Information Processing Letters 113 (2013) 381–385

Contents lists available at SciVerse ScienceDirect

Information Processing Letters

www.elsevier.com/locate/ipl

A strongly polynomial algorithm for criticality of branching processes and consistency of stochastic context-free grammars ☆

Javier Esparza [a,*], Andreas Gaiser [a,1], Stefan Kiefer [b,2]

[a] Technische Universität München, Germany
[b] University of Oxford, UK

ARTICLE INFO

Article history:
Received 30 August 2012
Accepted 13 February 2013
Available online 5 March 2013
Communicated by J. Torán

Keywords:
Algorithms
Branching processes
Stochastic context-free grammars
Exact matrix computations

ABSTRACT

We provide a strongly polynomial algorithm for determining whether a given multi-type branching process is subcritical, critical, or supercritical. The same algorithm also decides consistency of stochastic context-free grammars.

1. Introduction

Multi-type branching processes (MBPs) are stochastic processes modeling populations in which the individuals of a generation produce a random number of children of different *types* or *species* in the next generation. Individuals can be elementary particles, genes, animals, or program threads [1,2].

MBPs are classified into *subcritical*, *critical*, and *supercritical*, depending on the *spectral radius* of a certain matrix, the *expectation matrix*. This division plays a central role, since many theorems assume that the process belongs to one of these classes. In particular, criticality is strongly related to the *extinction probability*: under some weak conditions, the population of subcritical and critical processes goes ultimately extinct with probability 1, while for supercritical processes the extinction probability is strictly smaller than 1.

We study the computational complexity of the *classification problem*: deciding whether a given MBP is subcritical, critical, or supercritical. By definition, the problem consists of deciding if the spectral radius of the expectation matrix is smaller than, equal to, or larger than one [2]. Etessami and Yannakakis have observed in [3] that this problem reduces to feasibility of a linear programming problem (LP-problem). While LP-problems can be solved in polynomial time, no strongly polynomial algorithms are known: the number of arithmetic operations to be performed depends on the size of the input, which quickly degrades the performance of LP-based classificators. We show that LP can be avoided by reducing the classification problem to the problem of solving a system of linear equations. In particular, this leads to an algorithm with $O(n^3)$ arithmetical operations, where n is the dimension of the matrix, independently of the size of the entries.

Stochastic context-free grammars (SCFGs) are context-free grammars whose rules are weighted with probabilities. They are applied in diverse areas such as natural language

☆ Some results of this note are contained in the same authors' conference paper *Computing least fixed points of probabilistic systems of polynomials*, in: *Proceedings of STACS*, 2010, pp. 359–370.

* Corresponding author.
E-mail address: esparza@in.tum.de (J. Esparza).

[1] Supported by the DFG-Graduiertenkolleg 1480 (PUMA).
[2] Supported by the EPSRC.

0020-0190/$ – see front matter
http://dx.doi.org/10.1016/j.ipl.2013.02.015

processing [4], security [5], and biological sequence analysis [6]. An SCFG is called *consistent* if it generates a terminal string with probability 1. Consistency is a fundamental characteristic of SCFGs, and plays a central role in computations on SCFGs, see e.g. [7]. An SCFG is naturally associated to an MBP, so that the SCFG is consistent if and only if the extinction probability of the MBP is 1. As a consequence, our strongly polynomial algorithm also allows to decide consistency of SCFGs.

The note is organized as follows. In Section 2 we present our strongly polynomial algorithm for comparing the spectral radius to 1. Sections 3 and 4 explain the application of our algorithm to MBPs and SCFGs, respectively. Finally, in Section 5 we apply the algorithm to a neutron scattering process taken from [1], which, loosely speaking, studies when a ball of plutonium becomes an atomic bomb.

Related work. The computational complexity of problems related to branching process has been recently studied by Etessami, Stewart, and Yannakakis [8,9]. In particular, they prove that the extinction probability can be approximated in polynomial time. This is nicely complemented by our result, which shows how to decide in strongly polynomial time whether this probability is exactly 1.

2. Main result

For a square matrix M, we denote by $\rho(M)$ its *spectral radius*, i.e., the largest absolute value of the eigenvalues of M. A matrix is *nonnegative* if all its entries are nonnegative. In this section we show:

Theorem 2.1. *Given a nonnegative matrix $M \in [0,\infty)^{n\times n}$, one can decide in strongly polynomial time and with $O(n^3)$ arithmetic operations whether $\rho(M) < 1$ or $\rho(M) = 1$ or $\rho(M) > 1$.*

We need some notation. Let T be a finite set of indices with $|T| = n \geqslant 1$. For technical convenience we view (square) matrices as elements of $\mathbb{R}^{T\times T}$ and assume $M \in [0,\infty)^{T\times T}$. We write I for the identity matrix. We use bold letters for designating (column) vectors, e.g. $\boldsymbol{v} \in \mathbb{R}^T$. If the dimension is clear from the context, we write $\mathbf{0}$ (resp. $\mathbf{1}$) for the vector $(0,\dots,0)^\top$ (resp. $(1,\dots,1)^\top$), where $^\top$ denotes transpose. We write $\boldsymbol{v}_X$ for $\boldsymbol{v}(X)$ where $X \in T$. We write $\boldsymbol{v} = \boldsymbol{w}$ (resp. $\boldsymbol{v} \leqslant \boldsymbol{w}$ resp. $\boldsymbol{v} \prec \boldsymbol{w}$) if $\boldsymbol{v}_X = \boldsymbol{w}_X$ (resp. $\boldsymbol{v}_X \leqslant \boldsymbol{w}_X$ resp. $\boldsymbol{v}_X < \boldsymbol{w}_X$) holds for all $X \in T$. By $\boldsymbol{v} < \boldsymbol{w}$ we mean $\boldsymbol{v} \leqslant \boldsymbol{w}$ and $\boldsymbol{v} \neq \boldsymbol{w}$. For a nonnegative matrix M we define the matrix series $M^* := I + M + M^2 + \cdots$. We say M^* is finite if all its entries are finite, i.e., the series converges. To a nonnegative matrix $M \in [0,\infty)^{T\times T}$ we associate a directed graph *graph*(M) whose set of vertices is T and whose edges are (X, Y) whenever $M_{X,Y} > 0$. A nonnegative matrix M is *irreducible* if *graph*(M) is strongly connected. A matrix $M' \in \mathbb{R}^{T'\times T'}$ is a *principal submatrix* of $M \in \mathbb{R}^{T\times T}$ if $T' \subseteq T$ and $M'_{X,Y} = M_{X,Y}$ for all $X, Y \in T'$.

We start the proof of Theorem 2.1 by recalling three facts about nonnegative matrices. The first two are standard results, see e.g. [10], while the third follows from [10, Corollary 2.1.6].

Lemma 2.1. *Let $M \in [0,\infty)^{T\times T}$ be a nonnegative matrix.*

(a) *M^* is finite if and only if $\rho(M) < 1$.*
(b) *If M^* is finite, then $M^* = (I - M)^{-1}$.*
(c) *For all principal submatrices M' of M we have $\rho(M') \leqslant \rho(M)$. Furthermore, M has an irreducible principal submatrix M' with $\rho(M') = \rho(M)$.*

Consider the partition $T_1,\dots,T_N$ of T in strongly connected components of *graph*(M), and the corresponding principal submatrices $M^{(1)},\dots,M^{(N)}$ of M. By Lemma 2.1(c) we have $\rho(M^{(i)}) \leqslant \rho(M)$ for every $1 \leqslant i \leqslant N$. Moreover, since every irreducible principal submatrix of M is also a principal submatrix of $M^{(i)}$ for some $1 \leqslant i \leqslant N$, we also have $\rho(M) = \rho(M^{(j)})$ for some $1 \leqslant j \leqslant N$. Therefore,

$$\rho(M) = \max_{1\leqslant i\leqslant N} \rho\big(M^{(i)}\big). \tag{1}$$

Thanks to this equation, we can focus on irreducible nonnegative matrices. We recall four further facts from Perron–Frobenius theory, see [10, Chapter 2]:

Lemma 2.2. *Let $M \in [0,\infty)^{T\times T}$ be nonnegative and irreducible.*

(a) *$\rho(M)$ is a* simple *eigenvalue of M.*
(b) *There exists an eigenvector $\boldsymbol{v} \succ \mathbf{0}$ with $\rho(M)$ as eigenvalue.*
(c) *Every eigenvector $\boldsymbol{v} \succ \mathbf{0}$ has $\rho(M)$ as eigenvalue.*
(d) *For all $\alpha, \beta \in \mathbb{R} \setminus \{0\}$ and $\boldsymbol{v} > \mathbf{0}$: if $\alpha\boldsymbol{v} < M\boldsymbol{v} < \beta\boldsymbol{v}$, then $\alpha < \rho(M) < \beta$.*

Using these facts we prove:

Proposition 2.1. *Let $M \in [0,\infty)^{T\times T}$ be nonnegative and irreducible.*

(a) *Assume that there is $\boldsymbol{v} \in \mathbb{R}^T \setminus \{\mathbf{0}\}$ such that $(I - M)\boldsymbol{v} = \mathbf{0}$. If $\boldsymbol{v} \succ \mathbf{0}$ or $\boldsymbol{v} \prec \mathbf{0}$, then $\rho(M) = 1$; otherwise, $\rho(M) > 1$.*
(b) *Assume that $\boldsymbol{v} = \mathbf{0}$ is the only solution of $(I - M)\boldsymbol{v} = \mathbf{0}$; i.e., there exists a unique $\boldsymbol{x} \in \mathbb{R}^T$ such that $(I - M)\boldsymbol{x} = \mathbf{1}$. If $\boldsymbol{x} \geqslant \mathbf{1}$, then $\rho(M) < 1$; otherwise, $\rho(M) > 1$.*

Proof.

(a) From $(I - M)\boldsymbol{v} = \mathbf{0}$ it follows $M\boldsymbol{v} = \boldsymbol{v}$. So $\boldsymbol{v}$ is an eigenvector of M with eigenvalue 1, thus $\rho(M) \geqslant 1$.
  - Let $\boldsymbol{v} \succ \mathbf{0}$ or $\boldsymbol{v} \prec \mathbf{0}$. By Lemma 2.2(c), $\rho(M)$ is the eigenvalue of $\boldsymbol{v}$, and so $\rho(M) = 1$.
  - Let $\rho(M) \leqslant 1$, i.e., $\rho(M) = 1$. By Lemma 2.2(a) and (b), the eigenspace of the eigenvalue 1 is one-dimensional and contains a vector $\boldsymbol{x} \succ \mathbf{0}$. So $\boldsymbol{v} = \alpha \cdot \boldsymbol{x}$ for some $\alpha \in \mathbb{R} \setminus \{0\}$. Hence $\boldsymbol{v} \succ \mathbf{0}$ or $\boldsymbol{v} \prec \mathbf{0}$.

(b)
  - Let $\boldsymbol{x} \geqslant \mathbf{1}$. Then $M\boldsymbol{x} = \boldsymbol{x} - \mathbf{1} \prec \boldsymbol{x}$, so we have $\rho(M) < 1$ by Lemma 2.2(d).
  - Let $\rho(M) \leqslant 1$. Suppose for a contradiction that $\rho(M) = 1$. Then, by Lemma 2.2(a), the matrix M would have an eigenvector $\boldsymbol{v} \neq \mathbf{0}$ with eigenvalue 1, so $(I - M)\boldsymbol{v} = \mathbf{0}$, contradicting the assumption. So we have, in fact, $\rho(M) < 1$. By Lemma 2.1(a) and (b) this implies $\boldsymbol{x} = (I - M)^{-1}\mathbf{1} = M^*\mathbf{1} \geqslant \mathbf{1}$. □

We obtain as an immediate consequence:

Proposition 2.2. *Let $M \in [0,\infty)^{T\times T}$ be a nonnegative matrix. The following algorithm decides whether $\rho(M) < 1$ or $\rho(M) = 1$ or $\rho(M) > 1$:*

1. *Compute the partition $T_1, \dots, T_N$ of T in strongly connected components of graph(M), and the corresponding principal submatrices $M^{(1)}, \dots, M^{(N)}$ of M.*
2. *For each $M^{(i)}$, solve the system $(I - M^{(i)})\boldsymbol{v} = \boldsymbol{0}$ using Gaussian elimination.*
   - 2.1. *If there is a vector $\boldsymbol{v} \neq \boldsymbol{0}$ such that $(I - M^{(i)})\boldsymbol{v} = \boldsymbol{0}$, conclude $\rho(M^{(i)}) > 1$ or $\rho(M^{(i)}) = 1$ according to Proposition 2.1(a).*
   - 2.2. *If $\boldsymbol{v} = \boldsymbol{0}$ is the only solution of $(I - M^{(i)})\boldsymbol{v} = \boldsymbol{0}$, solve $(I - M^{(i)})\boldsymbol{v} = \boldsymbol{1}$ using Gaussian elimination, and conclude $\rho(M^{(i)}) < 1$ or $\rho(M^{(i)}) > 1$ according to Proposition 2.1(b).*
3. *Use Eq. (1) and the results of step 2 to conclude $\rho(M) < 1$, $\rho(M) = 1$, or $\rho(M) > 1$.*

Since the partition of a graph into strongly connected components can be computed in linear time by means of Tarjan's algorithm [11], the matrices $M^{(1)}, \dots, M^{(N)}$ can be computed in linear time. If the dimensions of these matrices are $n_1, \dots, n_N$, then we have $\sum_{i=1}^{N} n_i = n$, where $n = |T|$ is the dimension of M. Since Gaussian elimination of a rational n_i-dimensional linear equation system can be carried out in strongly polynomial time using $O(n_i^3)$ arithmetic operations (see e.g. [12]), steps 2 and 3 can be carried out using $O(n^3)$ arithmetic operations. This concludes the proof of Theorem 2.1.

Example 2.1. Consider the matrix

$$M = \begin{pmatrix} 0 & 0 & 4/9 & 2/9 \\ 0 & 0 & 0 & 4/5 \\ 2 & 0 & 0 & 0 \\ 0 & 5/4 & 0 & 0 \end{pmatrix}.$$

It has two strongly connected components, $T_A = \{1,3\}$ and $T_B = \{2,4\}$, with irreducible principal submatrices

$$M^{(A)} = \begin{pmatrix} 0 & 4/9 \\ 2 & 0 \end{pmatrix} \qquad M^{(B)} = \begin{pmatrix} 0 & 4/5 \\ 5/4 & 0 \end{pmatrix}.$$

The system $(I - M^{(A)})\boldsymbol{v} = \boldsymbol{0}$ has $\boldsymbol{v} = \boldsymbol{0}$ as only solution. Since the only $\boldsymbol{x}$ satisfying $(I - M^{(A)})\boldsymbol{x} = \boldsymbol{1}$ is $\boldsymbol{x} = (13,27)^\top \geqslant \boldsymbol{1}$, we have $\rho(M^{(A)}) < 1$. The system $(I - M^{(B)})\boldsymbol{v} = \boldsymbol{0}$ has a solution $\boldsymbol{v} = (4,5)^\top \succ \boldsymbol{0}$, and so $\rho(M^{(B)}) = 1$. Since $\rho(M) = \max\{\rho(M^{(A)}), \rho(M^{(B)})\}$, we conclude $\rho(M) = 1$.

3. Application to multi-type branching processes

As mentioned in the introduction, multi-type branching processes model populations in which the individuals of a generation produce a random number of children of different types in the next generation. Formally, a *population* over types $t_1, \dots, t_n$ is an element of $\mathbb{N}^n$; intuitively, $\boldsymbol{c} \in \mathbb{N}^n$ is the population containing $\boldsymbol{c}_i$ individuals of type t_i for each $i \in \{1, \dots, n\}$.

Let $\boldsymbol{z}^{(k)}$ denote the random variable modeling the population of the kth generation of a stochastic process, and let $\boldsymbol{c}^{(i,k,j)}$ denote the offspring (also a population) of the jth individual of type t_i in $\boldsymbol{z}^{(k)}$. Given an initial population $\boldsymbol{z}^{(0)}$, we have

$$\boldsymbol{z}^{(k+1)} = \sum_{i=1}^{n} \sum_{j=1}^{\boldsymbol{z}_i^{(k)}} \boldsymbol{c}^{(i,k,j)} \quad \text{for every } k \geqslant 0,$$

where $\boldsymbol{z}_i^{(k)}$ denotes the ith component of $\boldsymbol{z}^{(k)}$. If the $\boldsymbol{c}^{(i,k,j)}$ are i.i.d. over all $k \geqslant 0$ and $j \geqslant 1$, then the process $\{\boldsymbol{z}^{(k)}\}_{k=0}^{\infty}$ is called a *multi-type branching process* (MBP). In this case, for every vector $\boldsymbol{c} \in \mathbb{N}^n$ there is a fixed probability $p_{i,\boldsymbol{c}}$ that an individual of type t_i produces offspring $\boldsymbol{c}$. An MBP can be explicitly described by enumerating all the probabilities $p_{i,\boldsymbol{c}} > 0$, and implicitly described by giving functions $f_i : \mathbb{N}^n \to [0,1]$ such that $f_i(\boldsymbol{c}) = p_{i,\boldsymbol{c}}$.

A central parameter of branching processes with one single type is the expected number m of children of an individual, given by $m = \sum_{c=0}^{\infty} c p_c$, where p_c is the probability of generating c children. The process is called *subcritical*, *critical*, or *supercritical* if $m < 1$, $m = 1$, or $m > 1$, respectively. These definitions can be extended to the multi-type case. Let $m_{i,j}$ be the expected number of children of type j of an individual of type i, i.e., $m_{i,j} = \sum_{\boldsymbol{c} \in \mathbb{N}^n} \boldsymbol{c}_j p_{i,\boldsymbol{c}}$, and let M be the $n \times n$ matrix given by $M_{i,j} = m_{i,j}$.

Definition 3.1. An MBP is *subcritical*, *critical*, or *supercritical* if $m < 1$, $m = 1$, or $m > 1$, respectively, where $m = \rho(M)$.

Using the algorithm from Proposition 2.2 we can decide in strongly polynomial time if an MBP (given by the rational probabilities $p_{i,\boldsymbol{c}} > 0$) is subcritical, critical, or supercritical.

A fundamental quantity of an MBP and an initial population is the probability of its ultimate extinction. We define the *extinction probability* $q := \lim_{k\to\infty} \Pr(\boldsymbol{z}^{(k)} = \boldsymbol{0})$. The extinction probability is closely related to criticality [1–3]. We sketch the connection. Consider an MBP with initial population $\boldsymbol{z}^{(0)}$ and matrix M. W.l.o.g. we can assume that every type t_j is *reachable*, i.e., that $\Pr(\boldsymbol{z}_j^{(k)} > 0) > 0$ for some $k \geqslant 0$. (Unreachable types t_j can be removed from the description of the process without affecting its behavior.) Now, let us inductively define *mortality*: type t_j is *mortal* if $p_{j,\boldsymbol{0}} > 0$ or if $p_{j,\boldsymbol{c}} > 0$ for some population $\boldsymbol{c} \in \mathbb{N}^n$ containing only individuals of mortal types; otherwise, t_j is *immortal*. We have (see e.g. [3]): if some type is immortal (which can be easily decided in linear time), then $q < 1$; if all types are mortal, then $q = 1$ iff $\rho(M) \leqslant 1$.

Example 3.1. Consider a population of *commoners* and *nobles*, both of which can be *children* or *adults*. Common children (type 1) either die before reaching adulthood, become common adults, or become noble adults through marriage, with probabilities 3/9, 4/9, and 2/9, respectively. Noble children (type 2) either die or become noble adults with probabilities 1/5 and 4/5. Common adults (type 3) give birth to between 0 and 3 common children with probabilities 1/10, 1/10, 5/10, 3/10, respectively, and noble adults (type 4) give birth to between 0 and 3 noble children with probabilities 2/8, 3/8, 2/8, 1/8.

Let us compute some of the entries of M. For example, $m_{3,1}$ is the expected number of commoner children of a common adult, and so $m_{3,1} = 0 \cdot 1/10 + 1 \cdot 1/10 + 2 \cdot 5/10 + 3 \cdot 3/10 = 2$. Similarly, $m_{1,4}$ is the expected number of noble adults "generated" by a common child, and so $m_{1,4} = 1 \cdot 2/9 = 2/9$. In fact, M is the matrix of Example 2.1 with $\rho(M) = 1$.

Since all types are mortal and reachable (assuming the initial population contains commoners and nobles), it follows from the result above that we have $q = 1$, i.e., the population goes ultimately extinct almost surely.

4. Application to stochastic context-free grammars

Recall that a *stochastic context-free grammar* (SCFG) is a tuple $G = (V, \Sigma, R, X_1)$, where $V = \{X_1, \dots, X_n\}$ is a set of *variables* with a distinguished element X_1 called the *axiom*, Σ is a set of *terminals*, and R is a set of *production rules* $X_i \xrightarrow{p} \alpha$, where $\alpha \in (V \cup T)^*$ and $p \in [0, 1]$, such that $\sum_{X_i \xrightarrow{p} \alpha} p = 1$ for every variable X_i. The probability of a derivation of an SCFG is the product of the probabilities of its corresponding sequence of rules. As explained in [3], every SCFG induces an MBP: the types of the MBP are the variables of the SCFG, the initial population consists of an individual of type X_1, and for every $\boldsymbol{c} \in \mathbb{N}^n$ the probability $p_{i,\boldsymbol{c}}$ is defined as the probability that X_i generates in one step a string α with $(\boldsymbol{c}_1, \dots, \boldsymbol{c}_n)$ occurrences of the variables $(X_1, \dots, X_n)$, respectively. Observe that the branching process has no terminals and does not care about the order of variables, only about their multiplicities. Further, in the MBP all variables of a generation are "derived simultaneously", to produce the next generation. However, these differences are irrelevant as far as the generation of a terminal string is concerned, and we have [3]: the probability that an SCFG terminates (i.e., produces a string of terminals) is equal to the extinction probability of the induced MBP.

An SCFG G is called *consistent* if it terminates with probability 1. The algorithm sketched in Section 3 to decide if $q = 1$ can be easily turned into a strongly polynomial algorithm to decide the consistency of G (see [3], Fig. 7): first, remove all variables X_i that are not reachable from X_1, i.e., X_1 cannot generate any string containing at least one occurrence of X_i. If there is some useless variable left (i.e., some variable that cannot generate any string of terminals), then G is not consistent. Otherwise, compute the matrix M of the associated MBP. The grammar G is consistent iff $\rho(M) \leqslant 1$.

5. An example: neutron scattering process

To illustrate the interest of our result, we consider a classical problem of nuclear physics: determining the critical mass or, equivalently, the critical radius of a perfect sphere of plutonium.[3] Roughly speaking, the critical radius is the smallest radius that will cause a nuclear explosion. More precisely, recall that the explosion is produced by a chain reaction: spontaneous fission of an atom liberates neutrons, whose collisions with other atoms induce further fissions, etc. Following Harris [1], we model the ball by an MBP describing the population of atoms fissioning at different distances from the ball's center. Initially there is one free neutron in the ball. A chain reaction occurs if its line of descendants does not go ultimately extinct (physically, this is identical to all atoms in the ball fissioning in a very short time). Since the spontaneous fission rate is high (several hundred atoms per second per cm^3), even a small probability that one fission causes a chain reaction results in an explosion with large probability after a short time. So the critical radius is approximately given by the smallest radius such that $q < 1$.

Let us assume that the radius of the considered sphere is D, and that a neutron born at distance ξ from the center collides with an atom at distance η from the center with probability density $R(\xi, \eta)$. Let further p_k be the probability that a collision generates k neutrons ($k = 0$ means that no fission occurs). Harris uses the values $p_0 = 0.025$, $p_1 = 0.830$, $p_2 = 0.07$, $p_3 = 0.05$, $p_4 = 0.025$, $p_k = 0$ for $k > 4$, and also gives an expression for $R(\xi, \eta)$ (see [1], p. 86).

The probability that a neutron starting at distance ξ collides with an atom at a distance in the interval $[a, b]$ (with $0 \leqslant a \leqslant b \leqslant D$) and generates k neutrons can be expressed as

$$\theta(\xi, a, b, k) := p_k \cdot \int_a^b R(\xi, \eta)\, d\eta.$$

By discretizing the interval $[0, D]$ into n segments we obtain an MBP with n types $t_1, \dots, t_n$. An individual of type t_i represents a neutron whose distance from the center lies in between $(i-1)D/n$ and iD/n. The probabilities $p_{i,\boldsymbol{c}} > 0$ of the MBP are given by

$$p_{i,\boldsymbol{c}} = \begin{cases} \theta(\frac{(i-0.5)D}{n}, \frac{(j-1)D}{n}, \frac{jD}{n}, k), \\ \qquad \boldsymbol{c}_j = k \geqslant 1 \text{ and } \boldsymbol{c}_\ell = 0 \text{ for } \ell \neq j, \\ 1 - (1 - p_0) \cdot \int_0^D R(t\frac{(i-0.5)D}{n}, \eta)\, d\eta, \\ \qquad \boldsymbol{c} = \boldsymbol{0}, \\ 0, \quad \text{otherwise.} \end{cases}$$

Since all types of the MBP are mortal and all types are reachable from all types, checking whether $q = 1$ can be done by deciding whether $\rho(M) \leqslant 1$ for the square matrix M of the MBP as described in Section 3.

We take different discretizations $n = 25, 50, 75, 100, 150$ and combine our algorithm with binary search to determine the critical radius up to an error of 0.001, using the computer algebra system Maple. During the search, the algorithm analyzes MBPs that get closer and closer to being critical. The running times of our algorithm for the last (and most expensive) binary search step that decreases the interval to 0.001 are given in Table 1. We found the critical

[3] We assume room temperature, and so the density of plutonium is known.

Table 1
Runtime in seconds for the last step of the binary search described in the text.

n	25	50	75	100	150
Critical radius	2.9790	2.9809	2.9815	2.9815	2.9815
Precision	±0.0005	±0.0005	±0.0005	±0.0005	±0.0005
Our algorithm	< 1	< 1	< 1	1	4
Exact LP (Maple Simplex)	< 1	6	32	108	588
Exact LP (QSOpt_ex solver)	< 1	< 1	4	14	72

radius to be in the interval $[2.981, 2.982]$ (using the finest discretization $n = 150$). Harris [1] estimates 2.9.

We also measured the time required for analyzing the MBP in the last step of the binary search if we replace our algorithm by linear programming. We compared our algorithm to Maple's exact simplex package as well as the QSOpt_ex tool [13], a standalone exact LP solver. Our approach outperforms both by at least an order of magnitude.

References

[1] T.E. Harris, The Theory of Branching Processes, Springer, Berlin, 1963.
[2] K.B. Athreya, P.E. Ney, Branching Processes, Springer, 1972.
[3] K. Etessami, M. Yannakakis, Recursive Markov chains, stochastic grammars, and monotone systems of nonlinear equations, Journal of the ACM 56 (1) (2009) 1–66.
[4] C.D. Manning, H. Schütze, Foundations of Statistical Natural Language Processing, MIT Press, 1999.
[5] M. Weir, S. Aggarwal, B. de Medeiros, G. Glodek, Password cracking using probabilistic context-free grammars, in: 30th IEEE Symposium on Security and Privacy, 2009, pp. 391–405.
[6] R. Durbin, S. Eddy, A. Krogh, G. Mitchison, Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids, Cambridge University Press, 2003.
[7] M.-J. Nederhof, G. Satta, Computation of distances for regular and context-free probabilistic languages, Theoretical Computer Science 395 (2–3) (2008) 235–254.
[8] K. Etessami, A. Stewart, M. Yannakakis, Polynomial time algorithms for multi-type branching processes and stochastic context-free grammars, in: STOC, ACM, 2012, pp. 579–588.
[9] K. Etessami, A. Stewart, M. Yannakakis, Polynomial time algorithms for branching Markov decision processes and probabilistic min(max) polynomial Bellman equations, in: ICALP'12, in: Lecture Notes in Computer Science, vol. 7391, Springer, 2012, pp. 314–326.
[10] A. Berman, R.J. Plemmons, Nonnegative Matrices in the Mathematical Sciences, SIAM, 1994.
[11] R. Tarjan, Depth-first search and linear graph algorithms, SIAM Journal on Computing 1 (2) (1972) 146–160.
[12] M. Grötschel, L. Lovász, A. Schrijver, Geometric Algorithms and Combinatorial Optimization, Springer, 1993.
[13] QSOpt_ex solver, http://www2.isye.gatech.edu/~wcook/qsopt/ex/.