [5], where special types of perpetuities, namely the case $K = 1$, $b = 1$, $A_1 = U^a$ with $a > 0$ and U uniform $[0, 1]$ distributed is treated. It would be most deserving to have exact generators for more general equations of this form.

To solve our problem, we need to get detailed information on the fixed-point distributions, preferably an algebraic expression for the density if at least a density exists. Clearly, when the fixed-point equation characterizes the limit distribution $\mathcal{L}(X)$ of some limit law $X_n \to X$, the distribution $\mathcal{L}(X)$ cannot be used for approximating $\mathcal{L}(X_n)$ explicitly, as long as the density or distribution function of X cannot be approximated. We will develop suitable approximations in this paper. It should be noted that the fixed-point distribution may behave badly. For example, Chen, Goodman, and Zame [3] exhibited a fixed-point with a density on $[0, 1]$ that is not continuous on a dense subset of $[0, 1]$.

The present paper deals with density approximation and exact simulation from a class of fixed-points where a first general restriction is $K \geq 2$. We hope to report on progress in the case $K = 1$ elsewhere. We have to introduce a few restrictions on the class of fixed-point equations in order to guarantee algorithmic tractability. As shown below, all important known fixed-point equations arising in the probabilistic analysis of algorithms satisfy these conditions.

QUICKSORT, a sorting algorithm invented by Hoare [18, 21], sorts n numbers using C_n comparisons. It is known that $\mathbb{E}\, C_n \sim 2n \log n$ (Sedgewick [38, 39]). Hennequin [19, 20] showed that there is a limit law: $(C_n - \mathbb{E}\, C_n)/n \to X$ where $\to$ denotes convergence in distribution and X is a positive random variable. That proof was based on the method of moments. Régnier [33] used a martingale argument to prove that same limit law. The distribution of X was shown by Rösler [34] to satisfy the fixed-point equation

$$X \sim UX + (1-U)X' + 1 + 2\ln(U) + 2(1-U)\ln(1-U)$$

where U is a uniform $[0,1]$ random variable, X is unique subject to $\mathbb{E}\, X^2 < \infty$, and X and X' are i.i.d. This is precisely the format studied in this paper. Fill and Janson [11, 12, 13] studied the distribution of X in more detail. As announced above, the present paper develops computable approximations of the density of X, as a special case of a more general series of approximations.

A general theory for equation (1) seems, however, to be far away. The exact simulation from these distributions is dealt with in only one paper, by Devroye, Fill, and Neininger [6]. In that paper, an algorithm for the QUICKSORT case is developed that is based on an inequality due to Fill and Janson [13]. Related distributions include the limit distributions of the number of key exchanges of QUICKSORT, linear combinations of key exchanges and comparison. Several random trees, such as the random m-ary search tree, the random median-of-$(2k+1)$ search tree, and the random quadtree, see for the definitions Mahmoud [27], Sedgewick and Flajolet [40], Knuth [25], and Flajolet, Labelle, Laforest, and Salvy [14] for probabilistic analysis of quadtrees, have an important parameter, the total internal path length I_n (the sum of the distances from the nodes to the root), which satisfies $(I_n - \mathbb{E}\, I_n)/n \to X$ for a different limit law $\mathcal{L}(X)$. That was proved via the contraction method by Rösler [34, 36], Neininger [29], Neininger and Rüschendorf [30], Dobrow and Fill [8] (with the method of moments), Hwang and Neininger [22]. In all cases, $\mathcal{L}(X)$ satisfies the type of fixed-point equation studied in this paper. For the contraction method, see Rösler [34, 35], Rachev and Rüschendorf [32], Neininger and Rüschendorf [31] or Rösler and Rüschendorf [37].

Using this method the conditions

$$\xi := \sum_{r=1}^{K} \mathbb{E}\, A_r^2 < 1, \quad \mathbb{E}\, b^2 < \infty, \quad \mathbb{E}\, b = 0 \tag{2}$$

ensure that (1) has a unique fixed-point X in the space $\mathcal{M}_{0,2}$ of centered probability measures with finite second moments: see the "Contraction Lemma" in Rösler and Rüschendorf [37, Lemma 1, Theorem 3]. It is also well known that with the map T associated to (1), for every $\nu \in \mathcal{M}_{0,2}$,

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

with $\mathcal{M}$ the space of univariate probability measures and $Z^{(1)}, \ldots, Z^{(K)}, (A_1, \ldots, A_K, b)$ independent and $\mathcal{L}(Z^{(r)}) = \lambda$ for all r, we have $T^{(n)}(\nu) := T \circ \cdots \circ T(\nu) \to \mathcal{L}(X)$ in distribution. The second moments converge as well.

The exact definition of the equations (1) under consideration here is given in section 2. Roughly, we assume that the distributions of the coefficients $A_1, \dots, A_K, b$ are given by a Skorohod representation, i.e., by measurable functions $f_1, \dots, f_K, h : [0,1]^d \to \mathbb{R}$ such that $A_r \sim f_r(U), b \sim h(U)$ for a uniform $[0,1]^d$ distributed random vector U. Since it is well-known that any univariate distribution has a Skorohod representation of the given form this introduces no restrictions on the fixed-point equations. We do however impose some restrictions on some functional properties of $f_1, \dots, f_K, h$.

Consistent with the literature on non-uniform random variate generation, we assume that an infinite sequence of i.i.d. uniform $[0,1]$ random variates is available, that real numbers can be stored with infinite precision, and that standard arithmetic operations dealing with real numbers can be performed in one unit of time (see, e.g., Devroye [4]). We give a general approach for exact random variate generation from the fixed-points of equations (1) of the class to be specified, where for concrete applications certain parameters have to be adjusted and do these adjustments for the examples of the limit laws of the internal path lengths in random m-ary search trees, random median of $(2k+1)$ search trees, and random quadtrees, the other examples mentioned above being slight modifications. In fact, the algorithms developed here are solely based on addition, subtraction, multiplication, division, and comparisons of real numbers. We use a modified rejection method, similar to but different from that used for related problems in Devroye [5] and Devroye, Fill, and Neininger [6]. Since the density of $\mathcal{L}(X)$ cannot be computed exactly from the fixed-point equation, a convergent sequence of approximations is constructed to decide the outcome of a rejection test. Although our algorithm may be costly and not feasible for practical purposes, it is the first algorithm for exact finite time random variate generation for these fixed-point distributions.

The main ingredients of the present approach are firstly a technique based on a method of van der Corput and developed in Fill and Janson [11] to prove that the fixed-points under consideration have infinitely differentiable densities where explicit bounds on the densities and their derivatives are available. From these bounds the dominant, integrable curve needed for the rejection method are derived. Secondly, we define a sequence of discretized versions T_n of T as follows. Roughly, we use convergent discretizations $A_r^{(n)}$ of A_r and $b^{(n)}$ of b to define

$$T_n : \mathcal{M} \to \mathcal{M}, \quad \lambda \mapsto \mathcal{L}\left(\sum_{r=1}^{K} A_r^{(n)} Z^{(r)} + b^{(n)}\right),$$

with relations as for T such that we still have the analogous property

$$\mu_n := T_n \circ T_{n-1} \circ \cdots \circ T_1(\nu) \to X,$$

where the convergence is in distribution and with second moments for all $\nu \in \mathcal{M}_{0,2}$. This convergence is made quantitative using the minimal L_2 metric ℓ_2, which is defined by

$$\ell_2(\lambda, \nu) := \inf\{\|Z - Y\|_2 : \mathcal{L}(Z) = \lambda, \mathcal{L}(Y) = \nu\}, \quad \lambda, \nu \in \mathcal{M}_2,$$

where $\mathcal{M}_2$ is the space of probability distributions with finite second moment (see Bickel and Freedman [1] for properties of ℓ_2). Then, thirdly, using tools of Fill and Janson [13], a rate of convergence for (μ_n) in the ℓ_2-metric leads to a rate in the Kolmogorov metric and an explicit rate of convergence of approximations of the density of X, which are defined in terms of the distribution functions of the μ_n.

The discrete nature of the T_n enables us to calculate the distributions of μ_n algorithmically using only elementary operations when starting with a simple ν, e.g., the Dirac measure in zero. To reduce the computational complexity we will in fact not exactly use μ_n as defined above; for each $n \in \mathbb{N}$ we first further discretize μ_{n-1} to $\langle \mu_{n-1} \rangle$ and then iterate $\mu_n := T_n(\langle \mu_{n-1} \rangle)$, cf. (25),(26).

Another possible approach based on the iteration of T itself and numerical integration to obtain approximations of the density of X was posed in Fill and Janson [12].

The paper is organized as follows: In section 2 we define the class of equations (1) under consideration and introduce the concrete examples related to QUICKSORT and the internal path lengths of random search trees. In section 3 we prove that the fixed-points have C^∞ densities and give explicit bounds on the densities and their derivatives. These bound are made explicit for the examples mentioned. In section 4 we develop a general rate of convergence for $\mu_n \to X$ depending on the accuracy of the approximation of the discretizations $A_r^{(n)}$ and $b^{(n)}$ leading to an algorithmically computable sequence of approximations of the density of X needed for the decision of the outcome of the rejection test. The length of the paper is mostly explained by the need to compute all bounds explicitly. We will work out these explicit estimates for three examples. In section 5 all parts are put together, which, from a theoretical point of view, gives an exact simulation algorithm. Some remarks on the algorithm's complexity round out the paper.

2 Fixed-point equations and examples

We specify the type of fixed-point equation under consideration and give examples form the probabilistic analysis of algorithms.

2.1 Fixed-points

Throughout this paper we assume that $\mathcal{L}(X)$ satisfies

$$X \sim \sum_{r=1}^{K} A_r X^{(r)} + b, \tag{3}$$

as in (1), where the coefficients $A_1, \dots, A_K$ are given by measurable functions $f_1, \dots, f_K : [0,1]^d \to [0,1]$ such that $d \geq 1, K \geq 2$, and $A_r \sim f_r(U)$ with U uniform $[0,1]^d$ distributed, where we exclude the case $f_r = 0$ for some r. We assume moreover, that $\sum_{r=1}^K f_r = 1$. Our approach does not heavily rely on this condition; it could be replaced by other conditions. The present setting is chosen since all examples mentioned fit into this scheme. For the representation of b denote

$$S_{K-1} := \left\{ v \in [0,1]^{K-1} : \sum_{i=1}^{K-1} v_i \leq 1 \right\}, \quad f := (f_1, \dots, f_{K-1}).$$

Then we assume that we have $b \sim g(f(U))$ and $\mathbb{E}\, b = 0$ with a function $g : S_{K-1} \to \mathbb{R}$ being twice continuously differentiable (in particular bounded) such that its Hessian matrix

$$\mathrm{Hess}(g; v) := \left(\frac{\partial^2 g}{\partial v_i \partial v_j}(v) \right)_{i,j=1}^{K-1}$$

is for all $v \in f([0,1]^d) \subset S_{K-1}$ (positive or negative) definite, i.e., $\langle x, \mathrm{Hess}(g;v)\, x \rangle > 0$ (or < 0 respectively) for all $x \in \mathbb{R}^{K-1}$, where $\langle \cdot\,, \cdot \rangle$ denotes the standard inner product on $\mathbb{R}^{K-1}$. Then the fixed-point equation (3) takes the form

$$X \sim \sum_{r=1}^{K} f_r(U) X^{(r)} + g(f(U)), \tag{4}$$

with $U, X^{(1)}, \ldots, X^{(K)}$ independent, $U \sim \mathrm{unif}[0,1]^d$ and $X^{(r)} \sim X$ for all r.

In this situation the conditions (2) are satisfied. We assume that $\mathbb{E}\, X^2 < \infty$, so that $\mathcal{L}(X)$ is then the unique solution of (4) in $\mathcal{M}_{0,2}$.

The following conditions on $f_1, \ldots, f_K, g$ are assumed:

1. There exist $s, p_0 > 0$ and nonnegative functions D_1, D_2 such that for all $c > 0, p \geq p_0, t \geq Kc$ holds

$$\sum_{j=1}^{K} \lambda^d \left(\bigcap_{\substack{r=1 \\ r \neq j}}^{K} \{f_r \leq c/t\} \right) \leq \frac{D_1(c)}{t^s}, \tag{5}$$

$$\sum_{r=1}^{K} \int \mathbf{1}_{\{f_r \geq c/t\}} f_r^{-p}(u)\, du \leq \frac{D_2(p,c)}{t^{s-p}}, \tag{6}$$

   where λ^d denotes the d-dimensional Lebesgue measure.

2. There exists a $p_1 > p_0/K$ such that for all $0 < p < p_1$

$$M_p := \int_{[0,1]^d} \prod_{r=1}^{K} f_r^{-p}(u)\, du < \infty. \tag{7}$$

3. The cube $[0,1]^d$ can be decomposed (up to sets of Lebesgue measure zero) into measurable sets $(G_n)_{n \in \mathbb{N}}$, such that for all $n \in \mathbb{N}$ there exists a component $\ell = \ell(n)$, $1 \leq \ell \leq d$ such that the ℓ-cut $G_{n,\ell}(\tilde{u})$ of G_n,

$$\begin{aligned} G_{n,\ell}(\tilde{u}) &:= \{u_\ell \in [0,1] : [u_\ell, \tilde{u}] \in G_n\}, & (8)\\ [u_\ell, \tilde{u}] &:= (\tilde{u}_1, \ldots, \tilde{u}_{\ell-1}, u_\ell, \tilde{u}_\ell, \ldots, \tilde{u}_{d-1}), & (9) \end{aligned}$$

   is an interval and that the maps

$$u_\ell \mapsto f_r([u_\ell, \tilde{u}])$$

   are affine on $G_{n,\ell}(\tilde{u})$ for all $r = 1, \ldots, K$, at least one of these functions having nonzero derivative. Then we define

$$G'_{n,\ell} := \{\tilde{u} \in [0,1]^{d-1} : G_{n,\ell}(\tilde{u}) \neq \emptyset\}, \tag{10}$$

   and on $G'_{n,\ell}$ the function

$$\gamma(\tilde{u}) := \inf_{u_\ell \in G_{n,\ell}(\tilde{u})} \left| \left\langle \frac{\partial f}{\partial u_\ell}([u_\ell, \tilde{u}]), \mathrm{Hess}(g; f([u_\ell, \tilde{u}])) \frac{\partial f}{\partial u_\ell}([u_\ell, \tilde{u}]) \right\rangle \right| \tag{11}$$

   and assume

$$\sum_{n=1}^{\infty} \int_{G'_{n,\ell}} \frac{1}{\gamma^{1/2}(\tilde{u})} d\tilde{u} =: \Gamma < \infty. \tag{12}$$

The algorithm for perfect simulation form X is developed for all distributions $\mathcal{L}(X)$ that satisfy the conditions mentioned above.

Observe that the third condition restricts the admissible Skorohod representations. It is possible to extend our approximations and exact simulation algorithm to selected examples that are not locally affine on the cuts $G_{n,\ell}(\tilde{u})$, e.g., to the perpetuities mentioned in the introduction, where we have $K = 1$ and $A_1 = U^a$ for $a > 0$ and a uniform $[0, 1]$ distributed U. Presenting these generalizations would add little of substance to the paper. Note that one can find Skorohod representations that satisfy our third conditions even for non-affine functions of a uniform U. For example, for $A_1 = U^a$ with $a = 1/d$ for some $d \in \mathbb{N}$ we have the distributional identity $U^a \sim \max\{U_1, \ldots, U_d\}$, where the U_i's are independent uniform $[0, 1]$ random variables.

Throughout the following notations are used: X is the in $\mathcal{M}_{0,2}$ unique fixed-point of (4). By ϕ, μ, F, w its Fourier transform, distribution, distribution function, and density respectively are denoted. By H_n we denote the n-th harmonic number $H_n = \sum_{i=1}^{n} 1/i$.

2.2 Examples

The examples of limit laws of QUICKSORT cost measures and internal path lengths of random search trees fit into our setting with

$$g(v) = \kappa' \bar{g}(v) + \kappa \left(\sum_{r=1}^{K-1} (v_r \ln v_r) + \left(1 - \sum_{r=1}^{K-1} v_r\right) \ln \left(1 - \sum_{r=1}^{K-1} v_r\right)\right) \tag{13}$$

where $\kappa, \kappa' > 0$ are normalization constants and $\bar{g}(v)$ is either 1 or v or $v(1 - v)$ depending on the application. We treat the cases $\bar{g}(v) = 1$ or $= v$, the third case can be covered with slight modifications. We have

$$\mathrm{Hess}(g; v)_{ij} = \kappa \left(\frac{1}{v_K} + \delta_{ij} \frac{1}{v_i} \right)$$

with $v_K = 1 - \sum_{r=1}^{K-1} v_r$ and δ_{ij} denoting Kronecker's symbol. Using the relation $\sum_{r=1}^{K} \frac{\partial f_r}{\partial u_l} = 0$ we obtain for all $1 \leq l \leq d$:

$$\left\langle \frac{\partial f_r}{\partial u_l}, \mathrm{Hess}(g; f(\cdot)) \frac{\partial f_r}{\partial u_l} \right\rangle = \kappa \sum_{r=1}^{K} \frac{1}{f_r} \left(\frac{\partial f_r}{\partial u_l} \right)^2 .$$

We proceed by recalling the equations (4) for the limit laws of the internal path lengths of random m-ary search trees, median of $2k + 1$ search trees, and quadtrees and give choices for the quantities $\Gamma, s, p_0, D_1, D_2, p_1, M_p$ in (5)-(7),(12). For small parameters m, k, d these fixed-point equations, which define these limit laws, are presented in Figure 1.

2.2.1 m-ary search tree

For this limit distribution derived in [30] we have $K = m \geq 2$, $d = m - 1$, $\bar{g}(v) = 1$, $\kappa' = 1, \kappa = (H_m - 1)^{-1}$ and

$$(f_1, \ldots, f_m)(u) = (u_{(1)}, u_{(2)} - u_{(1)}, \ldots, 1 - u_{(m-1)}), \tag{14}$$

(i) QUICKSORT: Comparisons

$$X \sim UX^{(1)} + (1-U)X^{(2)} + \mathcal{E}(U),$$

$$\mathcal{E}(U) = 1 + 2(U\ln(U) + (1-U)\ln(1-U)).$$

(ii) ternary search tree

$$X \sim U_{(1)}X^{(1)} + (U_{(2)} - U_{(1)})X^{(2)} + (1-U_{(2)})X^{(3)} + \mathcal{E}(U),$$

$$\begin{aligned}\mathcal{E}(U) = 1 + \frac{6}{5}\Big(&U_{(1)}\ln(U_{(1)}) + (U_{(2)} - U_{(1)})\ln(U_{(2)} - U_{(1)}) \\ &+ (1-U_{(2)})\ln(1-U_{(2)})\Big).\end{aligned}$$

(iii) median of 3 search tree

$$X \sim \mathrm{med}(U_1,U_2,U_3)X^{(1)} + (1-\mathrm{med}(U_1,U_2,U_3))X^{(2)} + \mathcal{E}(U),$$

$$\begin{aligned}\mathcal{E}(U) = 1 + \frac{12}{7}\Big(&\mathrm{med}(U_1,U_2,U_3)\ln(\mathrm{med}(U_1,U_2,U_3)) \\ &+ (1-\mathrm{med}(U_1,U_2,U_3))\ln(1-\mathrm{med}(U_1,U_2,U_3))\Big).\end{aligned}$$

(iv) 2-dimensional quadtree

$$\begin{aligned}X \sim\ & U_1U_2X^{(1)} + U_1(1-U_2)X^{(2)} + (1-U_1)U_2X^{(3)} \\ &+ (1-U_1)(1-U_2)X^{(4)} + \mathcal{E}(U),\end{aligned}$$

$$\begin{aligned}\mathcal{E}(U) = 1 &+ U_1U_2\ln(U_1U_2) + U_1(1-U_2)\ln(U_1(1-U_2)) \\ &+ (1-U_1)U_2\ln((1-U_1)U_2) \\ &+ (1-U_1)(1-U_2)\ln((1-U_1)(1-U_2)).\end{aligned}$$

Figure 1: *Fixed-point equations for limit distributions of (i) the number of comparisons of* QUICKSORT *and the internal path lengths of (ii) random ternary search trees, (iii) random median of* 3 *search trees and (iv) random* 2*-dimensional quadtrees. med*(U_1, U_2, U_3) *and* $U_{(1)}, U_{(2)}$ *denote the median and the order statistics of* U_1, U_2, U_3 *and* U_1, U_2 *respectively.*

where $u_{(1)}, \ldots, u_{(m-1)}$ denote the order statistics of the components of $u \in [0,1]^{m-1}$. The conditions (5)-(7),(12) are satisfied as follows:
Ad (5): Note that

$$\begin{aligned}
\lambda^d\left(\bigcap_{\substack{r=1\\ r\neq j}}^{K}\{f_r \le c/t\}\right) &\le \lambda^d(\{f_r \le c/t\}) \\
&= \int_0^{c/y}(m-1)(1-x)^{m-2}dx \\
&= \left(1-\left(1-\frac{c}{t}\right)^{m-1}\right) \\
&\le (m-1)ct^{-1}.
\end{aligned}$$

Thus we choose $s := 1, D_1(c) := m(m-1)c$.
Ad (6): We have

$$\begin{aligned}
\int_{\{f_r \ge c/t\}} f_r^{-q}(u)du &= \int_{c/t}^{1} x^{-p}(m-1)(1-x)^{m-2}dx \\
&\le (m-1)\int_{c/t}^{1} x^{-p}dx \\
&\le \frac{m-1}{c^{p-1}(p-1)}\frac{1}{t^{1-p}},
\end{aligned}$$

for $p > 1$ which gives

$$p_0 := 1, \quad D_2(p,c) := \frac{m-1}{c^{p-1}(p-1)}.$$

Ad (7): Using that the joint distribution of the spacings $(U_{(1)}, U_{(2)}-U_{(1)}, \ldots, 1-U_{(m-1)})$ is Dirichlet $\mathrm{D}(1,\ldots,1)$ on the Simplex $\sum_{i=1}^m v_i = 1$ we obtain with the $(m-1)$-dimensional Hausdorff measure $\mathcal{H}$

$$\begin{aligned}
\int_{[0,1]^{m-1}}\prod_{i=1}^m f_i^{-p}(u)du &= (m-1)!\int_{\sum v_i=1}\prod_{i=1}^m v_i^{-p}d\mathcal{H}(v) \\
&= (m-1)!\frac{(\Gamma(1-p))^m}{\Gamma(m(1-p))}\int_{\sum v_i=1}\frac{\Gamma((m-1)(1-p))}{\Gamma(1-p)^m}\prod_{i=1}^m v_i^{-p}d\mathcal{H}(v) \\
&= (m-1)!\frac{(\Gamma(1-p))^m}{\Gamma(m(1-p))}
\end{aligned}$$

for $0 < p < 1$, the last integrand being the density of the Dirichlet $\mathrm{D}(1-p,\ldots,1-p)$ distribution. We obtain

$$p_1 := 1, \quad M_p := (m-1)!\frac{(\Gamma(1-p))^m}{\Gamma(m(1-p))}.$$

Ad (12): With the notation $u = [u_1, \tilde{u}]$ defined in (9) with $\tilde{u} \in [0,1]^{m-2}$ and $\tilde{u}_{(0)} := 0, \tilde{u}_{(m-1)} := 1$ on $\{\tilde{u}_{(j-1)} < u_1 < \tilde{u}_{(j)}\}$ we have

$$\frac{\partial f_r}{\partial u_1} = \begin{cases} 1 & r = j \\ -1 & r = j+1 \\ 0 & \text{otherwise} \end{cases}$$

for $j = 1, \dots, m-1$. This implies

$$\begin{aligned}
\kappa \sum_{r=1}^{m} \frac{1}{f_r} \left(\frac{\partial f_r}{\partial u_1} \right)^2 &= \kappa \sum_{j=1}^{m-1} \mathbf{1}_{\{\tilde{u}_{(j-1)} < u_1 < \tilde{u}_{(j)}\}} \sum_{r=1}^{m} \frac{1}{f_r} \left(\frac{\partial f_r}{\partial u_1} \right)^2 \\
&= \kappa \sum_{j=1}^{m-1} \mathbf{1}_{\{\tilde{u}_{(j-1)} < u_1 < \tilde{u}_{(j)}\}} \left(\frac{1}{u_1 - \tilde{u}_{(j-1)}} + \frac{1}{\tilde{u}_{(j)} - u_1} \right).
\end{aligned}$$

Note that

$$\inf_{\tilde{u}_{(j-1)} < u_1 < \tilde{u}_{(j)}} \left(\frac{1}{u_1 - \tilde{u}_{(j-1)}} + \frac{1}{\tilde{u}_{(j)} - u_1} \right) \geq \frac{4}{\tilde{u}_{(j)} - \tilde{u}_{(j-1)}},$$

thus, noting that a spacing between $m-1$ independent uniform $[0,1]$ random variables is beta$(1, m-2)$ distributed, we have

$$\begin{aligned}
\Gamma = \int_{[0,1]^{m-2}} \frac{1}{\gamma^{1/2}(\tilde{u})} d\tilde{u} &\leq \sum_{j=1}^{m-1} \int_{[0,1]^{m-2}} \frac{1}{2\sqrt{\kappa}} (\tilde{u}_{(j)} - \tilde{u}_{(j-1)})^{1/2} d\tilde{u} \\
&= \frac{m-1}{2\sqrt{\kappa}} \int_0^1 \sqrt{x}(1-x)^{m-3} dx \\
&= \frac{(m-1)(m-2)}{2\sqrt{\kappa}} B(3/2, m-2) \\
&= \frac{\sqrt{\pi}}{4\sqrt{\kappa}} \frac{\Gamma(m)}{\Gamma(m-1/2)}.
\end{aligned}$$

2.2.2 Median of $2k+1$ search tree

For this limit distribution derived in [36] we have $K = 2$, $d = 2k+1$, $\bar{g}(v) = 1$, $\kappa' = 1$, $\kappa = (H_{2k+2} - H_{k+1})^{-1}$ and $(f_1, f_2)(u) = (\mathrm{med}(u), 1 - \mathrm{med}(u))$, where $\mathrm{med}(u)$ denotes the median of the components of u.
Ad (5): Using that the median of $2k+1$ independent uniform $[0,1]$ random variables is beta$(k+1, k+1)$ distributed we find

$$\begin{aligned}
\lambda^d \left(\bigcap_{r \neq j} \{f_r \leq c/t\} \right) &\leq \lambda^d(\{f_r \leq c/t\}) \\
&= \int_0^{c/y} \frac{x^k (1-x)^k}{B(k+1, k+1)} dx \\
&\leq \frac{c^{k+1}}{(k+1)B(k+1, k+1)} t^{-(k+1)},
\end{aligned}$$

so we can choose

$$s := k+1, \quad D_1(c) = \frac{2c^{k+1}}{(k+1)B(k+1, k+1)}.$$

Ad (6): Observe that

$$\begin{aligned}\int_{\{f_r \geq c/t\}} f_r^{-q}(u)du &= \int_{c/t}^1 x^{-p}\frac{x^k(1-x)^k}{B(k+1,k+1)}dx\\ &\leq \frac{1}{B(k+1,k+1)}\int_{c/t}^1 x^{k-p}dx\\ &= \frac{1}{(k+1-p)B(k+1,k+1)}\left(1-\left(\frac{c}{t}\right)^{k+1-p}\right)\\ &\leq \frac{c^{k+1-p}}{(k+1-p)B(k+1,k+1)}\frac{1}{t^{k+1-p}},\end{aligned}$$

for all $p > k+1$. Thus we choose

$$p_0 := k+1, \quad D_2(p,c) := \frac{2c^{k+1-p}}{(k+1-p)B(k+1,k+1)}.$$

Ad (7): Evaluating a beta integral we easily obtain

$$p_1 := k+1, \quad M_p := \frac{B(k+1-p,k+1-p)}{B(k+1,k+1)}.$$

Ad (12): Denote

$$G_n = \{u \in [0,1]^{2k+1} : u_n = \mathrm{med}(u)\}$$

for $n = 1,\dots,2k+1$. Then with the notation in (8), (10) we obtain on $G'_{n,n}$

$$\gamma(\tilde{u}) = \inf_{u_n \in G_{n,n}(\tilde{u})} \kappa \sum_{r=1}^{2} \frac{1}{f_r}\left(\frac{\partial f_r}{\partial u_n}\right)^2 = \inf_{u_n \in G_{n,n}(\tilde{u})} \kappa\left(\frac{1}{u_n}+\frac{1}{1-u_n}\right) \geq 4\kappa,$$

which implies

$$\Gamma = \sum_{n=1}^{2k+1}\int_{G'_{n,n}} \frac{1}{c^{1/2}(\tilde{u})}d\tilde{u} \leq \frac{2k+1}{2\sqrt{\kappa}}.$$

2.2.3 Quadtree

For this limit distribution derived in [30] we have $d \geq 2$, the dimension of the quadtree, $K = 2^d$, $\bar{g}(v) = 1$, $\kappa' = 1$, $\kappa = 2/d$, and $(f_1,\dots,f_{2^d})(u)$ is the vector of the volumes of the quadrants in $[0,1]^d$ generated by the point u, see [30] for a formal definition.
For (5),(6) first note that the density φ_d and the distribution function F_d of the product of d independent unif$[0,1]$ distributed random variables is given by

$$\varphi_d(x) = \frac{1}{(d-1)!}\left(\ln\frac{1}{x}\right)^{d-1}, \quad F_d(x) = \sum_{j=1}^{d}\frac{1}{(j-1)!}\left(\ln\frac{1}{x}\right)^{j-1} x.$$

Furthermore we use the inequality

$$\forall\, \varepsilon > 0\, \forall\, d \geq 1\, \forall\, x \geq 1 : (\ln x)^d \leq \frac{d!}{\varepsilon^d}x^{\varepsilon}. \tag{15}$$

Ad (5): Using the inequality (15) with $\varepsilon = 1/d$ we obtain

$$\begin{aligned}
\lambda^d\left(\bigcap_{r\neq j}\{f_r \le c/t\}\right) &\le \lambda^d(\{f_r \le c/t\}) \\
&= \sum_{j=1}^{d} \frac{1}{(j-1)!}\left(\ln\frac{t}{c}\right)^{j-1}\frac{c}{t} \\
&\le \frac{c}{t}\sum_{j=1}^{d}\frac{1}{(j-1)!}\frac{(j-1)!}{(1/d)^{j-1}}\left(\frac{t}{c}\right)^{1/d} \\
&= c^{1-1/d}\frac{d^d-1}{d-1}t^{-(1-1/d)},
\end{aligned}$$

thus we set

$$s := 1 - 1/d, \quad D_1(c) = 2^d\frac{d^d-1}{d-1}c^{1-1/d}.$$

Ad (6): Using (15) with $\varepsilon = 1/d$, we observe the following:

$$\begin{aligned}
\int_{\{f_r\ge c/t\}} f_r^{-q}(u)du &= \int_{c/t}^{1} x^{-p}\frac{1}{(d-1)!}\left(\ln\frac{1}{x}\right)^{d-1} dx \\
&\le \frac{1}{(d-1)1}\int_{c/t}^{1} x^{-p}\frac{(d-1)!}{(1/d)^{d-1}}\left(\frac{1}{x}\right)^{1/d} dx \\
&= d^{d-1}\int_{c/t}^{1} x^{-p-1/d}dx \\
&= \frac{d^{d-1}}{1-p-1/d}\left(1-\left(\frac{c}{t}\right)^{1-p-1/d}\right) \\
&\le d^{d-1}\frac{c^{1-p-1/d}}{p+1/d-1}\frac{1}{t^{s-p}}.
\end{aligned}$$

We choose

$$p_0 := 1-\frac{1}{d}, \quad D_2(p,c) = 2^d d^{d-1}\frac{c^{1-p-1/d}}{p+1/d-1}.$$

Ad (7): We easily obtain

$$p_1 := 2^{-(d-1)}, \quad M_p := (B(1-p2^{d-1}, 1-p2^{d-1}))^d.$$

Ad (12): With some calculations involving the structure of the volumes generated by u, we note the following:

$$\kappa\sum_{r=1}^{2^d}\frac{1}{f_r}\left(\frac{\partial f_r}{\partial u_1}\right)^2 = \kappa\left(\frac{1}{u_1}+\frac{1}{1-u_1}\right) \ge \frac{8}{d},$$

which implies $\Gamma \le \sqrt{d/8}$.

2.2.4 Other examples

The limit distribution of the number of key comparisons of QUICKSORT is identical with the limit distribution of the internal path length of a random binary search tree. This is covered by m-ary search trees with $m = 2$ or median of $2k+1$ search trees with $k = 0$. The internal path length for random recursive trees (see [8, 26]) is covered with $K = 2$, $d = 1$, $\bar{g}(v) = v$ $\kappa' = 1$, $\kappa = 1$, and $(f_1, f_2)(u) = (u, 1-u)$. The choices can be made as the ones for the random binary search tree since $\bar{g}'' = 0$. Only the different value of κ has to be adjusted. The limit law for the number of key exchanges of QUICKSORT (see [22, 29]) involves the function $\bar{g}(v) = v(1-v)$ and can be treated with appropriate adjustments.

3 Densities and dominating curve

First we show that $\mathcal{L}(X)$, given in section 2.1, has an infinite differentiable density w, and that the density and all its derivatives are bounded. For this we use the approach of Fill and Janson [11]. The conditions (5)-(7),(12) are tailored to approach this method. Then a dominating integrable curve for w needed for the rejection method follows without work.

3.1 Properties of the density

Following Fill and Janson [11] we define $c_p \in [0, \infty]$ for $p > 0$ to be the smallest constants such that

$$|\phi(t)| \leq c_p |t|^{-p} \quad \text{for all } t \in \mathbb{R}.$$

Note that the sets $\{c \geq 0 : |\phi(t)| \leq c|t|^{-p} \text{ for all } t \in \mathbb{R}\}$, $p > 0$, contain their infima. The aim is show $c_p < \infty$ for p as large as possible with explicit bounds on c_p. If $c_p < \infty$ for all $p > 0$ it follows by the Fourier inversion formula that w is infinite differentiable and that all its derivatives are bounded. The following Theorem implies $c_p < \infty$ for all $p > 0$ in our situation:

Theorem 3.1 *We have with p_1, M_p as in (7), D_1, s, p_0, D_2 as in (5),(6), Γ as in (12),*

$$c_{1/2} \leq \sqrt{32}\,\Gamma \tag{16}$$

$$c_{Kp} \leq M_p c_p^K, \quad 0 < p < p_1, \tag{17}$$

$$c_{p+s} \leq \left(K^p c_p D_1(c_p^{1/p}) + (K-1)K^p c_p^2 D_2(p, c_p^{1/p})\right) \vee \left(K c_p^{1/p}\right)^{-(p+s)}, \tag{18}$$

for $p > p_0$.

Together with the trivial inequality $c_p \leq c_q^{p/q}$ for all $0 < p \leq q$ we obtain $c_p < \infty$ for all $p > 0$ by iterated, appropriate application of (16)-(18). First recall the following Lemma due to Fill and Janson [11]:

Lemma 3.2 *Let $z : [a, b] \to \mathbb{R}$ be twice continuously differentiable with $z'' \geq \gamma > 0$ or $z'' \leq -\gamma < 0$ on (a, b). Then*

$$\left| \int_a^b \exp(itz(x))\, dx \right| \leq \frac{\sqrt{32}}{\gamma^{1/2}} |t|^{-1/2}, \quad t \in \mathbb{R}. \tag{19}$$

Proof: Combine Lemmas 2.2 and 2.3 in Fill and Janson [11]. ■

Estimates for exponential integrals as in Lemma 3.2 are well-known in analytic number theory. The $\sqrt{32}$ may be replaced by 8 (Tenenbaum [42, Lemma 4.4]).

Proof of Theorem 3.1: Ad (16): With $W(u) := \sum_{r=1}^{K-1} x_r f_r(u) + x_K(1 - \sum_{r=1}^{K-1} f_r(u)) + g(f(u))$ for $x_1, \ldots, x_K \in \mathbb{R}$ we obtain by conditioning on the fixed-points,

$$|\phi(t)| \le \int_{\mathbb{R}^K} \left| \int_{[0,1]^d} \exp(itW(u))\, du \right| d(\mu \otimes \cdots \otimes \mu)(x_1, \ldots, x_K). \tag{20}$$

It is sufficient to obtain a bound for the inner integral. We have

$$\left| \int_{[0,1]^d} \exp(itW(u))\, du \right| \le \sum_{n=1}^{\infty} \int_{G'_{n,l}} \left| \int_{G_{n,l}(\tilde{u})} \exp(itW(u))\, du_l \right| d\tilde{u}. \tag{21}$$

For the inner integral note that $u_l \mapsto f_r([u_l, \tilde{u}])$ are affine for all $r = 1, \ldots, K$. On $G_{n,l} \times \{\tilde{u}\}$ we have therefore $\partial^2 f / \partial u_l^2 = 0$. This yields with the notation $x^- := (x_1 - x_K, \ldots, x_{K-1} - x_K)$

$$\begin{aligned}
\frac{\partial W}{\partial u_l} &= \left\langle x^- - (\nabla g) \circ f, \frac{\partial f}{\partial u_l} \right\rangle, \\
\left| \frac{\partial^2 W}{\partial u_l^2} \right| &= \left| \left\langle x^- - (\nabla g) \circ f, \frac{\partial^2 f}{\partial u_l^2} \right\rangle + \left\langle \frac{\partial f}{\partial u_l}, \mathrm{Hess}(g; f) \frac{\partial f}{\partial u_l} \right\rangle \right| \\
&= \left| \left\langle \frac{\partial f}{\partial u_l}, \mathrm{Hess}(g; f) \frac{\partial f}{\partial u_l} \right\rangle \right| \\
&\ge \gamma,
\end{aligned}$$

with γ defined in (11). Application of Lemma 3.2 implies

$$\left| \int_{G'_{n,l}(\tilde{u})} \exp(itW(u))\, du_l \right| \le \frac{\sqrt{32}}{\gamma^{1/2}} |t|^{-1/2}.$$

and with the outer integrations and summation in (20), (21), and with (12) it follows that

$$|\phi(t)| \le \sqrt{32}\Gamma |t|^{-1/2},$$

thus $c_{1/2} \le \sqrt{32}\Gamma$.
Ad (17): For $0 < p < p_1$, using (7), we have

$$|\phi(t)| \le \int_{[0,1]^d} \prod_{r=1}^{K} |\phi(f_r(u)t)|\, du \le \int_{[0,1]^d} \prod_{r=1}^{K} \frac{c_p}{f_r^p(u)|t|^p}\, du \le c_p^K M_p |t|^{-Kp}.$$

Ad (18): We assume $c_p < \infty$ for a $p > p_0$ and $t > K c_p^{1/p}$; in the case $0 < t < K c_p^{1/p}$ we have trivially $|\phi(t)| \le c_{p+s} |t|^{-(p+s)}$ since $|\phi(t)| \le 1$. For $t > K c_p^{1/p}$ we cannot have $f_r \le c_p^{1/p}/t$ for all $r = 1, \ldots, K$ since $\sum f_r = 1$. Thus we have only the two cases "all but one f_r are $\le c_p^{1/p}/t$" and "at least two f_r, f_q are $> c_p^{1/p}/t$". This yields

$$[0,1]^d = \left(\bigcup_{j=1}^{K} \bigcap_{\substack{r=1 \\ r \ne j}}^{K} \left\{ f_r \le \frac{c_p^{1/p}}{t} \right\} \right) \cup \left(\bigcup_{\substack{r,j=1 \\ r \ne j}}^{K} \left\{ f_r > \frac{c_p^{1/p}}{t}, f_j > \frac{c_p^{1/p}}{t} \right\} \right)$$

We denote the first of these two sets by B_1. The second one we intersect with $[0,1]^d = \cup_{q=1}^K \{f_q \geq 1/K\}$. It is easily seen that the second set is then a subset of

$$B_2 := \bigcup_{\substack{q,r=1\\ q\neq r}}^K \left\{ f_q \geq \frac{1}{K}, f_r > \frac{c_p^{1/p}}{t} \right\},$$

thus $[0,1]^d = B_1 \cup B_2$. Therefore, we have

$$|\phi(t)| \leq \int_{[0,1]^d} \prod_{r=1}^K \min\left\{ \frac{c_p}{(f_r(u)|t|)^p}, 1 \right\} du \leq \int_{B_1} + \int_{B_2} =: I + II.$$

For the estimate of I we note that $f_j(u) \geq 1-(K-1)c_p^{1/p}/t$ on $\cap_{r\neq j}\{f_r \leq c_p^{1/p}/t\}$, so that we obtain $f_j(u) \geq 1/K$ on this set. With (5), this yields

$$\begin{aligned} I &\leq \sum_{j=1}^K \int_{\cap_{r\neq j}\{f_r \leq c_p^{1/p}/t\}} \frac{c_p}{(f_j(u)t)^p} du \\ &\leq c_p K^p t^{-p} \sum_{j=1}^K \lambda^d \left(\bigcap_{\substack{r=1\\ r\neq j}}^K \left\{ f_r \leq \frac{c_p^{1/p}}{t} \right\} \right) \\ &\leq c_p K^p D_1(c_p^{1/p}) t^{-(p+s)}. \end{aligned}$$

For II we estimate first

$$\int_{\{f_q \geq 1/K, f_r > c_p^{1/p}/t\}} \frac{c_p^2}{(f_q(u)f_r(u))^p t^2}\, du \leq c_p^2 K^p t^{-2p} \int_{\{f_r > c_p^{1/p}/t\}} f_r^{-p}(u)\, du.$$

This yields, using (6),

$$\begin{aligned} II &\leq (K-1)c_p^2 K^p t^{-2p} \sum_{r=1}^K \int_{\{f_r > c_p^{1/p}/t\}} f_r^{-p}(u)\, du \\ &\leq (K-1)c_p^2 K^2 D_2(c_p^{1/p}) t^{-(p+s)}. \end{aligned}$$

The assertion follows. ■

3.2 The dominating curve

For a rejection algorithm a dominating, integrable curve q for the density w to be sampled from is necessary, such that from the distribution with density $q/\|q\|_1$ it is easy to sample. If Lipschitz- and moment-information on w is available a curve q can be constructed on the basis of Theorem 3.3 and Theorem 3.5 in Devroye [4, p. 315, p. 320]. For this we denote by $K_1, K_2, K_3 > 0$ constants with

$$\|w\|_\infty \leq K_1, \quad \|w'\|_\infty \leq K_2, \quad \mathbb{E}\, X^4 \leq K_3. \tag{22}$$

The existence of moments of all orders of X follows since the Laplace transform of X is finite in a neighborhood of 0, see Rösler [35]. Then a dominating, integrable curve for w is given by

$$q(x) := \min\left\{ K_1, \sqrt{2K_2K_3}x^{-2} \right\}, \quad x \in \mathbb{R}. \tag{23}$$

This follows from the general inequality $w(x) \le (2K_2 \min\{F(x), 1-F(x)\})^{1/2}$, cf. Theorem 3.5 in Devroye [4], where F is the distribution function of X, and, by Markov's inequality, $\min\{F(x), 1-F(x)\} \le \mathbb{E}\, X^4/x^4$.

A random variate with density $q/\|q\|_1$ is given by

$$S\frac{(2K_2K_3)^{1/4}}{K_1^{1/2}}\frac{U_1}{U_2}, \tag{24}$$

with U_1, U_2, S being independent, $U_1, U_2 \sim \text{uniform}[0,1]$ and S being an equiprobable random sign, cf. Theorem 3.3 in Devroye [4]. In our situation the following choices for K_1, K_2, K_3 are possible:

Lemma 3.3 *Define ξ as in (2) and $\xi_3 := \sum_{r=1}^K \mathbb{E}\, A_r^3$, $\xi_4 := \sum_{r=1}^K \mathbb{E}\, A_r^4$ and the c_p as in Lemma 3.1. [For a rough estimate ξ_3, ξ_4 may be replaced by ξ]. For the density w of X the inequalities in (22) are satisfied with*

$$\begin{aligned} K_1 &:= \frac{pc_p^{1/p}}{\pi(p-1)}, \quad p>1 \\ K_2 &:= \frac{1}{\pi}\left(c_p^{1/p} + \frac{c_p^{2/p}}{p-2}\right), \quad p>2 \\ K_3 &:= \frac{\|g\|_\infty^4}{1-\xi_4}\left(1 + \frac{1}{1-\xi} + \frac{1}{1-\xi_3} + \frac{K}{(1-\xi)(1-\xi_3)} + \frac{K(K-1)}{(1-\xi)^2}\right). \end{aligned}$$

Moreover we have

$$\|w''\|_\infty \le K_4 := \frac{1}{\pi}\left(c_p^{1/p} + \frac{c_p^{3/p}}{p-3}\right), \quad p>3.$$

Proof: By the Fourier inversion formula the k-th derivative $w^{(k)}$ satisfies

$$\|w^{(k)}\|_\infty \le \frac{1}{2\pi}\int_{-\infty}^{\infty} |t|^k|\phi(t)|dt, \quad k \in \mathbb{N}_0.$$

Splitting the domain of integration into $[-c_p^{1/p}, c_p^{1/p}]$ and its complement and using $|\varphi(t)| \le c_p|t|^{-p}$ we obtain

$$\|w^{(k)}\|_\infty \le \frac{1}{\pi}\left(c_p^{1/p} + \frac{c_p^{(k+1)/p}}{p-(k+1)}\right), \quad p>k+1.$$

This gives the choices for K_1, K_2 and the estimate for $\|w''\|_\infty$.

The moments of X can be calculated or estimated form the fixed-point equation. Using the independence assumptions and $\mathbb{E}\, X = 0$ we obtain with $|b| \le \|g\|_\infty$ and $|A_r| \le 1$ first $\mathbb{E}\, X^2 = \mathbb{E}\, X^2 \sum_{r=1}^K \mathbb{E}\, A_r^2 + \mathbb{E}\, b^2$, thus

$$\mathbb{E}\, X^2 \le \frac{\|g\|_\infty^2}{1-\xi}.$$

Then we have

$$\begin{aligned} \mathbb{E}\, X^3 &= \mathbb{E}\, b^3 + \mathbb{E}\, X^3 \sum_{r=1}^K \mathbb{E}\, A_r^3 + \mathbb{E}\, X^2 \sum_{r=1}^K \mathbb{E}\,[bA_r^2] \\ &\le \|g\|_\infty^3 + K\|g\|_\infty \mathbb{E}\, X^2 + \mathbb{E}\, X^3 \xi_3, \end{aligned}$$

thus

$$\mathbb{E}\, X^3 \leq \frac{\|g\|_\infty^3}{1-\xi_3}\left(1+\frac{K}{1-\xi}\right).$$

Expanding and estimating similarly the fourth moment of X leads to K_3. ■

Better bounds on K_1, K_2 are possible by refined decomposition of the range of integration and by better estimates of the c_p, see Fill and Janson [11].

In the examples on internal path lengths of m-ary search trees, median of $2k+1$ search trees and quadtrees ξ is given in (49), (50), and (51) respectively, $\|g\|_\infty$ is easily estimated since $|x\ln(x)| \leq 1/e$ for all $x \in [0,1]$.

4 Approximation of the density

As in section 3 the general part valid for all fixed-points as defined in section 2.1 is separated from the applications.

4.1 The approximating sequence

We assume that discretizations $A_r^{(n)}$ of A_r and $b^{(n)}$ of b are given satisfying conditions noted below. We define then discrete probability distributions $\mathcal{L}(X_n)$ for $n \geq 0$ by $X_0 := 0$ and for $n \geq 1$ recursively by

$$\widetilde{X}_n := \sum_{r=1}^K A_r^{(n)} X_{n-1}^{(r)} + b^{(n)}, \tag{25}$$

$$\mathcal{L}(X_n) := \mathcal{L}(\langle \widetilde{X}_n \rangle), \tag{26}$$

where $(A_1^{(n)}, \dots, A_K^{(n)}, b^{(n)}), X_{n-1}^{(1)}, \dots, X_{n-1}^{(K)}$ are independent with $X_{n-1}^{(r)} \sim X_{n-1}$ and $\langle \cdot \rangle$ denotes a further discretization step. We assume that we have the following pointwise accuracies of approximation:

$$\sum_{r=1}^K |A_r^{(n)} - A_r| \quad \leq \quad R_\Sigma(n), \tag{27}$$

$$\sum_{r=1}^K |A_r^{(n)} - A_r|^2 \quad \leq \quad R_\Sigma^{(2)}(n), \tag{28}$$

$$|b^{(n)} - b| \quad \leq \quad R_b(n), \tag{29}$$

$$|\widetilde{X}_n - \langle \widetilde{X}_n \rangle| \quad \leq \quad R_X(n), \tag{30}$$

$$\left| \sum_{r=1}^K \mathbb{E}\, A_r^{(n)} \right| \quad \leq \quad 1 - R_\Delta(n) \tag{31}$$

where $R_\Sigma, R_\Sigma^{(2)}, R_b, R_X, R_\Delta$ are functions on $\mathbb{N}$. Furthermore we denote by $C_A, C_A', \xi(n) \geq 0$ constants with

$$\sum_{r=1}^K \|A_r^{(n)}\|_2 \leq C_A, \qquad \sum_{\substack{r,s=1 \\ r \neq s}}^K \mathbb{E}\,[A_r^{(n)} A_s^{(n)}] \leq C_A', \qquad n \geq 1, \tag{32}$$

and

$$\xi^2(n) := \sum_{r=1}^K \|A_r^{(n)}\|_2^2, \tag{33}$$

where we recall that $\|X\|_2 = \sqrt{\mathbb{E}\, X^2}$. Then using $\mathbb{E}\, b = 0$ and (29) the means of X_n are estimated by

$$\begin{aligned}
|\,\mathbb{E}\, X_n| &\leq |\,\mathbb{E}\, \widetilde{X}_n| + |\,\mathbb{E}\,[X_n - \widetilde{X}_n]| \\
&\leq \left|\sum_{r=1}^K \mathbb{E}\, A_r^{(n)}\, \mathbb{E}\, X_{n-1}\right| + |\,\mathbb{E}\, b^{(n)}| + R_X(n) \\
&\leq \left|\sum_{r=1}^K \mathbb{E}\, A_r^{(n)}\right| |\,\mathbb{E}\, X_{n-1}| + R_b(n) + R_X(n) \\
&\leq \sum_{j=1}^n \left(\prod_{i=j}^{n-1} (1 - R_\Delta(i+1))\right) (R_b(j) + R_X(j)) =: M(n).
\end{aligned} \tag{34}$$

We start with the estimate

$$\begin{aligned}
\ell_2(X_n, X) &\leq \ell_2(X_n, \widetilde{X}_n) + \ell_2(\widetilde{X}_n, X) \\
&\leq R_X(n) + \ell_2(\widetilde{X}_n, X).
\end{aligned}$$

Using appropriate optimal couplings as it is common in the application of the contraction method, see, e.g., Rösler [36], we obtain

$$\begin{aligned}
\ell_2^2(\widetilde{X}_n, X) &\leq \left\|\sum_{r=1}^K A_r^{(n)} X_{n-1}^{(r)} + b^{(n)} - \sum_{r=1}^K A_r X^{(r)} - b\right\|_2^2 \\
&\leq \mathbb{E} \sum_{r=1}^K \left(A_r^{(n)} X_{n-1}^{(r)} - A_r X^{(r)}\right)^2 + \mathbb{E}\, (b^{(n)} - b)^2 \qquad (35) \\
&\quad + 2\, \mathbb{E} \sum_{r=1}^K \left(A_r^{(n)} X_{n-1}^{(r)} - A_r X^{(r)}\right) (b^{(n)} - b) \\
&\quad + \mathbb{E} \sum_{\substack{r,s=1 \\ r \neq s}}^K \left(A_r^{(n)} X_{n-1}^{(r)} - A_r X^{(r)}\right) \left(A_s^{(n)} X_{n-1}^{(s)} - A_s X^{(s)}\right) \\
&=: I + II + III + IV.
\end{aligned}$$

We have $II \leq R_b^2(n)$, and

$$\begin{aligned}
III &= 2\, \mathbb{E} \sum_{r=1}^K \left(A_r^{(n)} X_{n-1}^{(r)} - A_r X^{(r)}\right) (b^{(n)} - b) \\
&= 2\, \mathbb{E} \sum_{r=1}^K A_r^{(n)} X_{n-1}^{(r)} (b^{(n)} - b) \\
&\leq 2 \sum_{r=1}^K \|A_r^{(n)}\|_2 \|b^{(n)} - b\|_2\, \mathbb{E}\, X_{n-1} \\
&\leq 2 C_A R_b(n) M(n-1).
\end{aligned}$$

Analogously

$$\begin{aligned}
IV &= \mathbb{E}\sum_{\substack{r,s=1\\r\neq s}}^{K}\left(A_r^{(n)}X_{n-1}^{(r)}-A_rX^{(r)}\right)\left(A_s^{(n)}X_{n-1}^{(s)}-A_sX^{(s)}\right)\\
&= \sum_{\substack{r,s=1\\r\neq s}}^{K}\mathbb{E}\,[A_r^{(n)}A_s^{(n)}]\,\mathbb{E}\,[X_{n-1}]^2\\
&\leq C_A'M^2(n-1).
\end{aligned}$$

Finally, by the Cauchy-Schwarz inequality

$$\begin{aligned}
I &= \mathbb{E}\sum_{r=1}^{K}\left(A_r^{(n)}X_{n-1}^{(r)}-A_rX^{(r)}\right)^2\\
&= \mathbb{E}\sum_{r=1}^{K}\left(A_r^{(n)}(X_{n-1}^{(r)}-X^{(r)})-(A_r^{(n)}-A_r)X^{(r)}\right)^2\\
&= \sum_{r=1}^{K}\Bigg(\mathbb{E}\,(A_r^{(n)})^2\ell_2^2(X_{n-1},X)+\|A_r^{(n)}-A_r\|_2^2\,\mathbb{E}\,X^2\\
&\qquad +2\,\mathbb{E}\,[A_r^{(n)}(A_r^{(n)}-A_r)(X_{n-1}^{(r)}-X^{(r)})X^{(r)}]\Bigg)\\
&\leq \xi^2(n)\ell_2^2(X_{n-1},X)+R_\Sigma^{(2)}(n)\,\mathbb{E}\,X^2\\
&\qquad +2\sum_{r=1}^{K}\|A_r^{(n)}-A_r\|_2\|X^{(r)}\|_2\|A_r^{(n)}(X_{n-1}^{(r)}-X^{(r)})\|_2\\
&= \xi^2(n)\ell_2^2(X_{n-1},X)+R_\Sigma^{(2)}(n)\|X\|_2^2+2(R_\Sigma^{(2)}(n))^{1/2}\|X\|_2C_A\ell_2(X_{n-1},X).
\end{aligned}$$

We denote the prefactors and a constant used later by

$$\begin{aligned}
b_n &:= 2C_A\|X\|_2(R_\Sigma^{(2)}(n))^{1/2},\\
c_n &:= R_b^2(n)+2C_AR_b(n)M(n-1)+C_A'M^2(n-1)+R_\Sigma^{(2)}(n)\|X\|_2^2,\\
d_n &:= \max\left\{b_n/\xi,c_n^{1/2}\right\}.
\end{aligned}$$

Assume that there exists an $\ell\in\mathbb{N}$ such that for all $n\geq\ell$, $\xi(n)\in[\xi/2,(1+\xi)/2]$. Denote $\bar\xi:=(1+\xi)/2$. Then we obtain altogether

$$\begin{aligned}
\ell_2(X_n,X) &\leq R_X(n)+\sqrt{\xi^2(n)\ell_2^2(X_{n-1},X)+b_n\ell_2(X_{n-1},X)+c_n}\\
&\leq R_X(n)+\sqrt{(\xi(n)\ell_2(X_{n-1},X)+d_n)^2}\\
&= R_X(n)+d_n+\xi(n)\ell_2(X_{n-1},X)\\
&\leq \bar\xi^{n-\ell}\ell_2(X_\ell,X)+\sum_{i=0}^{n-1-\ell}\bar\xi^i(R_X(n-i)+d_{n-i})\\
&\leq \bar\xi^n\bar\xi^{-\ell}(\|X\|_2+\|X_\ell\|_2)+\sum_{i=0}^{n-1}\bar\xi^i(R_X(n-i)+d_{n-i}). \qquad (36)
\end{aligned}$$

In order to obtain explicit estimates we have to specify the functions $R_\Sigma, R_\Sigma^{(2)}, R_b, R_X$. We assume that for all $n \geq 1$,

$$R_\Delta(n) \geq \frac{1}{n}, \quad R_\Sigma(n) \leq C_\Sigma \frac{\ln(n)}{n}, \quad R_\Sigma^{(2)}(n) \leq C_\Sigma^{(2)} \frac{\ln(n)}{n^2}, \tag{37}$$

$$R_b(n) \leq \frac{C_b}{n^2}, \quad R_X(n) \leq \frac{C_X}{n^2}, \quad |\xi(n) - \xi| \leq \frac{C_\xi}{n}$$

with constants $C_\Sigma, C_\Sigma^{(2)}, C_b, C_X, C_\xi > 0$ and the contraction factor ξ given in (2).

In order to make the previous estimates explicit we start with two Lemmas:

Lemma 4.1 *For all $n \in \mathbb{N}$ we have*

$$\|X_n\|_\infty \leq Q_n := \begin{cases} (C_X + C_b + \|g\|_\infty)(n + C_\Sigma)\ln(n+1), & \text{if } 0 < C_\Sigma \leq 1, \\ \zeta(\lceil C_\Sigma \rceil)(C_X + C_b + \|g\|_\infty)(n + C_\Sigma)^{\lceil C_\Sigma \rceil}, & \text{if } C_\Sigma > 1, \end{cases}$$

where $\zeta(\cdot)$ denotes the Riemannian ζ-function, $\zeta(s) := \sum_{n \geq 1} n^{-s}$.

Proof: By definition of X_n,

$$\begin{aligned} \|X_n\|_\infty &\leq \|X_n - \tilde{X}_n\|_\infty + \|\tilde{X}_n\|_\infty \\ &\leq R_X(n) + \left\| \sum_{r=1}^K A_r^{(n)} X_{n-1}^{(r)} + b^{(n)} \right\|_\infty \\ &\leq R_X(n) + R_b(n) + \|b\|_\infty + \sum_{r=1}^K \|A_r^{(n)} - A_r\|_\infty \|X_{n-1}^{(r)}\|_\infty + \left\| \sum_{r=1}^K A_r X_{n-1}^{(r)} \right\|_\infty \\ &\leq C_X + C_b + \|g\|_\infty + (1 + R_\Sigma(n))\|X_{n-1}^{(r)}\| \\ &\leq \Big(R_X(n) + R_b(n) + \|g\|_\infty\Big) \sum_{j=1}^n \left(\prod_{i=j}^{n-1} (1 + R_\Sigma(i+1)) \right). \end{aligned}$$

With $R_\Sigma(n) \leq C_\Sigma / n$, we obtain

$$\prod_{i=j}^{n-1} (1 + R_\Sigma(i+1)) \leq \frac{(n + C_\Sigma)^{\lceil C_\Sigma \rceil}}{(j+1)^{\lceil C_\Sigma \rceil}}.$$

Thus,

$$\|X_n\|_\infty \leq \Big(C_X + C_b + \|g\|_\infty\Big)(n + C_\Sigma)^{\lceil C_\Sigma \rceil} \sum_{j=1}^n (j+1)^{-\lceil C_\Sigma \rceil},$$

which leads to the assertion. ■

Lemma 4.2 *We have*

$$\forall\, 0 < \bar{\xi} < 1, \forall\, n \geq 1 : \sum_{i=0}^{n-1} \frac{\bar{\xi}^i}{n-i} \leq \frac{1}{(1-\bar{\xi})^2} \frac{1}{n} \tag{38}$$

$$\forall\, 0 < \bar{\xi} < 1, \forall\, n \geq 1 : \bar{\xi}^n \leq \frac{1}{e \ln(1/\bar{\xi})} \frac{1}{n}. \tag{39}$$

Proof: For (38) note that $1/(n-i) \leq (i+1)/n$ for all $n \geq 1$ and $0 \leq i \leq n-1$. This implies

$$\sum_{i=0}^{n-1} \frac{\bar{\xi}^i}{n-i} \leq \frac{1}{n} \sum_{i=0}^{n-1} (i+1)\bar{\xi}^i \leq \frac{1}{n} \sum_{i=0}^{\infty} (i+1)\bar{\xi}^i \leq \frac{1}{(1-\bar{\xi})^2} \frac{1}{n}.$$

For (39) note that the function $x \mapsto x\bar{\xi}^x$, $x \geq 0$ has its maximum at $x = 1/\ln(1/\bar{\xi})$ which implies the assertion. ■

Lemma 4.3 *Let (X_n) be given by (25),(26) with $A_r^{(n)}, b^{(n)}, \langle \widetilde{X}_n \rangle$ satisfying (27)-(33) with R_Δ, R_Σ, $R_\Sigma^{(2)}$, R_b, R_X satisfying (37). Then, for all $n \geq 3$*

$$\ell_2(X_n, X) \leq C \frac{\ln(n)}{n},$$

where C is given by

$$C := \frac{\|g\|_\infty^2/(1-\xi) + \|X_\ell\|_\infty}{e \ln(1/\bar{\xi})\bar{\xi}^\ell} + \frac{\widetilde{C}}{(1-\bar{\xi})^2},$$

with $\widetilde{C}, \ell$ defined in (40), (41), $\bar{\xi} := (1+\xi)/2$ and $\|X_\ell\|_\infty$ estimated in Lemma 4.1.

Proof: With (34) and (37) it is

$$\begin{aligned} M(n) &\leq \sum_{j=1}^{n} \left(\prod_{i=j}^{n-1} \frac{i}{i+1} \right) (R_b(j) + R_X(j)) \\ &= \frac{1}{n} \sum_{j=1}^{n} j(R_b(j) + R_X(j)) \\ &\leq (C_b + C_X) \frac{H_n}{n} \\ &\leq (C_b + C_X) \frac{1+\ln(n)}{n}, \quad n \geq 1. \end{aligned}$$

Defining

$$\begin{aligned} \widetilde{C} := C_X + \max\Big\{ & 2C_A C_\Sigma^{(2)} \|X\|_2/\xi, \\ & (C_b^2 + C_\Sigma^{(2)} \|X\|_2^2 + 2C_A C_b (C_b + C_X) + C_A'(C_b + C_X)^2)^{1/2} \Big\}, \end{aligned} \quad (40)$$

we have

$$R_X(n-i) + d_{n-i} \leq \widetilde{C} \frac{1 \vee \ln(n-i)}{n-i} \leq \widetilde{C} \frac{1 \vee \ln(n)}{n-i}.$$

Set

$$\ell := \left\lceil \frac{2C_\xi}{\xi \wedge (1-\xi)} \right\rceil, \quad (41)$$

so that $\xi(n) \in [\xi/2, (1+\xi)/2]$ for $n \geq \ell$, and we obtain with (36) and Lemma 4.2,

$$\begin{aligned}
\ell_2(X_n, X) &\leq \bar{\xi}^n \bar{\xi}^{-\ell}(\|X\|_2 + \|X_\ell\|_2) + \widetilde{C} \sum_{i=0}^{n-1} \bar{\xi}^i \frac{1 \vee \ln(n)}{n-i} \\
&\leq \frac{1}{e \ln(1/\bar{\xi}) \bar{\xi}^\ell} \left(\frac{\|g\|_\infty^2}{1-\xi} + \|X_\ell\|_\infty \right) \frac{1}{n} + \frac{\widetilde{C}}{(1-\bar{\xi}^2)} \frac{1 \vee \ln(n)}{n},
\end{aligned}$$

which implies the assertion ■

In the following transposition of the ℓ_2 rate of convergence for (X_n) into a rate in the Kolmogorov metric we use an estimate of Lemma 5.1 in Fill and Janson [13]. The Kolmogorov metric is denoted by

$$\varrho(\lambda, \nu) := \sup_{x \in \mathbb{R}} |F_\lambda(x) - F_\nu(x)|,$$

where F_λ, F_ν denote the distribution functions of $\lambda, \nu \in \mathcal{M}$.

Lemma 4.4 *Let (X_n) and C be as in Lemma 4.3. Then, for all $n \geq 3$:*

$$\varrho(X_n, X) \leq 2(C\|w'\|_\infty)^{2/3} \left(\frac{\ln(n)}{n} \right)^{2/3}.$$

Proof: For the transposition of the ℓ_2 rate in Lemma 4.3 into a rate in the Kolmogorov metric we note that the bounded derivative of the density f implies that the modulus of continuity Δ_X of X is estimated by $\Delta_X(t) \leq \|w'\|_\infty t$ for all $t > 0$. Using the inequality

$$\varrho(X_n, X) \leq \ell_2^2(X_n, X) t^{-2} + \Delta_X(t),$$

valid for all $t > 0$ this implies

$$\varrho(X_n, X) \leq \frac{C^2 \ln^2(n)}{n^2} \frac{1}{t^2} + \|w'\|_\infty t,$$

for all $t > 0$. We choose

$$t = t_n = \left(\frac{C^2 \ln^2(n)}{\|w'\|_\infty n^2} \right)^{1/3}$$

which leads to the bound stated. ■

An approximation of w can now be constructed as in Theorem 6.1 in Fill and Janson [13]. In the proof we use a Taylor expansion of second order which improves the rate of convergence compared with the first order expansion used by Fill and Janson.

Theorem 4.5 *Let (X_n) and C be given as in Lemma 4.3 and denote by F_n the distribution functions of X_n. Define*

$$w_n(x) := \frac{F_n(x + \delta_n/2) - F_n(a - \delta_n/2)}{\delta_n}, \quad x \in \mathbb{R} \tag{42}$$

with

$$\delta_n = L\left(\frac{\ln(n)}{n}\right)^{2/9}, \tag{43}$$

with an $L > 0$*. Then*

$$\sup_{x\in\mathbb{R}} |w_n(x) - w(x)| \leq r_n := \left(\frac{4}{L}\Big(C\|w'\|_\infty\Big)^{2/3} + \frac{L^2\|w''\|_\infty}{24}\right)\left(\frac{\ln(n)}{n}\right)^{4/9}.$$

Proof: Let F denote the distribution function of X. By Taylor expansion we have $w(y) = w(x) + w'(x)(y-x) + (w''(\vartheta)/2)(y-x)^2$ with ϑ between x and y. This yields

$$\begin{aligned}
&|F(x+\delta_n/2) - F(x-\delta_n/2) - \delta w(x)| \\
\leq\ & \left|\int_{x-\delta_n/2}^{x+\delta_n/2} w(x) + w'(x)(y-x) + (w''(\vartheta)/2)(y-x)^2 - w(x)dy\right| \\
\leq\ & \frac{\|w''\|_\infty}{2}\int_{-\delta_n/2}^{\delta_n/2} y^2 dy \\
\leq\ & \frac{1}{24}\|w''\|_\infty \delta_n^3.
\end{aligned}$$

Thus with δ_n and w_n as given in the Lemma we obtain

$$\begin{aligned}
&\sup_{x\in\mathbb{R}} |w_n(x) - w(x)| \\
\leq\ & \sup_{x\in\mathbb{R}}\left\{\left|\frac{F_n(x+\delta_n/2) - F_n(x-\delta_n/2)}{\delta_n} - \frac{F(x+\delta_n/2) - F(x-\delta_n/2)}{\delta_n}\right|\right. \\
&\qquad \left. + \left|\frac{F(x+\delta_n/2) - F(x-\delta_n/2)}{\delta_n} - w(x)\right|\right\} \\
\leq\ & \frac{2}{\delta_n}\varrho(X_n, X) + \|w''\|_\infty \frac{\delta_n^2}{24} \\
\leq\ & \frac{2}{L}2\Big(C\|w'\|_\infty\Big)^{2/3}\left(\frac{\ln(n)}{n}\right)^{2/3} + \frac{\|w''\|_\infty}{24}L^2\left(\frac{\ln(n)}{n}\right)^{4/9} \\
\leq\ & \left(\frac{4}{L}\Big(C\|w'\|_\infty\Big)^{4/9} + \frac{L^2\|w''\|_\infty}{24}\right)\left(\frac{\ln(n)}{n}\right)^{4/9},
\end{aligned}$$

where we used Lemma 4.4. ■

Estimates for $\|w'\|_\infty, \|w''\|_\infty$ are given in Lemma 3.3.

4.2 Examples

For the examples of the sections 2.2.1-2.2.3 we define appropriate discretizations and show (37). The algorithmic computation of the distributions of the discretizations is done in the next section.

To define the discretized versions of $A_r = f_r(U)$ and $b = g(f(U))$ we denote, for $U = (U_1, \dots, U_d)$,

$$[U]_n := \left(\frac{\lfloor nU_1 \rfloor}{n}, \dots, \frac{\lfloor nU_d \rfloor}{n} \right) \tag{44}$$

and define for $r = 1, \dots K-1$

$$A_r^{(n)} := f_r([U]_n), \quad A_K^{(n)} := 1 - \frac{1}{n} - \sum_{r=1}^{K-1} f_r([U]_n). \tag{45}$$

For the discretization of b define first $\tilde{g}$ as g in (13) with the logarithm ln there replaced by the function $\breve{\ln}(x) := \ln(x)$ for $x \in (0,1)$ and $\breve{\ln}(x) := 0$ otherwise. Then it is $\tilde{g} = g$ on S_{K-1}. We define then

$$b^{(n)} := \tilde{g}(f([U]_s)) \tag{46}$$

with $s = s(n) := n^2 \lceil \ln(n) \rceil$ and the convention $\lceil \ln(n) \rceil := 1$ for $n = 1$. Furthermore we define

$$\langle \widetilde{X}_n \rangle := \frac{\lfloor n^2 \widetilde{X}_n \rfloor}{n^2}. \tag{47}$$

These choices can be used uniformly for all examples of the sections 2.2.1-2.2.3. For the verification of (37) we use a technical Lemma which allows us to treat the fact that $x \mapsto x\ln(x)$ has infinite derivative at $x = 0^+$.

Lemma 4.6 *With $\psi(x) := x\breve{\ln}(x)$ for $x \in \mathbb{R}$, we have*

$$|\psi(x) - \psi(y)| \le |x-y| \left(1 \vee \ln\left(\frac{1}{|x-y|} \right) \right), \quad x, y \in \mathbb{R}.$$

In particular, if $|x-y| \le c/n$ with $n, c \ge 1$ then

$$|\psi(x) - \psi(y)| \le c\, \frac{1 \vee \ln(n)}{n}.$$

Proof: For the first assertion distinguish the cases $|x-y| < 1/e$ and $\ge 1/e$. The second one follows directly from the first one. ■

4.2.1 m-ary search trees

Note that the discretization of U into $[U]_n$ preserves the ranks of the components so that with the f_r given in (14) we obtain

$$|f_r([U]_n) - f_r(U)| \le \frac{1}{n}, \quad r = 1, \dots, m.$$

Thus we may choose $C_\Sigma := C_\Sigma^{(2)} := m$. By Lemma 4.6 it is for all $r = 1, \dots m$,

$$|A_r^{(s)} \breve{\ln}(A_r^{(s)}) - A_r \ln(A_r)| \le \frac{1 \vee \ln(s)}{s} \le \frac{1 \vee (2\ln(n) + \ln(\ln(n)))}{n^2 \lceil \ln(n) \rceil} \le \frac{3}{n^2}, \tag{48}$$

thus we may choose $C_b := 3m$. Furthermore we choose $C_X := 1$ and for C_ξ note

$$\xi^2 = m\,\mathbb{E}\,U_{(1)}^2 = m\int_0^1 x^2(m-1)(1-x)^{m-2}\,dx = \frac{2}{m+1}. \tag{49}$$

Then,

$$|\xi(n)-\xi| = \frac{|\xi^2(n)-\xi^2|}{|\xi(n)-\xi|} \leq \frac{1}{\xi}\frac{m+2}{n} = C_\xi\frac{1}{n},$$

where $C_\xi := (m+2)\sqrt{(m+1)/2}$. Finally we can simply set $C_A := m$ and $C_A' := m(m-1)$ and (37) is satisfied.

4.2.2 Median of $(2\mathbf{k}+1)$ search tree

As for m-ary search trees the discretization preserves the ranks of the components and therefore also the median. This implies

$$|f_r([U]_n) - f_r(U)| \leq \frac{1}{n}, \quad r = 1,2.$$

We can choose $C_\Sigma := 2$, $\quad C_\Sigma^{(2)} := 2$ and by Lemma 4.6 similarly to (48) we obtain the choice $C_b := 6$. Furthermore $C_X := 1$ and for C_ξ note

$$\xi^2 = 2\,\mathbb{E}\,\mathrm{med}^2(U) = 2\int_0^1 x^2\frac{x^k(1-x)^k}{B(k+1,k+1)}\,dx = \frac{k+2}{2k+3}. \tag{50}$$

This yields $|\xi(n)-\xi| \leq 4/(\xi n) = C_\xi/n$ with $C_\xi := \sqrt{8(2k+3)/(k+2)}$. Finally $C_A := C_A' := 2$ completes the choices.

4.2.3 Quadtree

For quadtrees note thtn by induction we have for all $a_1,\dots,a_n,b_1,\dots,b_n \in [-1,1]$:

$$\left|\prod_{i=1}^K a_i - \prod_{i=1}^K b_i\right| \leq \sum_{i=1}^K |a_i - b_i|,$$

thus

$$|f_r([U]_n) - f_r(U)| \leq \frac{d}{n}, \quad r = 1,\dots,2^d,$$

the case $r = 2^d$ being also trivial. This yields

$$C_\Sigma = d2^d, \quad C_\Sigma^{(2)} := d^2 2^d,$$

and by Lemma 4.6 $C_b := 3d2^d$. Furthermore $C_X = 1$ and for C_ξ note

$$\xi^2 = 2^d\,\mathbb{E}\prod_{i=1}^d U_i^2 = \left(\frac{2}{3}\right)^d, \tag{51}$$

so that

$$|\xi(n) - \xi| \leq \frac{2d + d^2 2^d}{\xi n} = \frac{C_\xi}{n}$$

with $C_\xi := (2d + d^2 2^d)(3/2)^{d/2}$. Finally $C_A := 2^d, C'_A := 2^d(2^d - 1)$ completes the choices.

5 The rejection algorithm

The dominant curve and the associate sample needed for the rejection method were derived in section 3.2. It remains the problem of approximating the density w in order to decide the outcome of a rejection test.

Let K_2, K_4 be upper bounds for $\|w'\|_\infty, \|w''\|_\infty$, e.g., the choices given in Lemma 3.3. Then r_n given in Theorem 4.5 is estimated with the choice

$$L := \frac{96^{1/3}(CK_2)^{2/9}}{K_4^{1/3}} \tag{52}$$

by

$$r_n \leq R_n := \left(\frac{16}{3}\right)^{1/3} \left(CK_2\right)^{4/9} K_4^{1/3} \left(\frac{\ln(n)}{n}\right)^{4/9}. \tag{53}$$

Thus

$$\sup_{x \in \mathbb{R}} |w_n(x) - w(x)| \leq R_n,$$

with w_n given in (42) and L in the definition of δ_n there given by (52).

5.1 Algorithmic approximation of the density

For the computation of the approximations w_n of w we keep and update arrays $\mathcal{A}_n$ defined by

$$\mathcal{A}_n[k] := \mathbb{P}\left(X_n = \frac{k}{n}\right), \quad k \in \mathbb{Z},$$

so that $\mathcal{A}_n[k] \neq 0$ at most for $-Q_n \leq k \leq Q_n$ with Q_n given in Lemma 4.1. According to the recursive definition of X_n in (25), (26) and the choice of discretizations in (44)-(47) and with the notation $f_r^{(n)} := f_r$ for $r = 1, \ldots, K-1$ and $f_K^{(n)} := 1 - 1/n - \sum_{r=1}^{K-1} f_r$ we define first $\mathcal{A}_0[0] := 1, \mathcal{A}_0[k] := 0$ for $k \neq 0$ (which we call `initialize` $\mathcal{A}_0$) and for the update we assume that $\mathcal{A}_{n-1}$ is already given and $\mathcal{A}_n[k] := 0$ is initialized for all $k \in \mathbb{Z}$. Then we obtain $\mathcal{A}_n$ algorithmically by the procedure

`for` $i_1, \ldots, i_d = 0$ `to` $n^2\lceil\ln(n)\rceil - 1$ `do`

```
for j_1 ..., j_K = -Q_{n-1} to Q_{n-1} do
```

$$u := \frac{1}{n}\left(\left\lfloor \frac{i_1}{n\lceil \ln(n)\rceil}\right\rfloor, \ldots, \left\lfloor \frac{i_d}{n\lceil \ln(n)\rceil}\right\rfloor\right)$$

$$v := \frac{1}{n^2\lceil \ln(n)\rceil}(i_1, \ldots, i_d)$$

$$k := \frac{1}{n^2}\left\lfloor n^2\left(\sum_{r=1}^{K} f_r^{(n)}(u)j_r + \tilde{g}(f^{(n)}(v))\right)\right\rfloor$$

$$\mathcal{A}_n[k] := \mathcal{A}_n[k] + \left(n^2\lceil \ln(n)\rceil\right)^d \prod_{r=1}^{K} \mathcal{A}_{n-1}[j_r]$$

```
  enddo
enddo
```

We call this procedure update$(\mathcal{A}_{n-1}, \mathcal{A}_n)$. Then with the array $\mathcal{A}_n$ the discrete approximation w_n of w as in Theorem 4.5 is obtained by

$$w_n(x) := \frac{1}{\delta_n} \sum_{n(x-\delta_n/2)<k\leq n(x+\delta_n/2)} \mathcal{A}_n[k]. \tag{54}$$

5.2 The algorithm

Therefore, analogously to the algorithm in Devroye, Fill, and Neininger [6] the rejection algorithm looks as follows with w_n as in (54), δ_n there as in (43) with L as in (52), and R_n as in (53):

```
repeat
  generate indep. U unif[0,1] and X as in (24)
  T ← Uq(X)
  initialize A_0
  n ← 0
  repeat
    n ← n + 1
    update(A_{n-1}, A_n)
    Y ← w_n(X)
  until n ≥ 3 and |T − Y| ≥ R_n
  Accept = [T ≤ Y − R_n]
until Accept
return X
```

The correctness of the algorithm follows from von Neumann's rejection method, see [4].

5.3 Complexity

It is well-known that the expected number of (outer) loops of a rejection algorithm is the L_1-norm of the dominating curve, thus in our case this is $\|q\|_1 = 4K_1^{1/2}(2K_2K_3)^{1/4}$.

For the inner loop there is no universally accepted complexity measure. We propose for this to estimate the number of steps to approximate the density w up to an accuracy of $O(1/n)$. In the case

$0 < C_\Sigma \leq 1$ the `update` $(\mathcal{A}_{j-1}, \mathcal{A}_j)$ costs $O((j^2 \ln(j))^d (j \ln(j))^K) = O(j^{2d+K} (\ln(j))^{d+K})$ time units thus the computation of the array $\mathcal{A}_m$ takes time

$$O\left(\sum_{j=1}^{m} j^{2d+K} (\ln(j))^{d+K}\right) = O\Big(m^{2d+K+1} (\ln(m))^{d+K}\Big).$$

Since using $\mathcal{A}_m$ we can, by Lemma 4.5, approximate w up to a precision of $O((\ln(m)/m)^{4/9})$ we set $m = n^{9/4} \ln(n)$. This substitution implies that an approximation of w of the order $O(1/n)$ costs time

$$O\left(n^{(9/4)(2d+K+1)} (\ln(n))^{3d+2K+1}\right).$$

An analogous calculation leads in the case $C_\Sigma > 1$ to an approximation of the order $O(1/n)$ at the cost of

$$O\left(n^{(9/4)(2d+K\lceil C_\Sigma \rceil+1)} (\ln(n))^{3d+K\lceil C_\Sigma \rceil+1}\right).$$

For the special case of the limit law of the number of key comparisons of the QUICKSORT algorithm applied to a set of randomly permuted items we have $C_\Sigma = 1, d = 1, K = 2$, which gives an approximation of w at the order $O(1/n)$ at the cost of $O(n^{11.25} (\ln(n))^8)$. This improves the algorithm of Devroye, Fill and Neininger [6], where the approximation of w of the order $O(1/n)$ was calculated at the cost of $O(n^{36})$. However, the expected time taken by the inner loop in our algorithm is infinite. We do not know if a finite expected time algorithm exists that is allowed to use only the basic algebraic operations such as addition, comparison and multiplication. A solid lower bound theory for simulation algorithms is still lacking.

Acknowledgment

We are grateful to Jim Fill for making the technical reports [12, 13] available to us. We also thank the referee, whose comments are much appreciated.

References

[1] Bickel, P. J. and P. A. Freedman (1981). Some asymptotic theory for the bootstrap. *Ann. Statist. 9*, 1196–1217.

[2] Bougerol, P. and N. Picard (1992). Strict stationarity of generalized autoregressive processes. *Ann. Probab. 20*, 1714–1798.

[3] Chen, R., R. Goodman, and A. Zame (1984). Limiting distributions of two random sequences. *J. Multivariate Anal. 14*, 221–230.

[4] Devroye, L. (1986). *Nonuniform Random Variate Generation*. Springer-Verlag, New York.

[5] Devroye, L. (2001). Simulating perpetuities. *Methodol. Comput. Appl. Probab. 3*, 97–115.

[6] Devroye, L., J. A. Fill, and R. Neininger (2000). Perfect simulation from the Quicksort limit distribution. *Electron. Comm. Probab. 5*, 95–99 (electronic).

[7] de Bruijn, N. G. (1951). The asymptotic behaviour of a function occurring in the theory of primes. *J. Indian Math. Soc. (N.S.) 15*, 25–32.

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

[9] Embrechts, P. and C. M. Goldie, (1994). Perpetuities and random equations. *Asymptotic statistics (Prague, 1993)*, 75–86.

[10] Embrechts, P., C. Klüppelberg and T. Mikosch, (1997). *Modelling extremal events*, Springer, Berlin.

[11] Fill, J. A. and S. Janson (2000). Smoothness and decay properties of the limiting Quicksort density function. *Mathematics and Computer Science (Versailles, 2000)*, 53–64. Trends Math., Birkhäuser, Basel.

[12] Fill, J. A. and S. Janson (2001). Approximating the Limiting Quicksort Distribution. *Random Structures Algorithms 19*, 376–406.

[13] Fill, J. A. and S. Janson (2001). Quicksort Asymptotics. Technical Report #597, Department of Mathematical Sciences, The Johns Hopkins University. Available from `http://www.mts.jhu.edu/~fill/papers/quick_asy.ps`.

[14] Flajolet, P., G. Labelle, L. Laforest, and B. Salvy (1995). Hypergeometrics and the cost structure of quadtrees. *Random Structures Algorithms 10*, 117-144.

[15] Goldie, C. M. and R. Grübel (1996). Perpetuities with thin tails. *Adv. in Appl. Probab. 28*, 463–480.

[16] Goldie, C. M. and R. A. Maller (2000). Stability of perpetuities. *Ann. Probab. 28*, 1195–1218.

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

[18] Hoare, C. A. R. (1961). Algorithm 64, Quicksort. *Comm. ACM 4*, 321-322.

[19] Hennequin, P. (1989) Combinatorial analysis of quicksort algorithm. *RAIRO Informatique théorique et Applications 23*, 317–333.

[20] Hennequin, P. (1991) *Analyse en moyenne d'algorithme, tri rapide et arbres de recherche*, Ph.D. Thesis, Ecole Polytechnique.

[21] Hoare, C. A. R. (1962). Quicksort. *Comput. J.*, 10–15.

[22] Hwang, H.-K. and R. Neininger (2001). Phase change of limit laws in the quicksort recurrence under varying toll functions. Preprint available from `http://neyman.mathematik.uni-freiburg.de/homepages/neininger/qfinal.ps`

[23] Hwang, H.-K. and T.-H. Tsai (2000). Quickselect and Dickman function. *Combin. Probab. Comput.*, to appear.

[24] Kesten, H. (1973). Random difference equations and renewal theory for products of random matrices. *Acta Math. 131*, 207–248.

[25] Knuth, D. E. (1997). *The Art of Computer Programming. Vol. 3: Sorting and Searching.* Addison-Wesley.

[26] Mahmoud, H. M. (1991). Limiting distributions for path lengths in recursive trees. *Probab. Engrg. Inform. Sci. 5*, 53-59.

[27] Mahmoud, H. M. (1992). *Evolution of Random Search Trees.* John Wiley & Sons, New York, 1992.

[28] Mahmoud, H. M., R. Modarres, and R. T. Smythe (1995). Analysis of quickselect: An algorithm for order statistics. *RAIRO Informatique Théorique et Applications* **29**, 255-276.

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

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

[31] Neininger, R. and L. Rüschendorf (2001). A general contraction theorem and asymptotic normality in combinatorial structures. Preprint available from `http://neyman.mathematik.uni-freiburg.de/homepages/neininger/asynorm.ps`

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

[33] Régnier, M. (1989). A limiting distribution for quicksort. *RAIRO Informatique Théorique et Application 23*, 335–343.

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

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

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

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

[38] Sedgewick, R. (1977). The analysis of Quicksort programs. *Acta Informat. 7*, 327–355.

[39] Sedgewick, R. *Quicksort.* Gurland Publishing, NY, 1980.

[40] Sedgewick, R. and P. Flajolet (1996). *An introduction to the analysis of algorithms.*. Addison-Wesley, Amsterdam.

[41] Takács, L. (1955). On stochastic processes connected with certain physical recording apparatuses. *Acta Math. Acad. Sci. Hungar. 6*, 363–380.

[42] Tenenbaum, G. (1995). *Introduction to Analytic and Probabilistic Number Theory*. Translated from the second French edition (1995) by C. B. Thomas. Cambridge University Press, Cambridge.

[43] Vervaat, W. (1979). On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Adv. in Appl. Probab. 11,* 750–783.