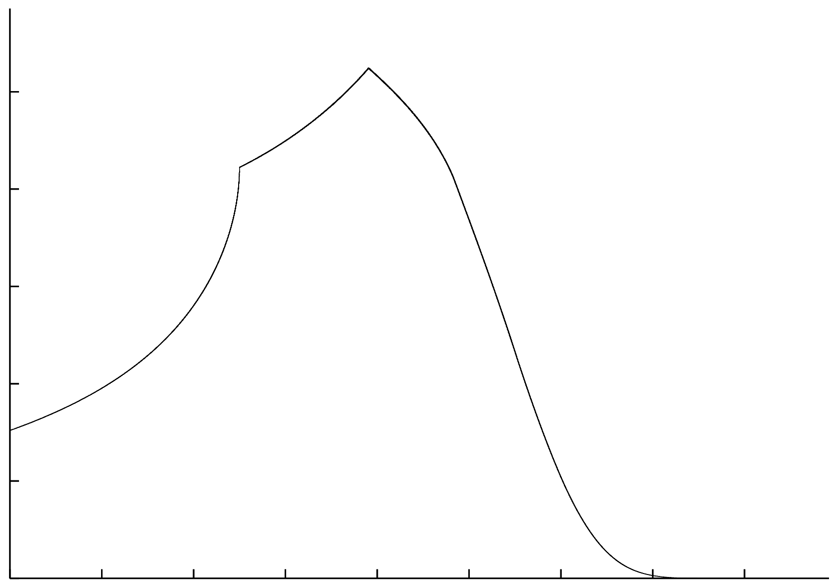
Johann Wolfgang Goethe-Universität Frankfurt am Main
Fachbereich Mathematik und Informatik
Institut für Stochastik und Mathematische Informatik
Prof. Dr. R. Neininger

Diplomarbeit

Approximating Perpetuities



Margarete Knape

Dezember 2006

Danksagung

An dieser Stelle möchte ich mich bei all jenen bedanken, die mir diese Diplomarbeit ermöglicht und zu ihrer Erstellung beigetragen haben.

Herrn Prof. Dr. R. Neininger danke ich für das interessante Thema und die hervorragende Betreuung. Sein Interesse am Fortgang der Arbeit, seine konstruktiven Anregungen und nicht zuletzt seine Zuversicht haben wesentlich zum Gelingen der Arbeit beigetragen.

Bei Herrn Dr. Ferebee möchte ich mich für sein Interesse bedanken und für seine Anregungen, die zum Fortschritt der Arbeit beigetragen haben. Auch Dr. Gaby Schneider, Tämur Ali Khan, Rupert Hartung und Martin Hutzenthaler danke ich für die moralische Unterstützung sowie für wertvolle Tipps.

Nicht zuletzt gilt mein besonderer Dank meinen Eltern, die mir das Studium ermöglicht haben, und Hartmut und Roswitha Knape und meiner Familie, die mich stets unterstützt und gerade in der letzten Zeit viel Rücksicht genommen haben.

Zusammenfassung

Als Perpetuity wird vor allem in der Versicherungs- und Finanzmathematik eine Zufallsvariable X auf $\mathbb{R}$ bezeichnet, deren Verteilung implizit durch eine stochastische Fixpunktgleichung der Form $X \stackrel{d}{=} AX + b$ charakterisiert ist. Dabei ist (A, b) ein Vektor von Zufallsvariablen, der unabhängig von X ist. Abhängigkeiten zwischen A und b sind jedoch erlaubt.

Bedingungen für die Existenz und Eindeutigkeit von Lösungen solcher Fixpunktgleichungen sind bereits seit längerem bekannt. Für eine große Klasse dieser Perpetuities existieren Tail-Abschätzungen. Ziel dieser Arbeit ist es, den zentralen Bereich solcher Verteilungen zu untersuchen. Dazu wird ein Algorithmus für die Approximation der Verteilungsfunktionen und gegebenenfalls der Dichten von einer möglichst großen Klasse solcher Perpetuities entwickelt. Um für diese Approximationen explizite Fehlerschranken anzugeben, muss der Stetigkeitsmodul der approximierten Funktion abgeschätzt werden. Für eine spezielle Klasse von Fixpunktgleichungen werden universelle Abschätzungen angegeben, im Allgemeinen muss eine solche Abschätzung jedoch für den Einzelfall hergeleitet werden. Dies wird exemplarisch an einem Beispiel aus der probabilistischen Analyse von Algorithmen durchgeführt, für das auch der Algorithmus implementiert und eine Tafel der Verteilungsfunktion generiert wird.

Um die Qualität der erhaltenen Fehlerschranken und die praktische Verwendbarkeit des Algorithmus zu beurteilen, werden abschließend einige Beispiele untersucht, in denen die Dichten oder zumindest gewisse Eigenschaften bereits bekannt sind. Hierbei zeigt sich, dass die theoretischen Fehlerschranken stets deutlich unterschritten werden und die Approximation in praktikabler Laufzeit bereits sehr gut Ergebnisse liefert.

Der verwendete Algorithmus beruht auf einem bekannten Verfahren, das jedoch für eine andere Klasse von Fixpunktgleichungen entwickelt wurde. Bei der Anpassung an den hier betrachteten Fall konnte eine wesentliche Verbesserung erreicht werden, die sich auch auf den ursprünglichen Algorithmus übertragen lässt.

Contents

1 Introduction **2**

2 Approximation **5**

2.1 Vervaat: conditions for existence and uniqueness of solutions 5

2.2 Rates of convergence 7

2.2.1 Minimal L_p metric 8

2.2.2 Kolmogorov metric 10

2.2.3 Approximation of the density 11

2.3 A class of perpetuities 13

2.3.1 A bound for the density 13

2.3.2 The modulus of continuity 14

2.4 Implementation 16

2.4.1 Algorithm 16

2.4.2 Complexity 17

3 Example: key exchanges in Quickselect **18**

3.1 Basic properties 19

3.2 An integral equation for the density 20

3.3 A bound for the density 22

3.4 The modulus of continuity 25

3.5 Implementation and explicit error bounds 28

4 Further examples **30**

4.1 Interval Splitting 30

4.1.1 $q = 1$ 31

4.1.2 $q = 1/2$ 34

4.1.3 $q = 0$ 37

4.2 Dickman distribution 38

Appendix A C++ Code **41**

Appendix B Table for distribution function of key exchanges **43**

References **44**

1 Introduction

In probability theory, a perpetuity denotes a random variable X in $\mathbb{R}$ that satisfies the stochastic fixed-point equation

$$X \stackrel{d}{=} AX + b, \tag{1}$$

where (A, b) is a vector of random variables which is independent of X, whereas dependence between A and b is allowed. The symbol $\stackrel{d}{=}$ denotes that left and right hand side in (1) are identically distributed. The notion of perpetuity extends directly to the multivariate case, where X and b are random vectors in $\mathbb{R}^d$ and A is a random $d \times d$ matrix. In th is work, we will however restrict ourselves to the univariate case.

Perpetuities arise in various different contexts:

- In discrete mathematics, perpetuities arise as the limit distributions of certain count statistics of decomposable combinatorial structures such as random permutations or random integers. In these areas, perpetuities (in particular the Dickman distribution) often arise via relationships to the GEM and Poisson-Dirichlet distributions. The Dickman distribution is a prototypical perpetuity, obtained from (1) by setting $A = b = U$ with U being uniformly distributed on the unit interval $[0, 1]$; see Arratia, Barbour, and Tavaré (2003) for perpetuities, GEM and Poisson-Dirichlet distribution in the context of combinatorial structures; see Donnelly and Grimmett (1993) for occurrences in probabilistic number theory.

- In the probabilistic analysis of algorithms, perpetuities come up as limit distributions of certain cost measures of recursive algorithms such as the selection algorithm Quickselect, see e.g. Hwang and Tsai (2002) or Mahmoud, Modarres, and Smythe (1995).

- In insurance and financial mathematics, a perpetuity represents the value of a commitment to make regular payments, where b represents the payment and A a discount factor both being subject to random fluctuation; see, e.g. Goldie and Maller (2000) or Embrechts, Klüppelberg, and Mikosch (1997, Section 8.4).

- Further, less systematic occurrences in connection with interval splitting procedures are discussed in Section 4.1.

As perpetuities are given implicitly by their fixed point characterization (1), properties of their distributions are not directly amenable. However, various questions about perpetuities have already been settled. Necessary and sufficient conditions on (A,b) for the fixed-point equation (1) to uniquely determine a probability distribution are discussed in Vervaat (1979) and Goldie and Maller (2000). Vervaat's argument can be found in Section 2.1. The tail behavior of perpetuities has been studied for certain cases in Goldie and Grübel (1996).

In the present work, we are interested in the central region of the distributions. The aim is to approximate perpetuities, in particular their distribution functions and their Lebesgue densities (if they exist).

To find a solution of (1), one approach is to define a mapping T on the space $\mathcal{M}$ of probability distributions, by

$$T : \mathcal{M} \to \mathcal{M}, \ \mu \mapsto \mathcal{L}(AY + b),$$

where Y is independent of (A,b), and $\mathcal{L}(Y) = \mu$. Then, $\mathcal{L}(X)$ is a fixed-point of T if and only if X satisfies (1). To approximate $\mathcal{L}(X)$, we iterate T, starting with some distribution μ_0. In Section 2.1, we discuss sufficient conditions on A and b for the convergence of this approximation and the uniqueness of the fixed-point and these are the cases considered subsequently. However, it is generally not possible to algorithmically compute the iterations of T exactly, when at least one of the occurring distributions is continuous. We will therefore follow the approach in Devroye and Neininger (2002) and use discrete approximations $(A^{(n)}, b^{(n)})$ of (A,b), which become more accurate for increasing n, to approximate T by a mapping $\widetilde{T}^{(n)}$, defined by

$$\widetilde{T}^{(n)} : \mathcal{M} \to \mathcal{M}, \ \mu \mapsto \mathcal{L}\big(A^{(n)}Y + b^{(n)}\big),$$

where again Y is independent of $(A^{(n)}, b^{(n)})$ and $\mathcal{L}(Y) = \mu$.

Although this approach can be translated into an algorithm, the running time of such an algorithm, starting with a simple distribution, e.g. the Dirac measure in $\mathbb{E}[X]$, is typically exponential. To allow for an efficient computation of the approximation, we introduce a further discretisation step $\langle\cdot\rangle_n$, explained in detail in Section 2.4 and define

$$T^{(n)} : \mathcal{M} \to \mathcal{M}, \ \mu \mapsto \mathcal{L}\big(\big\langle A^{(n)}Y + b^{(n)}\big\rangle_n\big),$$

where Y is independent of $(A^{(n)}, b^{(n)})$ and $\mathcal{L}(Y) = \mu$. In Section 2.2, we give conditions for $T^{(n)} \circ T^{(n-1)} \circ \cdots \circ T^{(1)}(\mu_0)$ to converge to the solution of (1). To this aim, we derive a rate of convergence in the minimal L_p metric ℓ_p, defined on the space $\mathcal{M}_p$ of probability measures on $\mathbb{R}$ with finite p-th moment by

$$\ell_p(\nu,\mu) := \inf\left\{\big\|V - W\big\|_p : \ \mathcal{L}(V) = \nu, \mathcal{L}(W) = \mu\right\}, \qquad \text{for } \nu,\mu \in \mathcal{M}_p \tag{2}$$

in Section 2.2.1. To get an explicit error bound for the distribution function, we then convert this in Section 2.2.2 into a rate of convergence in the Kolmogorov metric ϱ, defined by

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

where F_ν, F_μ denote the distribution functions of $\nu, \mu \in \mathcal{M}$. This implies explicit rates of convergence for distribution function and density, depending on the corresponding moduli of continuity of the fixed-point.

Such explicit rates for an approximation of the density can be used for perfect simulation from the distribution of the fixed-point using von Neumann's rejection method, see Devroye (2001), where the densities of certain perpetuities are approximated using a different approach based on characteristic functions and restricted to infinitely divisible distributions.

For the moduli of continuity needed, we find global bounds for perpetuities with $b \equiv 1$ in Section 2.3. For cases with random b, these moduli of continuity have to be derived individually. One example, connected to the selection algorithm Quickselect, is worked out in detail in Section 3, which is a main part of this work.

An implementation of an approximation of the form $T^{(n)} \circ \cdots \circ T^{(1)}(\mu_0)$ can be found in Section 2.4, where we also analyze its complexity. As a measure of the complexity of the approximation, we use the number of steps needed to obtain an accuracy of order $O(1/n)$. Although we generally follow the approach in Devroye and Neininger (2002), we can improve the complexity significantly by using different discretisations. For the approximation of the distribution function to an accuracy of $O(1/n)$ in a typical case, we obtain a complexity of $O(n^{1+\varepsilon})$ for any $\varepsilon > 0$. In comparison, the algorithm described in Devroye and Neininger (2002), which originally was designed for a different class of fixed-point equations, would lead to a complexity of $O(n^{4+\varepsilon})$, if applied to our cases. For the approximation of the density to an accuracy of order $O(1/n)$, we obtain a complexity of $O(n^{2+\varepsilon})$ compared to $O(n^{8+\varepsilon})$ for the algorithm in Devroye and Neininger (2002).

In Section 4, we apply the algorithm to some exemplary fixed-point equations, for which the solutions are more or less explicitly known. This enables us to compare the theoretical results to the actual error and to get an idea of the accuracy that can be attained with feasible running times.

2 Approximation

2.1 Vervaat: conditions for existence and uniqueness of solutions

Following Vervaat (1979), we find sufficient conditions on A and b for the existence and uniqueness of solutions of the fixed-point equation (1). A complete characterization of the existence of solutions of (1) can be found in Goldie and Maller (2000).

Theorem 2.1 (Vervaat 1979)**.** *Let A, b be real-valued random variables satisfying*

$$-\infty \leq \mathbb{E}\big[\log|A|\,\big] < 0 \qquad \textit{and} \qquad \mathbb{E}\big[\log^+|b|\big] < \infty,$$

where $\log$ *denotes the natural logarithm and* $\log^+ x := 0 \vee \log x$ *for* $x \in \mathbb{R}_+$*. Then, fixed-point equation* (1) *has a solution and this solution is unique in distribution.*

Proof. Let $(A_i, b_i)_{i\in\mathbb{N}}$ be a sequence of independent and identically distributed (iid) copies of (A, b), and $\mu := \mathbb{E}\big[\log|A|\big]$. Then, the strong law of large numbers implies

$$\frac{1}{n}\sum_{k=1}^{n} \log|A_k| \to \mu \quad \text{a.s.} \tag{3}$$

For any real-valued random variable X_0, we define the sequence $(X_n(X_0))_n$ by

$$X_n(X_0) := A_n \cdot X_{n-1}(X_0) + b_n \;\text{ for } n \geq 2, \qquad X_1(X_0) := A_1 X_0 + b_1, \tag{4}$$

so

$$X_n(X_0) = b_n + A_n b_{n-1} + A_n A_{n-1} b_{n-2} + \cdots + A_n \cdots A_2 b_1 + A_n \cdots A_1 X_0.$$

To show uniqueness of the limit, we compare with the sequence for a different starting point X_0', but the same sequence (A_i, b_i), and get

$$X_n(X_0) - X_n(X_0') = \left(\prod_{k=1}^{n} A_k\right)(X_0 - X_0').$$

Now (3) implies

$$\left|\prod_{k=1}^{n} A_k\right|^{1/n} \xrightarrow[n\to\infty]{} e^{\mu} \quad \text{a.s.}, \tag{5}$$

and using that μ is negative, we get

$$\prod_{k=1}^{n} A_k \xrightarrow[n\to\infty]{} 0 \quad \text{a.s.} \tag{6}$$

So if $\big(X_n(X_0)\big)_n$ converges in distribution for one X_0, it converges for every X_0, and this limit is unique in distribution. But if X is a solution of the fixed-point equation, we have $X_n(X) \stackrel{d}{=} X$ for all n, so this implies that the fixed-point of the equation is unique in distribution.

To show existence of the fixed-point, we define a new sequence of random variables $(Y_n)_{n\in\mathbb{N}}$ by

$$Y_0 := 0, \qquad Y_n := \sum_{k=1}^{n} A_1 \cdots A_{k-1} b_k \quad \text{for } n \geq 1$$

and observe that

$$\begin{aligned} X_n(X_0) &= b_n + A_n b_{n-1} + A_n A_{n-1} b_{n-2} + \cdots + A_n \cdots A_2 b_1 + A_n \cdots A_1 X_0 \\ &\stackrel{d}{=} b_1 + A_1 b_2 + A_1 A_2 b_3 + \cdots + A_1 \cdots A_{n-1} b_n + A_1 \cdots A_n X_0 \\ &= Y_n + \left(\prod_{k=1}^{n} A_k \right) X_0. \end{aligned}$$

So using (6) it is sufficient to show that the infinite series

$$\lim_{n\to\infty} Y_n = \sum_{k=1}^{\infty} A_1 \cdots A_{k-1} b_k \tag{7}$$

converges almost surely.

The b_k are iid, so $\mathbb{E}\left[\log^+ |b|\right] < \infty$ implies that for all $a > 1$,

$$\begin{aligned} \sum_{k=1}^{\infty} \mathbb{P}\left[|b_k| \geq a^k\right] &= \sum_{k=1}^{\infty} \mathbb{P}\left[\log|b_k| \geq k \log a\right] \\ &= \sum_{k=1}^{\infty} \mathbb{P}\left[\log^+|b| \geq k \log a\right] \\ &\leq \frac{1}{\log a} \mathbb{E}\left[\log^+|b|\right] \\ &< \infty. \end{aligned}$$

Using Borel-Cantelli for the events $A_k := \left\{|b_k| \geq a^k\right\}$, we get

$$\mathbb{P}\left[\left|b_n\right|^{1/n} \geq a \text{ infinitely often}\right] = 0$$

for all $a > 1$, so $\limsup \left|b_n\right|^{1/n} \leq 1$ almost surely. Combining this with (5) and using that μ is negative, we get

$$\limsup_{n\to\infty} \left|A_1 \cdots A_{n-1} b_n\right|^{1/n} \leq e^{\mu} < 1 \qquad \text{a.s.}$$

So by Cauchy's root criterion series (7) converges almost surely. □

Corollary 2.2. *If the conditions of Theorem 2.1 are satisfied, $X_n(X_0)$ converges in distribution for all real-valued random variables X_0.*

Remark 2.3. The conditions of Theorem 2.1 are satisfied, if

$$\|A\|_p < 1 \qquad \text{and} \qquad \|b\|_p < \infty \tag{8}$$

for any $p \geq 1$, as we can see using Jensen's inequality. Then the series given in (7) also converges in p-th mean and $\mathbb{E}[|X|^p]$ is finite. The moments $\mathbb{E}[X^j]$ for $j = 1, 2, \ldots, \lfloor p \rfloor$ are uniquely determined by

$$\mathbb{E}\left[X^j\right] = \sum_{k=0}^{j} \binom{j}{k} \mathbb{E}\left[A^k B^{j-k}\right] \mathbb{E}\left[X^k\right] \qquad \text{for } j = 1, 2, \ldots, \lfloor p \rfloor . \tag{9}$$

In the following, we will always work with fixed-point equations satisfying the conditions of this remark for some $p \in \mathbb{N}$.

2.2 Rates of convergence

To obtain an algorithmically computable approximation of the solution of the fixed-point equation (1), we use an approximation of the sequence defined in (4). Therefore we replace the iid. copies of (A, b) by a sequence of independent discrete approximations $(A^{(n)}, b^{(n)})$, converging to (A, b) in p-th mean for $n \to \infty$. To reduce the complexity, we introduce a further discretisation step $\langle \cdot \rangle_n$, that reduces the number of values attained by X_n. A concrete implementation of such discretisations can be found in the next section. Putting this together, we obtain

$$\widetilde{X}_n := A^{(n)} X_{n-1} + b^{(n)} \tag{10}$$

$$X_n := \langle \widetilde{X}_n \rangle_n . \tag{11}$$

We assume that the discretisations $A^{(n)}$, $b^{(n)}$ and $\langle \cdot \rangle_n$ satisfy

$$\left\| A^{(n)} - A \right\|_p \leq R_A(n) \tag{12}$$

$$\left\| b^{(n)} - b \right\|_p \leq R_b(n) \tag{13}$$

$$\left\| \langle \widetilde{X}_n \rangle_n - \widetilde{X}_n \right\|_p \leq R_X(n), \tag{14}$$

for some error functions R_A, R_b and R_X, which we will specify later.

Furthermore, we assume that there is some $\bar{\xi}_p < 1$, such that

$$\|A\|_p \leq \bar{\xi}_p \quad \text{and} \quad \left\| A^{(n)} \right\|_p \leq \bar{\xi}_p \quad \text{for all } n. \tag{15}$$

This is always possible for sufficiently large n, as we assume that $\|A\|_p < 1$, hence this is no real restriction on the choice of discretisations but can be reached by appropriately shifting the indices.

2.2.1 Minimal L_p metric

We now derive a rate of convergence for this discrete approximation in the ℓ_p metric. To find explicit estimates, we have to specify functions R_A, R_b, and R_X in (12)–(14), and we will carry this out for polynomial discretisation ($R(n) = O(n^{-r})$) as well as exponential discretisation ($R(n) = O(\gamma^{-n})$).

For simplicity, we use the shorthand notation $\ell_p(X, Y) := \ell_p(\mathcal{L}(X), \mathcal{L}(Y))$.

Lemma 2.4. *Let* (X_n) *be defined by* (10) *and* (11), $\bar{\xi}_p$ *as in* (15). *Then*

$$\ell_p(X_n, X) \leq \bar{\xi}_p^n \, \|X - \mathbb{E}X\|_p + \sum_{i=0}^{n-1} \bar{\xi}_p^i \, R(n-i), \tag{16}$$

where $R(n) := R_X(n) + R_A(n) \|X\|_p + R_b(n)$ *for the error functions in* (12)–(14).

Proof. We have

$$\begin{aligned} \ell_p(X_n, X) &\leq \ell_p(X_n, \widetilde{X}_n) + \ell_p(\widetilde{X}_n, X) \\ &\leq \left\| \langle \widetilde{X}_n \rangle_n - \widetilde{X}_n \right\|_p + \ell_p(\widetilde{X}_n, X). \end{aligned} \tag{17}$$

The first term is bounded by (14) and for the second term we get

$$\begin{aligned} \ell_p(\widetilde{X}_n, X) &\leq \left\| \widetilde{X}_n - X \right\|_p = \left\| A^{(n)} X_{n-1} + b^{(n)} - AX - b \right\|_p \\ &\leq \left\| A^{(n)} X_{n-1} - AX \right\|_p + \left\| b^{(n)} - b \right\|_p \\ &= \left\| A^{(n)}(X_{n-1} - X) - (A - A^{(n)})X \right\|_p + \left\| b^{(n)} - b \right\|_p \\ &\leq \left\| A^{(n)} \right\|_p \|X_{n-1} - X\|_p + \left\| A - A^{(n)} \right\|_p \|X\|_p + \left\| b^{(n)} - b \right\|_p, \end{aligned}$$

where in the last step we use that $A^{(n)}$ and $(X_{n-1} - X)$ as well as $(A - A^{(n)})$ and X are independent by assumption.

Now we use the important property of ℓ_p that the infimum in definition (2) is attained, see Bickel and Freedman (1981). We use a so called optimal coupling of X_{n-1} and X, for example by taking $U \sim \mathrm{unif}[0, 1]$ independent of $\left(A^{(n)}, b^{(n)}\right)$ and (A, b) and then setting $X_{n-1} := F_{X_{n-1}}^{-1}(U)$ and $X := F_X^{-1}(U)$, where F_Y^{-1} is the generalized inverse of the distribution function of Y, and get $\|X_{n-1} - X\|_p = \ell_p(X_{n-1}, X)$. Combining this with (17) and using the bounds given in (12)–(15), we obtain

$$\ell_p(X_n, X) \leq R_X(n) + \bar{\xi}_p \, \ell_p(X_{n-1}, X) + R_A(n) \|X\|_p + R_b(n),$$

and the claim then follows by induction, finally using $X_0 = \mathbb{E}[X]$. □

To make this estimates explicit we have to specify functions for $R_A(n)$, $R_b(n)$, and $R_X(n)$. We will do so in two different ways, one representing a polynomial discretisation of the corresponding random variables and one representing an exponential discretisation. Although we get better asymptotic results for the second case, we will see in the examples that when we are actually implementing the approximation on recent hardware, the first approach is superior.

Corollary 2.5. *Let* (X_n) *be defined by* (10) *and* (11), $\bar{\xi}_p$ *as in* (15), *and assume*

$$R_A(n) \le C_A \cdot \frac{1}{n^r}, \qquad R_b(n) \le C_b \cdot \frac{1}{n^r}, \qquad R_X(n) \le C_X \cdot \frac{1}{n^r},$$

for some $r \ge 1$. *Then, we get*

$$\ell_p(X_n, X) \le C_r \cdot \frac{1}{n^r},$$

where

$$C_r := \frac{r \cdot \|X - \mathbb{E}X\|_p}{e^r \cdot \log\big(1/\bar{\xi}_p\big)} + \frac{r!\,\Big(C_X + C_b + C_A \,\|X\|_p\Big)}{\big(1-\bar{\xi}_p\big)^{r+1}} \tag{18}$$

Proof. Using Lemma 2.4 we get

$$\ell_p(X_n, X) \le \bar{\xi}_p^n \,\|X - \mathbb{E}X\|_p + (C_X + C_A \|X\|_p + C_b) \sum_{i=0}^{n-1} \frac{\bar{\xi}_p^i}{(n-i)^r}. \tag{19}$$

To see that both summands are of order n^{-r}, we can extend the argumentation in Lemma 4 of Devroye and Neininger (2002) to get

$$\bar{\xi}_p^n \le \frac{r}{\mathrm{e}^r \log\big(1/\bar{\xi}_p\big)} \cdot \frac{1}{n^r} \quad \text{and} \tag{20}$$

$$\sum_{i=0}^{n-1} \frac{\bar{\xi}_p^i}{(n-i)^r} \le \frac{r!}{\big(1-\bar{\xi}_p\big)^{r+1}} \cdot \frac{1}{n^r} \qquad \text{for } 0 < \bar{\xi}_p < 1,\ n \ge 1. \tag{21}$$

□

Remark 2.6. The estimates in (20) and (21) are not sharp. However, here we are only interested in the order of magnitude. When evaluating the error in the examples, we will always use equation (19) directly.

Corollary 2.7. *Let* (X_n) *be defined by* (10) *and* (11), $\bar{\xi}_p$ *as in* (15), *and assume*

$$R_A(n) \le C_A \cdot \frac{1}{\gamma^n}, \qquad R_b(n) \le C_b \cdot \frac{1}{\gamma^n}, \qquad R_X(n) \le C_X \cdot \frac{1}{\gamma^n},$$

for some $\gamma < 1/\bar{\xi}_p$. *Then, we get*

$$\ell_p(X_n, X) \le C_\gamma \cdot \frac{1}{\gamma^n},$$

where

$$C_\gamma := \|X - \mathbb{E}X\|_p + \frac{\left(C_X + C_b + C_A \|X\|_p\right)}{1 - \bar{\xi}_p \gamma}. \tag{22}$$

Proof. Using Lemma 2.4 we get

$$\ell_p(X_n, X) \le \bar{\xi}_p^n \|X - \mathbb{E}X\|_p + (C_X + C_A \|X\|_p + C_b)\gamma^{-n} \sum_{i=0}^{n-1} \bar{\xi}_p^i \gamma^i, \tag{23}$$

and the assumption on γ implies that both summands are $O(\gamma^{-n})$ with the constant given in the lemma. □

2.2.2 Kolmogorov metric

If we know some properties of the distribution of the fixed point, we can transform the rate of convergence in the ℓ_p metric into a rate for the Kolmogorov metric.

Lemma 2.8. *Let* X_n *be defined by* (10) *and* (11) *and* X *have a bounded density* f_X. *Then, the distance in the Kolmogorov metric can be bounded by*

$$\varrho(X_n, X) \le \left(C_r \, (p+1)^{1/p} \|f_X\|_\infty\right)^{p/(p+1)} \cdot n^{-r \cdot p/(p+1)}$$

with $r \ge 1$ *and* C_r *defined in* (18).

Proof. We use Lemma 5.1 in Fill and Janson (2002), which states, that for X with bounded density f_X and any Y,

$$\varrho(Y, X) \le \left((p+1)^{1/p} \|f_X\|_\infty \cdot \ell_p(Y, X)\right)^{p/(p+1)} \quad \text{for } p \ge 1.$$

Using Corollary 2.5, we get the stated result. □

In some cases, we can give a similar bound, although the density of X is not bounded or no explicit bound is known. Instead, it is sufficient to know a bound for the modulus of continuity of the distribution function F_X of X, defined by

$$\Delta_{F_X}(\delta) := \sup_{x\in\mathbb{R}} \left|F_X(x+\delta) - F_X(x)\right| .$$

Remark 2.9. Lemma 5.1 in Fill and Janson (2002) can easily be extended to cases, when the modulus of continuity of the distribution function of X can be bounded by $\Delta_{F_X}(\delta) \le c \cdot \delta^\alpha$ for some $\alpha \in (0,1]$, $c \in \mathbb{R}_+$. Then,

$$\varrho(Y,X) \le \left(\left(\frac{p}{\alpha}+1\right)^{1/p} \cdot c^{1/\alpha} \cdot \ell_p(Y,X)\right)^{p\alpha/(p+\alpha)} \qquad \text{for } p \ge 1.$$

2.2.3 Approximation of the density

To approximate the density of the fixed-point, we define

$$f_n(x) = \frac{F_n(x+\delta_n) - F_n(x-\delta_n)}{2\delta_n}, \tag{24}$$

where F_n is the distribution function of X_n. For this approximation we can give a rate of convergence, depending on the modulus of continuity of the density of the fixed-point, which is defined by

$$\Delta_{f_X}(\delta) := \sup_{\substack{u,v\in\mathbb{R}\\ |u-v|\le\delta}} \left|f_X(u) - f_X(v)\right|$$

Lemma 2.10. *Let X have a density f_X with modulus of continuity Δ_{f_X} and let (X_n) be defined by (10) and (11). Then, for f_n defined by (24),*

$$\left\|f_n - f_X\right\|_\infty \le \frac{1}{\delta_n}\, \varrho(X_n, X) + \Delta_{f_X}(\delta_n) \qquad \textit{for } \delta_n > 0.$$

Proof.

$$\begin{aligned} \left|f_n(x) - f_X(x)\right| \le &\left|\frac{F_n(x+\delta_n) - F_n(x-\delta_n)}{2\delta_n} - \frac{F(x+\delta_n) - F(x-\delta_n)}{2\delta_n}\right| + \\ &+ \left|\frac{F(x+\delta_n) - F(x-\delta_n)}{2\delta_n} - f_X(x)\right| \\ &\le \frac{1}{\delta_n}\, \varrho(X_n, X) + \frac{1}{2\delta_n}\int_{-\delta_n}^{\delta_n} \left|f_X(x+y) - f_X(x)\right| dy \\ &\le \frac{1}{\delta_n} \varrho(X_n, X) + \frac{1}{\delta_n}\int_0^{\delta_n} \Delta_{f_X}(y)\, dy \end{aligned}$$

□

Corollary 2.11. *Let X have a bounded density f_X, which is Hölder continuous with exponent α, i.e. $\Delta_{f_X}(\varepsilon) \le c \cdot \varepsilon^\alpha$ for some $c > 0$, $\alpha \in (0,1]$. Let (X_n) be an approximation of X as in Corollary 2.5 and f_n be defined by (24) with*

$$\delta_n := L \cdot n^{-r/(\alpha+1)\cdot p/(p+1)}$$

with an $L > 0$. Then, we obtain

$$\|f_n - f_X\|_\infty \le \left(\left(C_r \, (p+1)^{1/p} \, \|f_X\|_\infty \right)^{p/(p+1)} \Big/ L + c\, L^\alpha \right) \cdot n^{-\alpha/(\alpha+1)\cdot r \cdot p/(p+1)}$$

with C_r as defined in (18).

Proof. Combine Lemma 2.8 and Lemma 2.10. □

Remark 2.12. Similarly, we get for exponential discretisation as in Corollary 2.7 with

$$\delta_n := L \cdot \gamma^{-1/(\alpha+1)\cdot n \cdot p/(p+1)},$$

with an $L > 0$, the bound

$$\|f_n - f_X\|_\infty \le \left(\left(C_\gamma \, (p+1)^{1/p} \, \|f_X\|_\infty \right)^{p/(p+1)} \Big/ L + c\, L^\alpha \right) \cdot \gamma^{n\cdot\alpha/(\alpha+1)\cdot p/(p+1)},$$

with C_γ as defined in (22).

Remark 2.13. If X is bounded and bounds for the density f_X and its modulus of continuity are known explicitly, the last result is strong enough to allow, in principle, perfect simulation using von Neumann's rejection method as carried out in Devroye and Neininger (2002). However, we will see in the examples in Section 3 and 4, that the resulting running time is too slow for practical purposes.

Remark 2.14. Lemma 2.10 can be improved for cases, when $X \ge c$ almost surely for some $c \in \mathbb{R}$. If $f_X(c)$ can be approximated at least to an accuracy of $\varrho(X_n, X)/(2\delta_n)$, we approximate the density by

$$f_n(x) := \begin{cases} 0 & \text{for } x < 0, \\ f_X(0) & \text{for } 0 \le x \le \delta_n, \\ \dfrac{F_n(x+\delta_n) - F_n(x-\delta_n)}{2\delta_n} & \text{otherwise.} \end{cases}$$

and can use for the bound the (possibly smaller) modulus of continuity $\Delta^{(c)}_{f_X}$ of f_X on $[c, \infty)$, defined by

$$\Delta^{(c)}_{f_X}(\delta) := \sup_{\substack{u,v \ge c \\ |u-v| \le \delta}} \left| f_X(u) - f_X(v) \right|. \tag{25}$$

To make the bounds of this section explicit, we need a bound for the absolute value and modulus of continuity of the density of the fixed-point. For a simple class of fixed-point equations, we will give universal bounds in the next section. For more complicated cases, those properties have to be derived individually, which we will do for one example in Section 3.

2.3 A class of perpetuities

For fixed-point equations of the form

$$X \stackrel{d}{=} AX + 1 \qquad \text{with } A \geq 0, \tag{26}$$

where A and X are independent, we can bound the density and modulus of continuity of X using the corresponding values of A.

2.3.1 A bound for the density

Lemma 2.15. *Let X satisfy fixed-point equation* (26) *and A have a density f_A. Then X has a density f_X satisfying*

$$f_X(u) = \int_1^\infty \frac{1}{x}\, f_A\!\left(\frac{u-1}{x}\right) f_X(x)dx, \qquad \textit{for } u \geq 1, \tag{27}$$

and $f_X(u) = 0$ otherwise.

Proof. From the fixed-point equation we can see that $X \geq 1$ almost surely. Now let $\mathbb{P}_X$ be the distribution of X; by conditioning on X, we get for any borel set B:

$$\begin{aligned}
\mathbb{P}[X \in B] &= \int_1^\infty \mathbb{P}[Ax + 1 \in B]\, d\mathbb{P}_X(x) \\
&= \int_1^\infty \int_B f_{xA+1}(u) du\, d\mathbb{P}_X(x) \\
&= \int_1^\infty \int_B \frac{1}{x}\, f_A\!\left(\frac{u-1}{x}\right) du\, d\mathbb{P}_X(x) \\
&= \int_B \int_1^\infty \frac{1}{x}\, f_A\!\left(\frac{u-1}{x}\right) d\mathbb{P}_X(x)\; du,
\end{aligned}$$

where we have used Fubini in the last step, because the integrand is product measurable. The claim follows, as this is just the definition of Lebesgue density. □

Corollary 2.16. *Let A have a bounded density f_A. Then X has a density f_X satisfying*

$$\|f_X\|_\infty \leq \|f_A\|_\infty .$$

Proof. Using Lemma 2.15 we get

$$\|f_X\|_\infty \leq \|f_A\|_\infty \cdot \mathbb{E}\left[\frac{1}{X}\right],$$

but $X \geq 1$ implies $\mathbb{E}[1/X] \leq 1$, so the claim follows. □

2.3.2 The modulus of continuity

Corollary 2.17. *Let A have a density f_A, and Δ_{f_A} be its modulus of continuity. Then X has a density f_X, and its modulus of continuity satisfies*

$$\Delta_{f_X}(\delta) \leq \Delta_{f_A}(\delta) \quad \text{for all } \delta > 0.$$

Proof. Using (27), we obtain for any $u, v \in \mathbb{R}$

$$\left|f_X(u) - f_X(v)\right| \leq \int_1^\infty \frac{1}{x} f_X(x) \left| f_A\left(\frac{u-1}{x}\right) - f_A\left(\frac{v-1}{x}\right)\right| dx.$$

But $x \geq 1$ and the modulus of continuity Δ_{f_A} is monotonically increasing by definition, so we can bound

$$\left| f_A\left(\frac{u-1}{x}\right) - f_A\left(\frac{v-1}{x}\right)\right| \leq \Delta_{f_A}\left(\frac{|u-v|}{x}\right) \leq \Delta_{f_A}(|u-v|),$$

and putting this into the last inequality, we get

$$\left|f_X(u) - f_X(v)\right| \leq \mathbb{E}\left[\frac{1}{X}\right] \cdot \Delta_{f_A}(|u-v|).$$

To finish the proof, we notice that $\mathbb{E}[1/X] \leq 1$, because $X \geq 1$ almost surely and take the supremum over all suitable u, v. □

We can extend this result to many practical examples, where A has jumps at points in a set $\mathcal{I}_A$. We use the modulus of continuity $\Delta^{(0)}_{f_A}$ of A and $\Delta^{(1)}_{f_X}$ of X on $[0,\infty)$ and $[1,\infty)$, respectively, as in Remark 2.14 and denote by J_{f_A} the jump function of f_A, defined by

$$J_{f_A}(t) = f_A(t) - \lim_{s \uparrow t} f_A(s).$$

Lemma 2.18. *Let A have a bounded density f_A, which is a càdlàg function and let $\mathcal{I}_A$ be its (countable) set of points of discontinuity. Furthermore let J_{f_A} be the jump function of f_A and $\widetilde{\Delta}^{(0)}_{f_A}$ the modulus of continuity of f_A after removing all jumps. Then,*

$$\Delta^{(1)}_{f_X}(\delta) \le \widetilde{\Delta}^{(0)}_{f_A}(\delta) + \|f_X\|_\infty \sum_{s\in\mathcal{I}_A} \frac{|J_{f_A}(s)|}{s} \cdot \delta \qquad \text{for } \delta > 0.$$

Proof. We first assume that f_A has just one jump in $s_0 > 0$, the general case then follows by induction. So for $1 \le u < v$, we have

$$\big|f_X(u) - f_X(v)\big| \le \int_1^\infty \frac{1}{x} f_X(x) \left| f_A\Big(\frac{u-1}{x}\Big) - f_A\Big(\frac{v-1}{x}\Big)\right| dx.$$

We denote by $\widetilde{f}_A := f_A - \mathbb{1}_{[s_0,\infty)} J_{f_A}(s_0)$ the function remaining after removing the jump in s_0 and define

$$\alpha := \frac{u-1}{s_0} \vee 1, \quad \beta := \frac{v-1}{s_0} \vee 1$$

and divide the range of integration into three parts $(1,\alpha], [\alpha,\beta], [\beta,\infty)$. Now, in the first and last part, f_A equals $\widetilde{f}_A$ and for $x \in [\alpha,\beta]$ we have

$$\left| f_A\Big(\frac{u-1}{x}\Big) - f_A\Big(\frac{v-1}{x}\Big)\right| \le \left| \widetilde{f}_A\Big(\frac{u-1}{x}\Big) - \widetilde{f}_A\Big(\frac{v-1}{x}\Big)\right| + \big|J_{f_A}(s_0)\big|.$$

Putting everything together we get

$$\begin{aligned}
\big|f_X(v) - f_X(u)\big| &\le \\
&\le \int_1^\infty \frac{1}{x} f_X(x) \left| \widetilde{f}_A\Big(\frac{u-1}{x}\Big) - \widetilde{f}_A\Big(\frac{v-1}{x}\Big)\right| dx + \int_\alpha^\beta \frac{1}{x} f_X(x) \big|J_{f_A}(s_0)\big| dx \\
&\le \int_1^\infty \frac{1}{x} f_X(x) \left| \widetilde{f}_A\Big(\frac{u-1}{x}\Big) - \widetilde{f}_A\Big(\frac{v-1}{x}\Big)\right| dx + \|f_X\|_\infty \frac{v-u}{s_0} \big|J_{f_A}(s_0)\big|,
\end{aligned}$$

now if $\widetilde{f}_A$ has no more jumps, we can follow the argumentation in Corollary 2.17 to bound the integral by $\widetilde{\Delta}^{(0)}_{f_A}(v-u)$, otherwise we first repeat this strategy, each time removing one jump and adding one summand on the right. Finally the claim follows by taking the supremum over all $v - u \le \delta$. □

2.4 Implementation

In this section, we will give an algorithm for an approximation satisfying the assumptions in the last section for many important cases. We assume that the distributions of A and b are given by Skorohod representations, i.e. by measurable functions $\varphi, \psi : [0,1] \to \mathbb{R}$, such that $A \stackrel{d}{=} \varphi(U)$ and $b \stackrel{d}{=} \psi(U)$ for $U \sim \mathrm{unif}[0,1]$. Furthermore, we assume that $\|\varphi\|_\infty \leq 1$ and $\|\psi\|_\infty < \infty$ and that both functions are Lipschitz continuous and can be evaluated in constant time. Now we define the discretisation $\langle\cdot\rangle_n$ by

$$\langle Y \rangle_n := \lfloor s(n) \cdot Y \rfloor / s(n), \tag{28}$$

where $s(n)$ can be either polynomial, i.e. $s(n) = n^r$ or exponential, $s(n) = \gamma^n$. Defining

$$\begin{aligned} A^{(n)} &:= \varphi(\langle U \rangle_n) \quad \text{and} \\ b^{(n)} &:= \psi(\langle U \rangle_n)\,, \end{aligned}$$

the conditions on φ and ψ ensure that Corollary 2.5 and 2.7 can be applied.

2.4.1 Algorithm

We keep the distribution of X_n in an array $\mathcal{A}_n$, where $\mathcal{A}_n[k] := \mathbb{P}[X_n = k/s(n)]$ for $k \in \mathbb{Z}$. Note however, that as A and b are bounded, $\mathcal{A}_n[k] = 0$ at least for $|k| > Q_n$, where Q_n is given by the recursive definition $Q_n = \|A\|_\infty Q_{n-1} + \|b\|_\infty$ and $Q_0 = \|X_0\|_\infty = \mathbb{E}[X]$. The core of the implementation is the following update procedure:

procedure UPDATE($\mathcal{A}_{n-1}, \mathcal{A}_n$)
  for $i \leftarrow 0$ **to** $s(n) - 1$ **do**
    for $j \leftarrow -s(n-1) \cdot Q_{n-1}$ **to** $s(n-1) \cdot Q_{n-1} - 1$ **do**

$$\begin{aligned} u &\leftarrow \frac{i}{s(n)} \\ k &\leftarrow \left\lfloor s(n) \cdot \left(\varphi(u) \cdot \frac{j}{s(n-1)} + \psi(u) \right) \right\rfloor \\ \mathcal{A}_n[k] &\leftarrow \mathcal{A}_n[k] + \frac{1}{s(n)} \cdot \mathcal{A}_{n-1}[j] \end{aligned}$$

    end for
  end for
end procedure

The complete code for polynomial discretisation for the example in Section 3, implemented in C++ , can be found in Appendix A.

To approximate the density as in (24) with $\delta_n = d/s(n)$ for some $d \in \mathbb{N}$, we compute a new array $\mathcal{D}_n$ by setting

$$\mathcal{D}_n[k] = \frac{s(n)}{2d} \sum_{j=k-d+1}^{k+d} \mathcal{A}_n[j].$$

2.4.2 Complexity

To measure the complexity of our algorithm, we estimate the number of steps needed to approximate the distribution function and the density up to an accuracy of $O(1/n)$. For the case that X has a bounded density f_X which is Hölder continuous, we will give asymptotic bounds for polynomial as well as exponential discretisation.

Lemma 2.19. *Let X be the solution of the fixed-point equation $X \overset{d}{=} AX + b$, where X and (A, b) are independent and A and b satisfy the assumptions made at the beginning of this section. Furthermore assume that X has a bounded density f_X, which is Hölder continuous with exponent $\alpha \in (0, 1]$, i.e. $\Delta_{f_X}(\delta) \leq c \cdot \delta^\alpha$ for some $c > 0$. Using the algorithm described above with polynomial discretisation $s(n) = n^r$, we can then approximate the distribution function of X to an accuracy of $O(1/n)$ in time $O\big(n^{(2+2/r)\cdot(p+1)/p}\big)$ and the density f_X to the same accuracy in time*

$$O\left(n^{2\,(1+1/\alpha)\cdot(r+1)/r\cdot(p+1)/p}\right).$$

Using exponential discretisation $s(n) = \gamma^n$, approximation to the same accuracy takes time $O\big(n^{(p+1)/p}\log n\big)$ for the distribution function and

$$O\big(n^{(1+1/\alpha)(p+1)/p} \cdot \log n\big)$$

for the density.

Proof. In an execution of update($\mathcal{A}_{k-1}, \mathcal{A}_k$), we have $s(k)$ runs of the outer loop. The assumptions on A and b ensure that $Q_k = O(k)$, so we have $O(k \cdot s(k))$ runs of the inner loop and get for the whole procedure time $O(k \cdot s(k)^2)$. Hence, finding $\mathcal{A}_N$ costs time

$$O\left(\sum_{k=1}^{N} k\, s(k)^2\right) = O\big(N^2 \cdot s(N)^2\big)\,. \tag{29}$$

For discretisations with $s(n) = n^r$ we get a running time of $O(N^{2r+2})$ and Corollary 2.5 ensures that the error of this approximation of the distribution function is of order $O(N^{-rp/(p+1)})$. Setting $n = N^{rp/(p+1)}$, we can see that an approximation of the distribution function to an accuracy of $O(1/n)$ is possible in the time stated in the lemma.

The conditions on the density of X ensure that we can apply Corollary 2.11, and by setting $n = N^{\alpha/(\alpha+1)\cdot r\cdot p/(p+1)}$ this implies the stated bound on the time needed for an approximation of the density to an accuracy of $O(1/n)$.

For $s(n) = \gamma^n$, equation (29) together with Corollary 2.7 implies that we can get an approximation with an error of $O(\gamma^{-Np/(p+1)})$ in time $O(N^2 \cdot \gamma^N)$. This together with Remark 2.12 again ensures the stated running time for an approximation to an accuracy of $O(1/n)$. □

However, in the next section we will see that for the given example and feasible running times, we can get better bounds by using polynomial discretisation than by using exponential discretisation and that the optimal values for p and r are rather small, see Table 1.

3 Example: key exchanges in Quickselect

In this section, we will apply our algorithm to the fixed-point equation

$$X \stackrel{d}{=} UX + U(1-U), \tag{30}$$

where U and X are independent, $U \sim \text{unif}[0,1]$. This equation appears in the analysis of the selection algorithm Quickselect, which is an algorithm to select the element of rank k in a list of n distinct entries and works similar to the sorting algorithm Quicksort. The asymptotic distribution of the number of key exchanges executed by Quickselect, when acting on a random equiprobable permutation of length n and selecting an element of rank $k = o(n)$ can be characterized by the above fixed-point equation, see Hwang and Tsai (2002).

We can use the algorithm of Section 2.4 to get a discrete approximation of the fixed-point. The plot of a histogram, generated using the code in Appendix A with $N = 80$, $r = 3$, can be found in Figure 1.

In the following, we will work out in detail how the bounds in Section 2 can be made explicit for this example. Therefore, we will first derive the needed properties of the solution of the fixed-point equation and after this sketch the implementation and give explicit error bounds for the approximation of the distribution function and density.

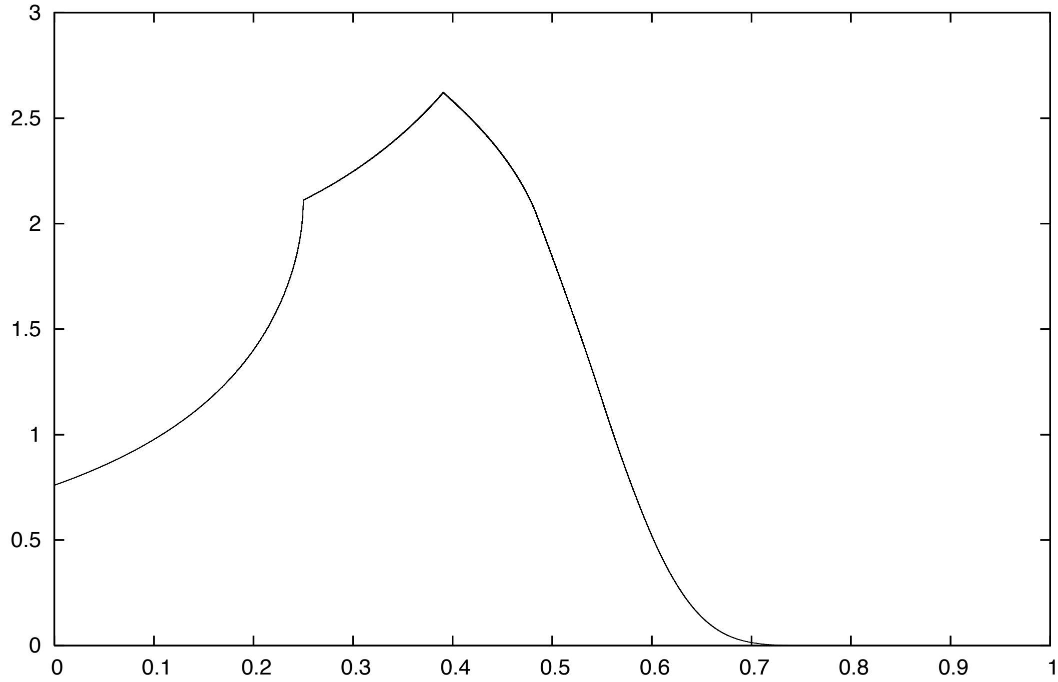


Figure 1: Histogram of approximation for $X = UX + U(1-U)$

3.1 Basic properties

Here, we will derive some basic properties of the limiting distribution, the most important being that X is concentrated on $[0,1]$.

Lemma 3.1. *Let X be a solution of* (30)*. Then, we have $0 \le X \le 1$ almost surely, and the moments are recursively given by $\mathbb{E}[X^0] = 1$ and*

$$\mathbb{E}\big[X^k\big] = (k+1)!\,(k-1)!\sum_{j=0}^{k-1}\frac{\mathbb{E}[X^j]}{j!(2k-j+1)!}, \quad k \ge 1, \tag{31}$$

in particular, $\mathbb{E}[X] = 1/3$.

Proof. The conditions in Remark 2.3 are apparently satisfied, so the sequence given in (4) converges in distribution to the unique fixed-point. But in this sequence, if $0 \le X_0 \le 1$, we get $0 \le X_n(X_0) \le 1$ for all n, hence the same must hold for the limit and therefore the fixed-point almost surely.

To find the moments, we use Remark 2.3 and notice that $\mathbb{E}\big[U^k(1-U)^{k-j}\big]$ is just the Beta-function $\mathrm{B}(k+1, k-j+1)$, so we get

$$\begin{aligned}\mathbb{E}\left[X^k\right] &= \frac{1}{1-\mathbb{E}[U^k]} \cdot \sum_{j=0}^{k-1} \binom{k}{j} \mathbb{E}\left[X^j\right] \mathrm{B}(k+1, k-j+1) \\ &= \frac{k+1}{k} \cdot \sum_{j=0}^{k-1} \frac{k!}{j!(k-j!)} \frac{k!(k-j)!}{(2k-j+1)!} \mathbb{E}\left[X^j\right].\end{aligned}$$

□

Lemma 3.2. *Let X be a solution of* (30)*. Then, for all $\kappa \in \mathbb{N}$ and $\varepsilon > 0$,*

$$\mathbb{P}[X \geq 1-\varepsilon] \leq 2^{(\kappa^2-\kappa)/4} \cdot \varepsilon^{\kappa/2}.$$

Proof. Using that X is concentrated on $[0,1]$, it is easy to show that for all $\varepsilon > 0$,

$$\begin{aligned}\mathbb{P}[X \geq 1-\varepsilon] &= \mathbb{P}[UX + U(1-U) \geq 1-\varepsilon] \\ &\leq \mathbb{P}[X \geq 1-2\varepsilon] \cdot \mathbb{P}\left[U \geq 1-\sqrt{\varepsilon}\right],\end{aligned}$$

and this inequality can be translated into

$$\mathbb{P}[X \geq 1-2\varepsilon] \geq \frac{\mathbb{P}[X \geq 1-\varepsilon]}{\sqrt{\varepsilon}}. \qquad (32)$$

Applying (32) κ times, we get

$$1 \geq \mathbb{P}[X \geq 1-2^\kappa\varepsilon] \geq \frac{\mathbb{P}[X \geq 1-\varepsilon]}{2^{\kappa(\kappa-1)/4} \cdot \varepsilon^{\kappa/2}}.$$

□

3.2 An integral equation for the density

Lemma 3.3. *Let X be a solution of* (30)*. Then X has a Lebesgue density f satisfying $f(t) = 0$ for $t < 0$ or $t > 1$ and*

$$f(t) = 2\int_{p_t}^{t} g(x,t) f(x) dx + \int_t^1 g(x,t) f(x) dx \qquad \text{for } t \in [0,1], \qquad (33)$$

where

$$p_t := 2\sqrt{t} - 1, \qquad g(x,t) := \frac{1}{\sqrt{(1+x)^2 - 4t}}. \qquad (34)$$

Proof. Let $\mathbb{P}_X$ be the distribution of X. Then we get for any Borel set B by conditioning on X

$$\begin{aligned}\mathbb{P}[X \in B] &= \mathbb{P}\big[UX + U(1-U) \in B\big] \\ &= \int_0^1 \mathbb{P}\big[Ux + U(1-U) \in B\big]\, d\mathbb{P}_X(x) \\ &= \int_0^1 \int_B \varphi_x(t) dt\, d\mathbb{P}_X(x) \\ &= \int_B \int_0^1 \varphi_x(t) d\mathbb{P}_X(x)\, dt\end{aligned}$$

where φ_x is a Lebesgue density of $(1+x)U - U^2$. The last step is valid by Fubini's theorem as $(x,t) \mapsto \varphi_x(t)$ is product measurable, cf. (36).

Hence, the function f satisfying

$$f(t) = \int_0^1 \varphi_x(t) f(x) dx \tag{35}$$

is a Lebesgue-density of X.

To find φ_x, we observe that $(1+x)U - U^2 \leq (1+x)^2/4$ and get

$$\begin{aligned}&\mathbb{P}\big[(1+x)U - U^2 \leq t\big] = \\ &\quad = \mathbb{P}\left[U \leq \frac{1+x-\sqrt{(1+x)^2-4t}}{2} \quad \vee \quad U \geq \frac{1+x+\sqrt{(1+x)^2-4t}}{2}\right] \\ &\quad = \begin{cases} 0 & \text{for } t < 0, \\ \dfrac{1+x-\sqrt{(1+x)^2-4t}}{2} & \text{for } 0 \leq t < x, \\ 1 - \sqrt{(1+x)^2-4t} & \text{for } x \leq t \leq (1+x)^2/4, \\ 1 & \text{otherwise.} \end{cases}\end{aligned}$$

To get a density, we differentiate with respect to t and rewrite as a function of x yielding

$$\varphi_x(t) = \begin{cases} \dfrac{2}{\sqrt{(1+x)^2-4t}} & \text{for } 2\sqrt{t}-1 < x \leq t, \\ \dfrac{1}{\sqrt{(1+x)^2-4t}} & \text{for } t < x \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{36}$$

Putting this into (35) we get the stated integral equation. □

Remark 3.4. The integral of $g(x,t)$ with respect to x can explicitly be evaluated:

$$\int g(x,t)\,dx = \log\left(1 + x + \sqrt{(1+x)^2 - 4t}\right) \tag{37}$$

Remark 3.5. For $t = 0$ we get

$$f(0) = \mathbb{E}\left[\frac{1}{1+X}\right] = 0.759947956\ldots$$

Proof. From the integral equation (33) we get for $t = 0$,

$$f(0) = \int_0^1 \frac{1}{1+x} f(x)dx,$$

and by expanding the geometric series we get

$$\mathbb{E}\left[\frac{1}{1+X}\right] = \sum_{k=0}^{\infty} (-1)^k\, \mathbb{E}\left[X^k\right],$$

which we can calculate to any accuracy using for the k-th moments the formula given in Lemma 3.1. □

3.3 A bound for the density

In order to use Lemma 2.8 to bound the deviation of our approximation, we need an explicit bound for the density of the distribution of the fixed point. We will derive a rather rough bound here and will see later, that we can use the resulting bound for our approximation to improve it.

Lemma 3.6. *Let f be the density of X as in Lemma 3.3. Then*

$$\|f\|_\infty \leq 18.$$

Proof. To get an explicit bound for $t \in [0,1]$ we simplify the integral equation and get

$$f(t) \leq 2 \int_{p_t}^1 g(x,t) f(x) dx \tag{38}$$

We already know $f(t)$ for $t < 0$, and we can easily bound $g(x,t)$, if x clearly stays away from p_t. Hence we split the integral into a left part for which we already have a bound for f and a right part, in which we can bound g. For any $\gamma \in (p_t, 1]$, we have

$$f(t) \leq 2 \int_{p_t}^{\gamma} g(x,t)dx + 2 \int_{\gamma}^{1} g(x,t)f(x)dx, \tag{39}$$

where in the second integral, we can use that g is decreasing in x for any fixed t and bound $g(x,t) \leq g(\gamma, t)$.

For $t < 1/4$, we can use that p_t is negative, and set $\gamma = 0$. So the first integral vanishes and only the second remains and we get

$$\begin{aligned} f(t) &\leq 2 \int_0^1 g(x,t)f(x)dx \\ &\leq 2\; g(0,t) \int_0^1 f(x)dx \\ &= \frac{1}{\sqrt{\frac{1}{4} - t}}. \end{aligned} \tag{40}$$

To go on, we set $\gamma = \gamma_t := (p_t + t)/2$ and get with (39)

$$f(t) \leq 2\; \mu_t \int_{p_t}^{\gamma_t} g(x,t)dx + 2\; g(\gamma_t, t) \int_{\gamma_t}^{1} f(x)dx,$$

where $\mu_t := \sup\{f(\tau) :\; \tau \in (p_t, \gamma_t)\}$.

We can calculate the first integral using the integral of g given in (37),

$$\int_{p_t}^{\gamma_t} g(x,t)dx = \log\left(1 + \frac{(1-\sqrt{t})^2 + (1-\sqrt{t})\sqrt{1+6\sqrt{t}+t}}{4\sqrt{t}}\right) =: h(t), \tag{41}$$

and for the second, we obtain

$$\int_{\gamma_t}^{1} f(x)dx \leq \int_{p_t}^{1} f(x)dx = \mathbb{P}\Big[X \geq 1 - 2\Big(1 - \sqrt{t}\Big)\Big].$$

Putting everything together we get

$$f(t) \leq 2\, \mu_t\, h(t) + 4\, \frac{\mathbb{P}\big[X \geq 1 - 2(1-\sqrt{t})\big]}{(1-\sqrt{t})\sqrt{1+6\sqrt{t}+t}}. \tag{42}$$

For $t = 1/4$ we get $\gamma_{1/4} = 1/8$, and $\mu_{1/4} \leq 2\sqrt{2}$ by (40), so

$$f(t) \leq 4\sqrt{2}\; \log\left(1 + \frac{1+\sqrt{17}}{8}\right) + \frac{16}{\sqrt{17}} \leq 7. \tag{43}$$

From the integral equation we get for $0 \leq s < t \leq 1/4$

$$\begin{aligned} f(t) - f(s) = & \int_0^1 \big(g(x,t) - g(x,s)\big) f(x) dx + \\ & + \int_0^s \big(g(x,t) - g(x,s)\big) f(x) dx + \int_s^t g(x,t) f(x) dx \\ & > 0, \end{aligned}$$

so f is strictly increasing on $[0, 1/4]$. Therefore, the bound for $t = 1/4$ extends to all $t \in [0, 1/4] =: I_0$. To go on, we recursively define

$$b_i := \left(\frac{1 + b_{i-1}}{2} \right)^2, \quad b_0 = 0,$$

and

$$\begin{aligned} I_{2k-1} &:= \left(b_k, \frac{b_k + b_{k+1}}{2} \right], \\ I_{2k} &:= \left(\frac{b_k + b_{k+1}}{2}, b_{k+1} \right]. \end{aligned}$$

For each interval I_n we find a corresponding bound M_n for f, using that $p_{b_i} = b_{i-1}$ and therefore $(p_t, \gamma_t) \subset I_{n-1} \cup I_{n-2}$ for $t \in I_n$.

Furthermore we get for $1/4 \leq t \leq 1$ by differentiating the function h defined in (41)

$$h'(t) = c_t \cdot \left(\frac{d}{dt} \frac{(1-\sqrt{t})^2}{4\sqrt{t}} + \frac{d}{dt} \frac{(1-\sqrt{t})\sqrt{1+6\sqrt{t}+t}}{4\sqrt{t}} \right),$$

where $c_t \geq 1$. But the first term is negative and for the second observe that

$$\begin{aligned} \frac{d}{dt}(1-\sqrt{t})\sqrt{1+6\sqrt{t}+t} &= \frac{\big(1-\sqrt{t}\big)\big(3+\sqrt{t}\big) - \big(1+6\sqrt{t}+t\big)}{2\sqrt{t}\sqrt{1+6\sqrt{t}+t}} \\ &= \frac{1-4\sqrt{t}-t}{\sqrt{t}\sqrt{1+6\sqrt{t}+t}} \\ &< 0, \end{aligned}$$

hence $h(t)$ is decreasing.

The second term in (42) can be bounded using Lemma 3.2 with $\kappa = 2$ to get

$$4\,\frac{\mathbb{P}\big[X \geq 1 - 2(1-\sqrt{t})\big]}{(1-\sqrt{t})\sqrt{1+6\sqrt{t}+t}} \leq 4\,\frac{\mathbb{P}\big[X \geq 1 - 2\,(1-\sqrt{t})\big]}{2\,(1-\sqrt{t})} \leq 4\sqrt{2}. \tag{44}$$

So for $t \in I_n = (\alpha_n, \beta_n]$ we have

$$f(t) \leq M_n := \left\lceil 2\,h(\alpha_n)\,\max\{M_{n-1}, M_{n-2}\} + 4\sqrt{2}\right\rceil. \tag{45}$$

Evaluating this we get

$$\begin{aligned} M_0 &= 7, \\ M_1 &= 13, \\ M_2 &= 17, \\ M_3 &= 18, \\ M_4 &= 17. \end{aligned}$$

But for $t > b_3$ we have $h(t) < 2/7$ so (M_n) is decreasing for $n \geq 4$. □

3.4 The modulus of continuity

In order to use Lemma 2.10 to bound the deviation of our approximation of the density of the fixed-point, we will now derive a bound for the modulus of continuity of the density.

Lemma 3.7. *Let f be the density of X as in Lemma 3.3. Then f is Hölder continuous on $[0,1]$ with Hölder exponent $\frac{1}{2}$:*

$$\big|f(t) - f(s)\big| \leq 9\,\|f\|_\infty \sqrt{t-s}, \qquad \text{for } 0 \leq s < t \leq 1. \tag{46}$$

Proof. Using the integral equation given in Lemma 3.3, we have

$$\begin{aligned} \big|f(t) - f(s)\big| \leq 2\left|\int_{p_t}^{t} g(x,t)f(x)dx - \int_{p_s}^{s} g(x,s)f(x)dx\right| + \\ + \left|\int_{t}^{1} g(x,t)f(x) - \int_{s}^{1} g(x,s)f(x)dx\right|. \end{aligned} \tag{47}$$

For $0 \leq s < t \leq 1$ we use the integral of g given in (37) to obtain

$$\begin{aligned} \int_s^t g(x,s)dx &= \log(1 + t + \sqrt{(1+t)^2 - 4s}) - \log(2) \\ &\leq \log(1 + \sqrt{t-s}) \\ &\leq \sqrt{t-s}. \end{aligned}$$

Hence,

$$0 \le \int_s^t g(x,s) f(x) dx \le \|f\|_\infty \sqrt{t-s},$$

and

$$\begin{aligned} 0 \le \int_t^1 \big(g(x,t) - g(x,s)\big) f(x) dx &\le \|f\|_\infty \left(\int_t^1 g(x,t)dx - \int_t^1 g(x,s)dx\right) \\ &\le \|f\|_\infty \left(\int_t^1 g(x,t)dx - \int_s^1 g(x,s)dx + \int_s^t g(x,s)dx\right) \\ &\le \|f\|_\infty \Big(\log(1+\sqrt{1-t}) - \log(1+\sqrt{1-s}) + \sqrt{t-s}\Big) \\ &\le \|f\|_\infty \sqrt{t-s}. \end{aligned}$$

Combining this, we get for the second term in (47)

$$\begin{aligned} &\left|\int_t^1 g(x,t) f(x) dx - \int_s^1 g(x,s) f(x) dx\right| = \\ &\quad = \left|\int_t^1 \big(g(x,t) - g(x,s)\big) f(x) dx - \int_s^t g(x,s) f(x) dx\right| \\ &\quad \le \|f\|_\infty \sqrt{t-s}. \end{aligned}$$

To bound the first term in (47), we split the interval at 1/4. For $s < t \le 1/4$ we use that p_s and p_t are negative and $f(x)$ is increasing, and get

$$\begin{aligned} \left|\int_{p_t}^t g(x,t) f(x) dx - \int_{p_s}^s g(x,s) f(x) dx\right| &= \int_0^t g(x,t) f(x) dx - \int_0^s g(x,s) f(x) dx \\ &\le f(t)\left(\int_0^t g(x,t)dx - \int_0^s g(x,s)dx\right) \\ &= f(t)\Big(\log\big(1+\sqrt{1-4s}\big) - \log\big(1+\sqrt{1-4t}\big)\Big) \\ &\le 2 f(t)\sqrt{t-s} \\ &\le 2 f(1/4)\sqrt{t-s}. \end{aligned}$$

For $1/4 \le s < t \le 1$ we get

$$\begin{aligned} 0 \le \int_{p_t}^t g(x,t) f(x) dx - \int_{p_t}^s g(x,s) f(x) dx &\le \|f\|_\infty \left(\int_{p_t}^t g(x,t)dx - \int_{p_t}^s g(x,s)dx\right) \\ &= \|f\|_\infty \log\left(1 + \frac{\sqrt{t-s}}{\sqrt{t}}\right) \\ &\le \frac{\|f\|_\infty}{\sqrt{t}}\sqrt{t-s} \\ &\le 2\|f\|_\infty \sqrt{t-s} \end{aligned}$$

and

$$0 \le \int_{p_s}^{p_t} g(x,s)f(x)dx \le \|f\|_\infty \int_{p_s}^{p_t} g(x,s)dx$$
$$= \|f\|_\infty \log\left(1 + \frac{\sqrt{t}-\sqrt{s}+\sqrt{t-s}}{\sqrt{s}}\right)$$
$$\le \frac{2\,\|f\|_\infty}{\sqrt{s}}\sqrt{t-s}$$
$$\le 4\,\|f\|_\infty \sqrt{t-s},$$

hence

$$\left|\int_{p_t}^{t} g(x,t)f(x)dx - \int_{p_s}^{s} g(x,s)f(x)dx\right| =$$
$$= \left|\int_{p_t}^{t} g(x,t)f(x)dx - \int_{p_t}^{s} g(x,s)f(x)dx - \int_{p_s}^{p_t} g(x,s)f(x)dx\right|$$
$$\le 4\,\|f\|_\infty \sqrt{t-s}.$$

□

Remark 3.8. The last lemma cannot be substantially improved, as in $t = 1/4$, the density $f(t)$ cannot be Hölder continuous with Hölder exponent $1/2+\varepsilon$ for any $\varepsilon > 0$.

Proof. Using the integral of g given in Remark 3.4, we find that

$$\int_0^t g(x,t)dx = \log(2) - \log\bigl(1 + 2\sqrt{1/4-t}\bigr),$$

and together with the fact that f is increasing on $[0,1/4]$, we get

$$\frac{f(1/4)-f(1/4-h)}{h^{1/2+\varepsilon}} = \frac{1}{h^{1/2-\varepsilon}}\left(\int_0^{1/4} g(x,1/4)\,f(x)dx - \int_0^{1/4-h} g(x,1/4-h)\,f(x)dx\right)$$
$$+ \frac{1}{h^{1/2+\varepsilon}}\left(\int_0^1 \Bigl(g(x,1/4) - g(x,1/4-h)\Bigr)f(x)dx\right)$$
$$\ge \frac{1}{h^{1/2+\varepsilon}}\left(\int_0^{1/4}\Bigl(g(x,1/4) - \mathbb{1}_{[0,1/4-h]}(x)\; g(x,1/4-h)\Bigr)f(x)dx\right)$$
$$\ge \frac{f(0)}{h^{1/2+\varepsilon}}\left(\int_0^{1/4} g(x,1/4)\,dx - \int_0^{1/4-h} g(x,1/4-h)\,dx\right)$$
$$= \frac{f(0)\,\log(1+2\sqrt{h})}{h^{1/2+\varepsilon}}$$
$$\longrightarrow \infty \quad \text{as } h \downarrow 0$$

□

3.5 Implementation and explicit error bounds

We can now combine the bounds for the density and its modulus of continuity with Lemma 2.8 and Lemma 2.10 to bound the deviation of an approximation using the algorithm of Section 2.4 from the solution of the fixed-point equation.

We use the algorithm of Section 2.4 with $N = 80$ and discretisation to $s(n) = n^3$ steps. The implementation in C++ can be found in Appendix A and some remarks on why we chose these values will be made at the end of this section.

To approximate the density f we follow the approach of Remark 2.14 and set

$$f_n(x) := \begin{cases} f(0) & \text{for } 0 \leq x \leq \delta_n, \\ \dfrac{F_n(x+\delta_n) - F_n(x-\delta_n)}{2\delta_n} & \text{for } \delta_n < x \leq 1, \\ 0 & \text{otherwise,} \end{cases}$$

where $f(0)$ is given in Remark 3.5.

Corollary 3.9. *We have $\varrho(X_{80}, X) \leq 1.162 \cdot 10^{-4}$, and $\|f_{80} - f\|_\infty \leq 0.931$. Furthermore, we can improve the bound of Lemma 3.6 and get $\|f\|_\infty \leq 3.561$.*

Proof. We have $C_A = C_b = C_X = 1$, hence combining Lemma 3.6 and Lemma 2.8, we obtain

$$\varrho(X_n, X) \leq \left(\left(\bar{\xi}_p^n \|X\|_p + \left(2 + \|X\|_p\right) \sum_{i=0}^{n-1} \frac{\bar{\xi}_p^i}{(n-i)^r} \right) \cdot (p+1)^{1/p} \cdot \|f\|_\infty \right)^{p/(p+1)}.$$

The moments of X can be computed using Lemma 3.1 and we set $[U]_n := \lfloor n^3 U \rfloor / n^3$, hence

$$\bar{\xi}_p = \|U\|_p = \left(\frac{1}{p+1} \right)^{1/p}.$$

Optimizing over p for $n = 80$, $r = 3$, and $\|f\|_\infty \leq 18$ yields

$$\varrho(X_{80}, X) \leq 5.1842 \cdot 10^{-4} \tag{48}$$

for $p = 12$.

For the density we use Remark 2.14 and as we can give $f(0)$ with the needed accuracy using Remark 3.5, we obtain

$$\|f_n - f\|_\infty \leq \frac{1}{\delta_n} \varrho(X_n, X) + 9 \|f\|_\infty \sqrt{\delta_n},$$

and optimizing over δ_n, using for the Kolmogorov metric the bound in (48), yields

$$\|f_{80} - f\|_\infty \leq 4.512$$

for $\delta_{80} = 3.44 \cdot 10^{-4}$ (averaging 352 values).

We can now use this to improve our bound for $\|f\|_\infty$: Reading off the maximal value of our approximation ($\|f_{80}\|_\infty \leq 2.630$), we can now bound

$$\|f\|_\infty \leq \|f_{80}\|_\infty + \|f_{80} - f\|_\infty \leq 7.142,$$

and this in turn enables us to improve our bounds for the approximation, leading to $\varrho(X_{80}, X) \leq 2.2085 \cdot 10^{-4}$ and $\|f_{80} - f\|_\infty \leq 1.8331$ for $\delta_{80} = 3.6 \cdot 10^{-4}$. Repeating this strategy a few times, we get the stated values for $p = 13$ and $\delta_{80} = 3.7 \cdot 10^{-4}$ (averaging 378 values). □

In Table 1, the resulting error bounds for several possible discretisations with similar running time can be found. Of the given possibilities, the one chosen for the evaluation above seems to be optimal with respect to the theoretical bounds. Note however, that "comparable running time" is no precise notion and depends on many parameters not listed here, including for example the implementation, programming language and hardware architecture used, hence the explicit values chosen for N for the different discretisations can certainly be discussed. On the other hand, especially in the lower half of the table, changing N by ± 1 does change the running time a lot more than the bounds.

Discret.	N	$\varrho(X_N, X)$	opt. p	$s(N)$
n	22000	0.00178	14	22000
n^2	430	0.00025	16	184900
n^3	80	0.00012	13	512000
n^4	30	0.00050	3	810000
1.5^n	35	0.00070	3	1456110
1.7^n	27	0.00187	2	1667712

Table 1: table of bounds for $\varrho(X_n, X)$ for comparable total running time

(Using a realistic bound of $\|f\|_\infty \leq 2.7$ would give $\varrho(X_{80}, X) \leq 8.9809 \cdot 10^{-5}$ ($p = 13$) and $\|f_{80} - f\|_\infty \leq 0.7101$ ($\delta_{80} = 3.8 \cdot 10^{-4}$, 390 values))

4 Further examples

In this chapter we apply our algorithm for approximating distribution function and density of fixed points from Section 2.4 to various other fixed-point equations. For some of these equations, solutions are more or less explicitly known by expressions for their densities or relations satisfied by their densities. Thus, for these equations we can compare the approximations of our algorithms with the true densities and distribution functions and evaluate the error being made. This enables us to get an idea of the quality of the general error bounds proved in Section 2. It appears that in these examples the error bounds from Section 2 are rather loose and that the approximation is much better than indicated by our bounds.

4.1 Interval Splitting

Fixed-point equations of the form studied here arise in the analysis of nested random intervals. These are sequences $\big([L_n, R_n]\big)$ of random intervals defined by $[L_0, R_0] := [0,1]$ and some randomized recursive procedure. One is interested in the limit X to which the intervals shrink almost surely. For example, Chen, Goodman, and Zame (1984) and Chen, Lin, and Zame (1981) considered the following recursive interval splitting procedure: Fix $q \in [0,1]$ and set $[L_0, R_0] := [0,1]$. If $[L_n, R_n]$ is already defined, then split $[L_n, R_n]$ by an independent and uniformly on $[L_n, R_n]$ distributed random variable Y and choose independently the larger of the two subintervals $[L_n, Y], [Y, R_n]$ with probability q to be $[L_{n+1}, R_{n+1}]$, otherwise the smaller one. In the papers mentioned, the authors prove that $([L_n, R_n])$ shrinks to a limit X almost surely, where X has the beta$(2,2)$ distribution if $q = 1$ and the arcsin$(=$ beta$(1/2, 1/2))$ distribution if $q = 1/2$ (see also Devroye, Letac, and Seshadri (1986)).

In the analysis of this interval splitting procedure it is convenient to represent a unif$[0,1]$ distributed random variable in the form

$$G\frac{1+U}{2} + (1-G)\frac{1-U}{2}$$

with independent G, U with $U \sim$ unif$[0,1]$ and $G \sim$ Be$(1/2)$, the Bernoulli distribution with probability $1/2$ on the point 1. It was shown in Neininger (2001), where mainly rates of convergence of such interval splitting schemes are estimated, that the point X to which the intervals shrink has a distribution that can be characterized as the fixed-point of (1), where

$$A = G' \frac{1+U}{2} + (1-G') \frac{1-U}{2},$$
$$b = G'(1-G) \frac{1-U}{2} + (1-G')G \frac{1+U}{2},$$

with G, G' and U independent and $G' \sim \mathrm{Be(q)}$, $G \sim \mathrm{Be}(1/2)$, $U \sim \mathrm{unif}[0,1]$.

Subsequently, we will apply our method to approximate these fixed-points for the cases $q = 1$, $q = 1/2$, and $q = 0$. Since we know the fixed-points in the cases $q = 1$ and $q = 1/2$ to be the beta$(2,2)$ distribution and the arc sine distribution respectively, we can explicitly quantify the distance of our approximations to the true density and distributions and also compare these errors with the error bounds implied by our general estimates of Section 2, see Sections 4.1.1 and 4.1.2. Section 4.1.3 has plots for the case $q = 0$. Here the limit X has no well-known distribution and it seems to be difficult to derive explicit expressions for characteristics. Properties of this distribution and generalizations were derived in Herz (1988).

4.1.1 $q = 1$

We first look at the case $q = 1$, where we can simplify the fixed-point equation to get

$$X \stackrel{d}{=} \frac{1+U}{2} X + G \frac{1-U}{2},$$

where G, U, and X are independent, $G \sim \mathrm{Be}(1/2)$ and $U \sim \mathrm{unif}[0,1]$. To approximate the fixed-point, we modify the algorithm of Section 2.4 by evaluating the two cases $G = 0$ and $G = 1$ in the inner loop. And as the approximated function is symmetric, we use a symmetric discretisation for (A, b) instead of (28), setting

$$\langle U \rangle_n := (2 \lfloor s(n)U \rfloor + 1)/2s(n). \tag{49}$$

Doing so, we get for $n = 50$ and $s(n) = n^3$ the distribution function shown in Figure 2.

To compute the bounds as given in Section 2, we can set $C_A = C_b = 1/4$, $\bar{\xi}_p = \|A\|_p$, and A is uniform distributed on $[1/2, 1]$, so

$$\|A\|_p^p = \frac{2^{p+1} - 1}{2^p\,(p+1)} \qquad \text{for } p \in \mathbb{N},$$

and we know that X is beta$(2,2)$ distributed, so we can give the moments directly,

$$\|X\|_p^p = \prod_{s=0}^{p-1} \frac{2+s}{4+s}, \qquad \text{for } p \in \mathbb{N}.$$

Furthermore, we know the density $f(x) = 6\,x(1-x)$, so $\|f\|_\infty \leq 1.5$. We can now use Lemma 2.8 and Corollary 2.5 to obtain

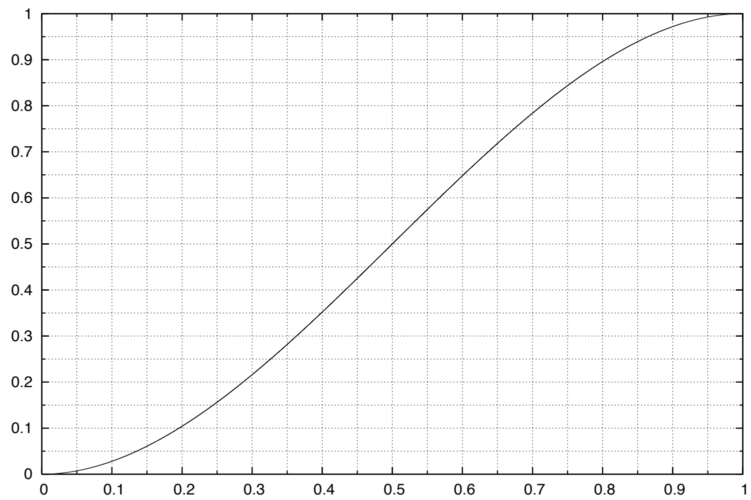


Figure 2: Distribution function of approximation for $q = 1$

$$\varrho(X_n, X) \le \left((p+1)^{1/p} \cdot 1.5 \cdot \left(\|A\|_p^n \|X\|_p + \frac{5 + \|X\|_p}{4} \sum_{i=0}^{n-1} \frac{\|A\|_p^i}{(n-i)^3} \right) \right)^{\frac{p}{p+1}},$$

which we can evaluate for $n = 50$ and minimize over p to get $p_{\min} = 5$ and

$$\varrho(X_{50}, X) \le 0.001043. \tag{50}$$

As we know the limit distribution, we can now compare this bound to the actual error. We are approximating a continuous, monotone function by a step function with step size $1/n$, so the maximal deviation will occur at the borders of the steps. A plot of the maximal distance between the discrete distribution function and the theoretical distribution function on each step can be found in Figure 3.

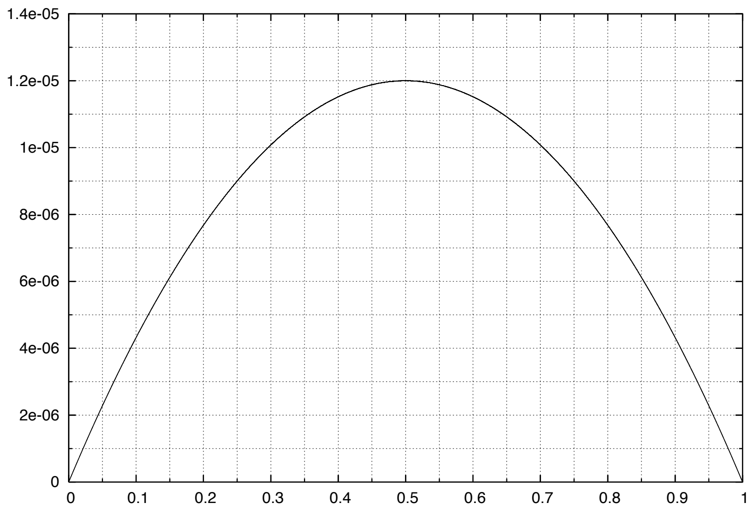


Figure 3: Error of distribution function of approximation for $q = 1$

It is in fact quite exactly what we would expect for a discretisation of step size $1/n^3$, applying our discretisation $\langle \cdot \rangle_n$ to the fixed-point X, where the error is maximal at the left border of a step, and there about equal to the value of the derivative on this step divided by n^3. In Figure 4, the deviation of our approximation from such

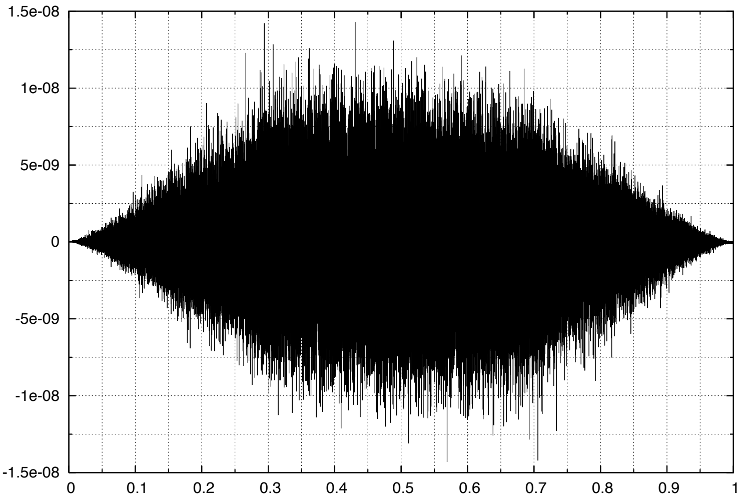


Figure 4: Deviation of discrete distribution function from discretisation of beta$(2, 2)$-distribution function

a discretisation can be found. Putting this together, the error of our approximation is at most $1.2015 \cdot 10^{-5}$, which is significantly less than the stated bound.

Now we look at the density. The histogram of the discrete approximation is shown in Figure 5. The modulus of continuity of the density of the beta$(2, 2)$ distribution

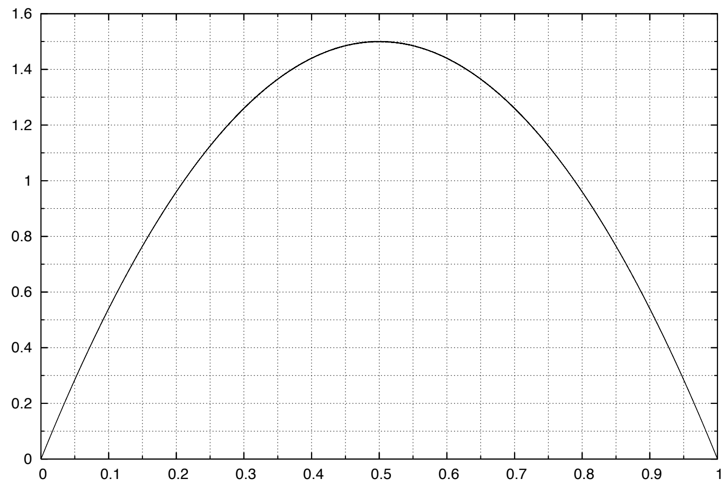


Figure 5: Histogram of approximation for $q = 1$

can be bounded by $\Delta_f(\varepsilon) \leq 6\,\varepsilon$ for all positive ε. So for the function f_n, which we get by averaging over $2\delta_n$ as in (24), we get with Lemma 2.10

$$\|f_n - f\|_\infty \leq \frac{1}{\delta_n} \varrho(X_n, X) + 6\,\delta_n.$$

We evaluate for $n = 50$, use the bound in (50), and minimizing over δ_n we obtain

$$\|f_n - f\|_\infty \leq 0.1583$$

for $\delta_{50} = 0.01318$, so we take the average over 3 296 values. The deviation of this approximation from the density is shown in Figure 6.

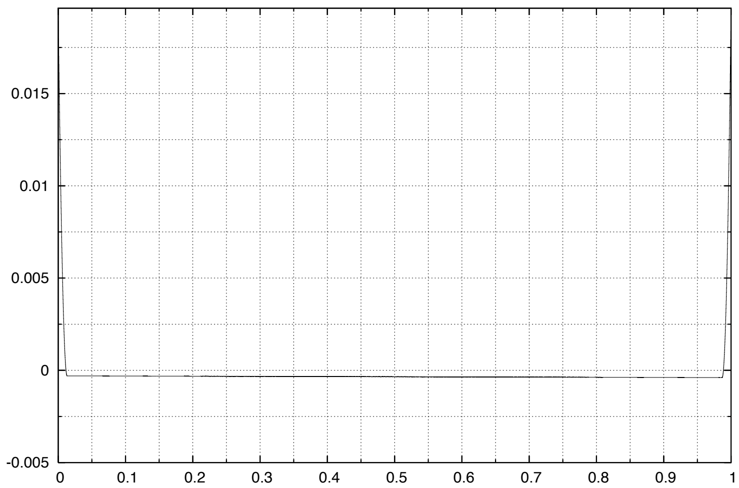


Figure 6: deviation of approximation from density for $q = 1$

We seem to take the average over way to many values, so the smoothing enlarges the error, especially at the borders of the domain, instead of reducing it.

4.1.2 $q = 1/2$

For the case $q = 1/2$, the algorithm is similar to the previous case, we only have to take into account two Bernoulli distributed variables, so we now evaluate four cases in the inner loop. Again, we use the symmetric discretisation $\langle U \rangle_n$ given in (49). With this, we get for $n = 50$ and $s(n) = n^3$ the distribution function shown in Figure 7.

For the error bounds, we again get $C_A = C_b = 1/4$. This time, A is uniform distributed on $[0, 1]$, hence $\bar{\xi}_p^p = \|A\|_p^p = 1/(p+1)$. Furthermore, we know that X is beta$(1/2, 1/2)$ distributed, hence $\|X\|_p^p = \prod_{q=0}^{p-1}(1/2+r)/(1+r)$. In this example, X has no bounded density, but the modulus of continuity of the distribution function

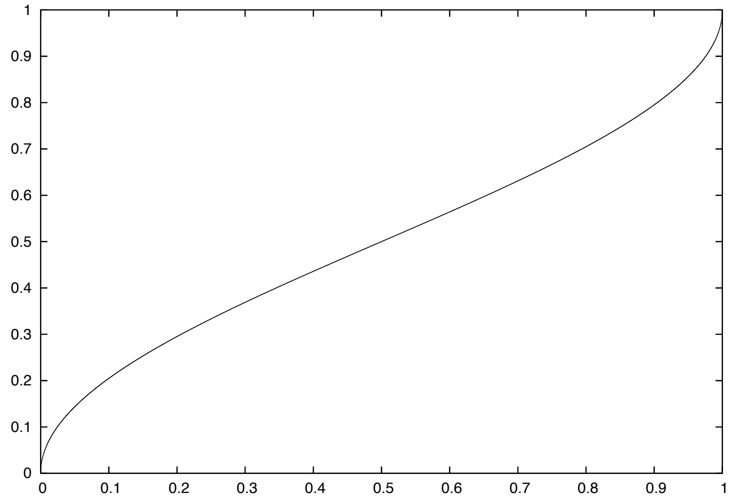


Figure 7: Distribution function of approximation for $q = 1/2$

of X can be bounded by $\Delta_X(\delta) \leq \sqrt{8}/\pi \cdot \sqrt{\delta}$ (see Neininger (2001, page 805)). Using Remark 2.9, we get

$$\varrho(X_n, X) \leq \left((2p+1)^{1/p} \cdot 8/\pi^2 \cdot \left(\|X\|_p \, \|A\|_p^n + \frac{5 + \|X\|_p}{4} \sum_{i=0}^{n-1} \frac{\|A\|_p^i}{(n-i)^3} \right) \right)^{\frac{p}{2p+1}},$$

and minimizing over p for $n = 50$ yields $p_{\min} = 7$ and $\varrho(X_{50}, X) \leq 0.01142$.

The distribution function of X is $F_X(x) = 2/\pi \arcsin(\sqrt{x})$ for $x \in [0,1]$, and comparing this to the discrete distribution function, we get a maximal deviation of $1.8 \cdot 10^{-3}$ at 0, whereas it is of significantly lower order in the main part. A plot of this deviation can be found in Figure 8.

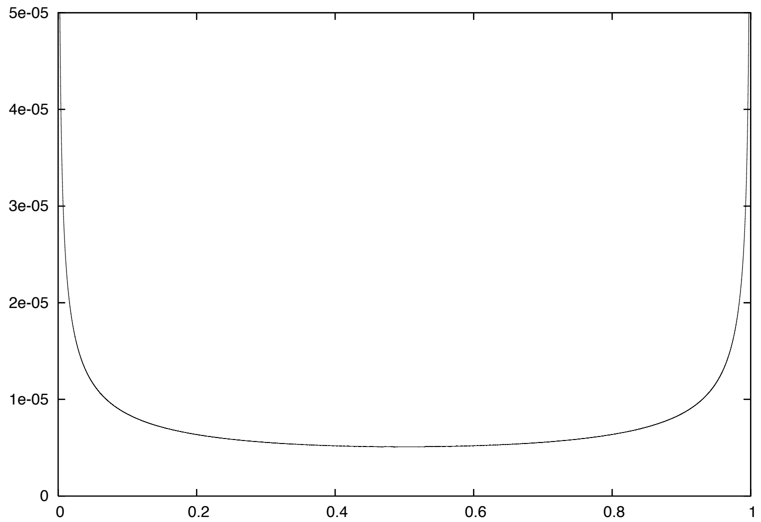


Figure 8: Error of distribution function for $q = 1/2$

We get an approximation of the density using the histogram shown in Figure 9, but

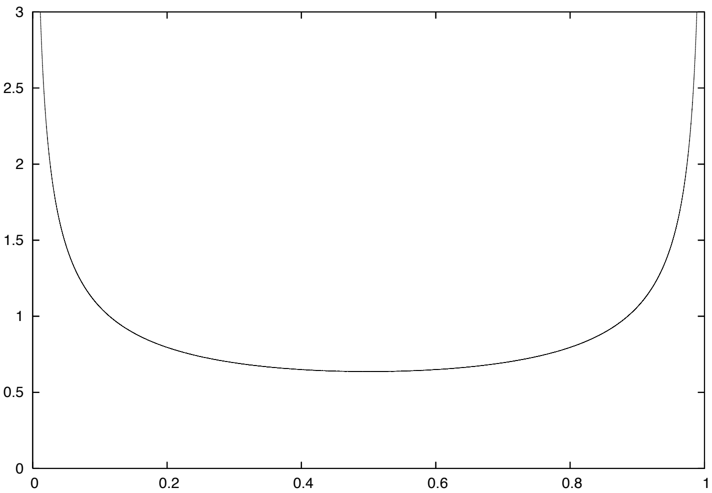


Figure 9: Histogram of approximation for $q = 1/2$

we cannot give a global bound for the error, as the density grows to infinity at 0 and 1. Hence, our theoretical bound is not applicable here, but in Figure 10 we can see, that the deviation of the histogram from a corresponding discretisation of the density of X on the interval $[0.005, 0.995]$ is quite small, so the algorithm still works well in this case.

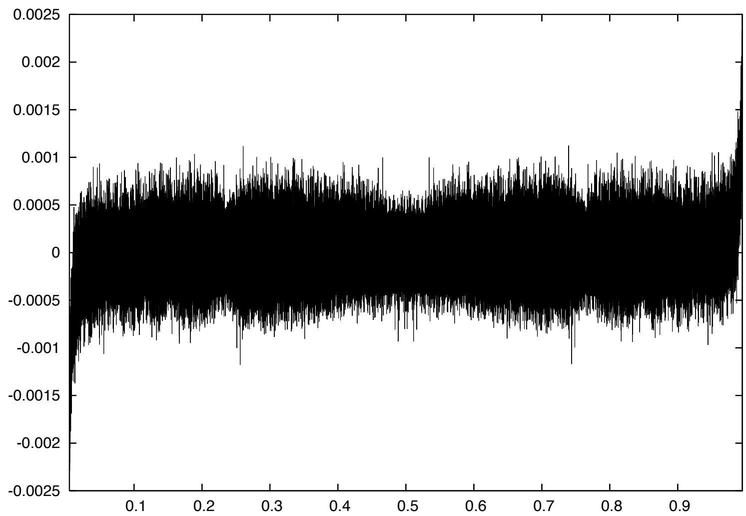


Figure 10: Deviation of histogram from discretisation of density for $q = 1/2$, on $[0.005, 0.995]$

4.1.3 $q = 0$

For $q = 0$ we can simplify the fixed-point equation to get

$$X \stackrel{d}{=} \frac{1-U}{2} X + G \frac{1+U}{2},$$

where G, U, and X are independent, $G \sim \mathrm{Be}(1/2)$, $U \sim \mathrm{unif}[0, 1]$.

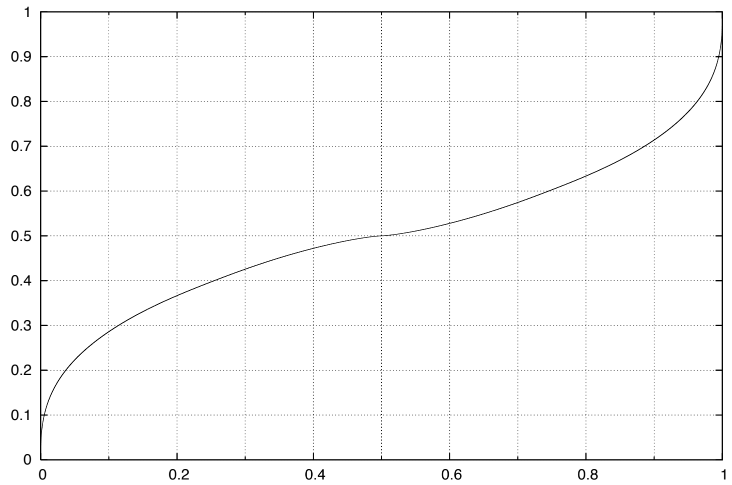


Figure 11: Distribution function of approximation for $q = 0$

We can apply our algorithm as in the previous examples and get the discrete distribution function and the histogram shown in Figure 11 and Figure 12 respectively. However, we cannot give error bounds for this case, because the density is unbounded and we do not know the modulus of continuity of the distribution function of X.

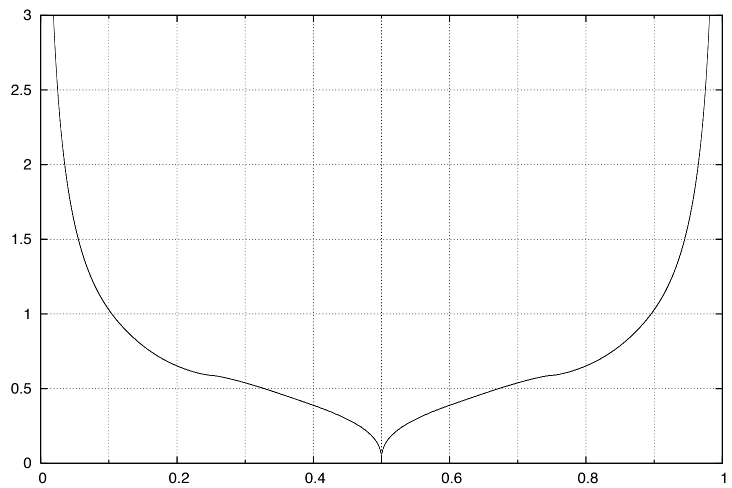


Figure 12: Histogram of approximation for $q = 0$

4.2 Dickman distribution

The Dickman distribution $\mathcal{L}(X)$ is given as the unique solution of the stochastic fixed-point equation

$$X \stackrel{d}{=} UX + U,$$

with X and U independent and U being unif$[0,1]$ distributed. The shifted random variable $Y = X + 1$ satisfies the equation

$$Y \stackrel{d}{=} UY + 1,$$

hence for the (shifted) Dickman distribution, the bounds of Section 2.3 can be applied.

The density f_X of X can be described by a delayed differential equation. We have $f_X(x) = e^{-\gamma}\varphi(x)$ with Euler's constant γ, where φ is given by $\varphi(x) = 0$ for $x < 0$, $\varphi(x) = 1$ for $0 \le x \le 1$ and

$$\varphi'(x) = -\frac{\varphi(x-1)}{x}, \quad x > 1. \tag{51}$$

For properties of φ see Tenenbaum (1995, § III.5.4). The Dickman distribution originated in the analysis of largest prime factors of random integers in Dickman (1930), but later on appeared in various areas of mathematics, for example in the analysis of the selection algorithm Quickselect, which we already encountered in Section 3. It is the limit distribution of the number of key comparisons when acting on a random equiprobable permutation of length n and selecting a rank k of order $k = o(n)$, see Hwang and Tsai (2002), where also references to various further occurrences of the Dickman distribution can be found.

To approximate Y, we use the algorithm of Section 2.4, again using $s(n) = n^3$, $n = 50$, and $\langle U \rangle_n$ as in (49), but as Y is not bounded, our running time is now of order $O(n^8)$, because $Q_n = O(n)$.

To compute the error bounds, we use $C_A = 1/2$, $C_b = 0$, and $\bar{\xi}_p^p = \|U\|_p^p = 1/(p+1)$ and obtain with (19)

$$\ell_p(Y_n, Y) \le \|Y\|_p \frac{1}{(p+1)^{n/p}} + \left(\frac{2 + \|Y\|_p}{2}\right) \sum_{i=1}^{n-1} \frac{1}{(p+1)^{i/p}(n-i)^3}.$$

We use Remark 2.3 to compute the moments of Y

$$\mathbb{E}[Y^p] = \frac{(p+1)!}{p} \sum_{k=0}^{p-1} \frac{\mathbb{E}\left[Y^k\right]}{(k+1)!(p-k)!}$$

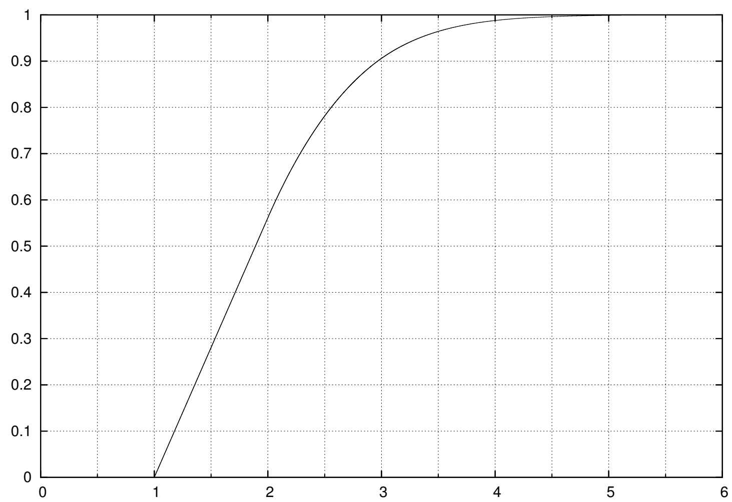


Figure 13: Approximation of distribution function for $Y \overset{d}{=} UY+1$

and using $\|f_Y\|_\infty \leq e^{-\gamma}$ we get $p_{\min} = 7$ and

$$\varrho(Y_{50}, Y) \leq 2.677 \cdot 10^{-4}.$$

To compute the actual error, we have to evaluate the distribution function of the dickman distibution. But this time, the density is only given implicitly by the delayed differential equation (51). The solution of this equation can be found iteratively, but the result gets complicated rather soon. So we evaluated here only the interval $[1, 3]$ and the deviation of our approximation on this interval can be found in Figure 14. The error is again clearly smaller than the bound, but this time,

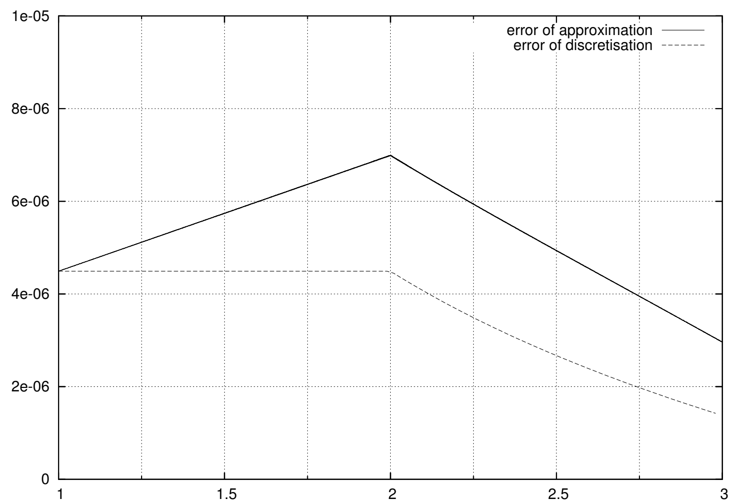


Figure 14: Deviation of approximation of distribution function

we do have some "systematic" error, so the approximation differs from a direct discretisation of the limiting function, as indicated in Figure 14.

The histogram of the discrete approximation can be found in Figure 15.

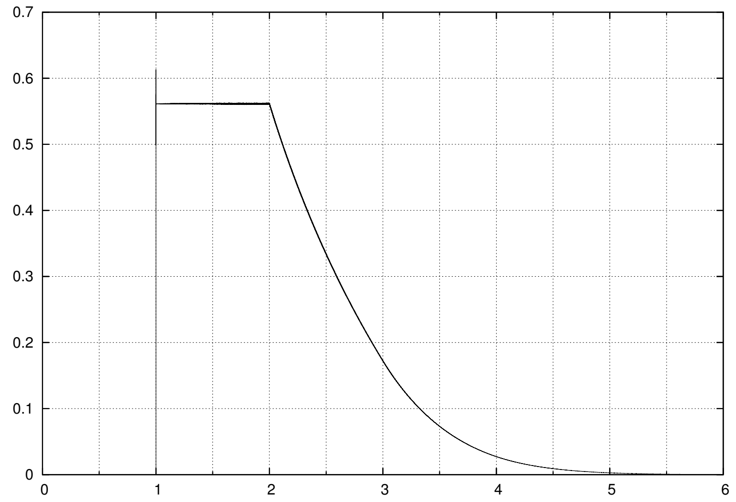


Figure 15: Histogram of approximation for $Y \overset{d}{=} UY+1$

To approximate the density, we use Remark 2.14 with $c = 1$. We have $f_X(1) = \mathrm{e}^{-\gamma}$ and the modulus of continuity of the density is bounded by $\Delta^{(1)}_{f_X}(\varepsilon) \leq \mathrm{e}^{-\gamma}\varepsilon$, and minimizing over δ_n we get

$$\|f_n - f_X\|_\infty \leq 0.02452$$

for $\delta_n = 0.021835$, taking the average over 5458 values. In Figure 16, we can see that again the main error of this approximation is induced by the smoothing and that the actual deviations are significantly smaller than the error bound.

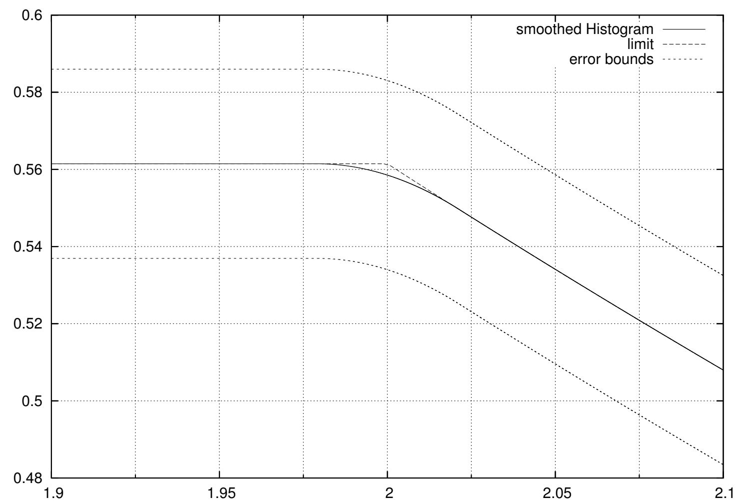


Figure 16: approximation of density with error bounds for $Y \overset{d}{=} UY+1$

Appendix A C++ Code

```
#include <iostream>
#include <vector>
#include <math.h>

typedef std::vector<double> Vector;

static unsigned int r; // Parameter for discretisation

inline unsigned int numSteps(unsigned int n) {
  // returns the number of steps per unit.
  return static_cast<unsigned int>(pow(n,r));
}

Vector update(const Vector &v, unsigned int n) {
  // This is equivalent to the pseudocode in section 2.4.1.
  // To reduce running time, operations were drawn out of
  // the loops where possible and trailing zeros removed.
  Vector res;
  const unsigned int stepCount = numSteps(n);
  const double newStepSize = 1.0/stepCount;
  const double oldStepSize = 1.0/numSteps(n - 1);
  while (res.size() < stepCount) {res.push_back(0.0);}
  for (unsigned int i = 0; i < stepCount; ++i) {
    const double phu = i * oldStepSize;
    const double psu = i * (1.0 - i * newStepSize);
    for (unsigned int j=0; j < v.size(); ++j){
      unsigned int k;
      k = static_cast<unsigned int>(phu * j + psu);
      res[k] += v[j];
    }
  }
  Vector::iterator last = res.begin();
  Vector::iterator it;
  for (it = res.begin(); it != res.end(); ++it) {
    if ((*it *= newStepSize) != 0) {last = it;}
  }
  res.erase(++last, res.end());
  return res;
}
```

```
Vector distrFun(const Vector &v) {
  // calculate distribution function for probabilities in v
  Vector F;
  Vector::const_iterator it = v.begin();
  F.push_back(*it++);
  for ( ; it != v.end(); ++it) {
    F.push_back(F.back() + *it);
  }
  return F;
}

inline void print(const Vector &v) {
  Vector::const_iterator it = v.begin();
  for ( ; it != v.end(); ++it) {
    std::cout << *it << std::endl;
  }
}

int main(int, char **) {
  unsigned int N;
  std::cin >> N >> r; //parameters are read from StdIn
  Vector cur;
  cur.push_back(1.0); //initialize with [1] for n=1
  unsigned int n;
  for (n = 1; n < N; ) {
    cur = update(cur, ++n);
  }
  std::cout.precision(10);
  print(cur);
  std::cout << std::endl << std::endl;
  print(distrFun(cur));
  std::cout << std::endl << std::endl;
  return 0;
}
```

Appendix B Table for distribution function of key exchanges in Quickselect

Table 2 was generated using the code given in Appendix A with $N = 80$ and $r = 3$.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	0.000	0.008	0.016	0.024	0.032	0.040	0.049	0.058	0.067	0.076
0.10	0.086	0.096	0.106	0.117	0.127	0.139	0.150	0.162	0.175	0.188
0.20	0.202	0.216	0.231	0.247	0.265	0.284	0.305	0.326	0.348	0.370
0.30	0.392	0.415	0.438	0.461	0.485	0.509	0.534	0.558	0.584	0.610
0.40	0.636	0.661	0.687	0.711	0.735	0.759	0.782	0.804	0.825	0.846
0.50	0.865	0.883	0.899	0.914	0.928	0.940	0.951	0.961	0.969	0.976
0.60	0.982	0.986	0.990	0.993	0.995	0.997	0.998	0.999	0.999	1.000
0.70	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.80	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.90	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 2: distribution function of $X \stackrel{d}{=} UX + U(1-U)$

References

Arratia, R., Barbour, A. D., and Tavaré, S. (2003) Logarithmic combinatorial structures: a probabilistic approach. EMS Monographs in Mathematics. *European Mathematical Society (EMS), Zürich*, 2003. xii+363 pp.

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

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

Chen, R., Lin, E., and Zame, A. (1981) Another arc sine law. *Sankhyā Ser. A* **43**, 371–373.

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

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

Devroye, L., Letac, G., and Seshadri, V. (1986) The limit behavior of an interval splitting scheme. *Statist. Probab. Lett.* **4**, 183–186.

Devroye, L. and Neininger, R. (2002) Density approximation and exact simulation of random variables that are solutions of fixed-point equations. *Adv. in Appl. Probab.* **34**, 441–468.

Dickman, K. (1930) On the frequency of numbers containing prime factors of a certain relative magnitude. *Arkiv for Matematik, Astronomi och Fysik* **22**, 1–14.

Donnelly, P. and Grimmett, G. (1993) On the asymptotic distribution of large prime factors. *J. London Math. Soc.* **47**, 395–404.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997) Modelling extremal events. For insurance and finance. Applications of Mathematics (New York), 33. *Springer-Verlag, Berlin*, 1997. 648 pp.

Fill, J. A. and Janson, S. (2002) Quicksort asymptotics. *J. Algorithms* **44**, 4–28.

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

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

Herz, C. (1988) Splitting intervals. *Statist. Probab. Lett.* **7**, 3–7.

Hwang, H.-K. and Tsai, T.-H. (2002) Quickselect and the Dickman function. *Combin. Probab. Comput.* **11**, 353–371.

Mahmoud, H., Modarres, R., and Smythe, R.T. (1995) Analysis of QUICKSELECT: an algorithm for order statistics. *RAIRO Inform. Théor. Appl.* **29**, 255–276.

Neininger, R. (2001) Rates of convergence for products of random stochastic 2×2 matrices. *J. Appl. Probab.* **38**, 799–806.

Tenenbaum, G. (1995) Introduction to analytic and probabilistic number theory. Translated from the second French edition (1995) by C. B. Thomas. Cambridge Studies in Advanced Mathematics, 46. *Cambridge University Press, Cambridge*, 1995. xvi+448 pp.

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

Xuan, T. Z. (1993) On the asymptotic behavior of the Dickman-de Bruijn function. *Math. Ann.* **297**, 519–533.