Restart Strategies in a Continuous Setting

Jan-Hendrik Lorenz¹

Accepted: 27 March 2021 / Published online: 5 May 2021

© The Author(s) 2021

Abstract

Restarting is a technique frequently employed in randomized algorithms. After some number of computation steps, the state of the algorithm is reinitialized with a new, independent random seed. Luby et al. (Inf. Process. Lett. **47**(4), 173–180, 1993) introduced a universal restart strategy. They showed that their strategy is an optimal universal strategy in the worst case. However, the optimality result has only been shown for discrete processes. In this work, it is shown that their result does not translate into a continuous setting. Furthermore, we show that there are no (asymptotically) optimal strategies in a continuous setting. Nevertheless, we obtain an optimal universal strategy on a restricted class of continuous probability distributions. Furthermore, as a side result, we show that the expected value under restarts for the lognormal distribution tends towards 0. Finally, the results are illustrated using simulations.

Keywords Restarts · Universal strategy · Randomized algorithm

1 Introduction

Numerous algorithms use restarts to boost their performance: If the algorithm fails to find a solution, the algorithm's state is reset to its initial state. Restarts are usually governed by a discrete quantity. For example, stochastic local search algorithms restart if the number of local search steps exceeds a certain threshold and backtracking algorithms commence a restart after a certain number of backtracking steps. In settings such as these, the results by Luby et al. [12] can be applied. They studied the case of Las Vegas algorithms, where restarts can be triggered at discrete time steps. They proved that restarts potentially help to improve the expected runtime of this type of algorithms. In their work, they introduced two of the most important restart strategies: The fixed-cutoff strategy and Luby's (universal) strategy. They

---

✉ Jan-Hendrik Lorenz
jan-hendrik.lorenz@uni-ulm.de

¹ Institute of Theoretical Computer Science, Ulm University, 89069 Ulm, Germany

showed that the fixed-cutoff strategy is optimal w.r.t. the expected runtime for some parameter t^* . However, in practice, t^* is rarely known and hard to determine. Luby's strategy solves this problem: In expectation, it is only worse by a logarithmic factor as compared to the optimal fixed-cutoff strategy. Gomes et al. [6] use randomization and restarts to increase the performance of SAT and CSP-solvers by several orders of magnitude. Nowadays, virtually every state-of-the-art SAT solver utilizes restarts [1]. In an empirical study, Huang [7] found that Luby's strategy is the best performing strategy in practice. Luby's strategy is also beneficial in other contexts, such as $(1 + 1)$ evolutionary algorithms [5].

Albeit both the fixed-cutoff strategy and Luby's strategy have only been introduced for discrete quantities, the restart paradigm can also be applied to continuous quantities to minimize the expected value of this quantity. Eliazar et al. [2] examined enzymes searching for a target in a DNA strand. One of the search strategies of the enzyme is arbitrarily relocating after some time. This relocation corresponds to a restart in the algorithmic setting. Evans and Majumdar [3] analyzed general diffusion processes and found that random resets can be beneficial in expectation. Both works study the effect of restarts on the expected value of a continuous quantity: time. Generally, restarts can help to minimize continuous resources such as money, amount of fuel, traveled distances, emitted greenhouse gases, and so forth. In such scenarios, it seems natural to use established strategies from algorithmics such as the fixed-cutoff and Luby's strategy. However, this raises the question of whether these are also good strategies in a continuous setting since their performance guarantees have only been shown for the discrete case.

Since (most) computers are discrete systems working with finite precision, it may seem unreasonable to consider real-valued restart times for computational purposes. Nevertheless, the results presented here are statements about all universal strategies, including those that can be represented by a digital computer. Therefore, when facing a situation involving a computer controlling a real-time system, e.g., transmission times in a network, the optimal restart time may be an irrational value. This kind of scenario is relevant, e.g., for distributed computations.

Our contribution Luby et al. [12] prove the optimality of Luby's strategy by upper and lower bounds in their work. In this work, we show that the upper bounds do not apply in a continuous setting. Furthermore, we show that there is no reasonable upper bound for either Luby's strategy or any other universal strategy. These results show that in a continuous framework, optimal strategies cannot exist.

However, we determine classes of distributions for which optimal universal strategies exist. These classes are defined by the optimal restart times t^* of the fixed-cutoff strategy. We prove that if t^* is not too close to zero, then there are optimal strategies defined on these classes. Furthermore, we improve the known bounds of Luby's strategy and, therefore, implicitly show that it is an optimal strategy on those classes.

Finally, the quality of Luby's strategy is checked by means of simulations. We use Lévy-Smirnov, generalized Pareto, and lognormal distributions. As a spin-off result, we also show that the expected value under restarts of the lognormal distribution approaches zero under certain circumstances.

2 Preliminaries

Let X be a positive, continuous random variable. In the following, F_X denotes the cumulative distribution function, the survival function $1 - F_X(x)$ is denoted by $S_X(x)$, and f_X denotes the density function of X . If it is evident within the context, then the subscripts are omitted.

2.1 Fundamental Restart Strategies

In this section, we formally define the concept of restarts and universal strategies.

Definition 1 [12] Let $\mathcal{A}$ be a randomized algorithm, and x be its input. Let $L = (t_1, t_2, \dots)$ be an arbitrary sequence of numbers with $t_i \in \mathbb{R}_+ \cup \{\infty\}$ and consider the following modified algorithm $\mathcal{A}_L$: Let $\mathcal{A}(x)$ run for t_i time-units in the i -th try. If a solution is found, return it and quit; otherwise, reset $\mathcal{A}(x)$ and proceed with the $(i + 1)$ -st try with new and independent random choices. Then, L is called a **restart strategy**. Let X be a random variable that describes the runtime behavior of $\mathcal{A}$. In the following, X_L describes the runtime behavior of $\mathcal{A}_L$. If L is not dependent on the algorithm $\mathcal{A}$ and its input x , then L is called a **universal strategy**.

Luby et al. [12] introduced two essential restart strategies. The first one is called the fixed-cutoff strategy. The fixed-cutoff strategy consists of always restarting after some time t has passed.

Definition 2 [12] Let X be a random variable which is defined on $\mathbb{R}_+$ and let $L = (t_1, t_2, \dots)$ be a restart strategy with $t_i = t_1$ for all $i \in \mathbb{N}$. Then, L is called a **fixed-cutoff strategy**. Note that t_1 uniquely determines a fixed-cutoff strategy. We often write X_{t_1} to denote a random variable which describes the runtime with fixed-cutoff restarts using L .

Luby et al. [12] showed that the fixed-cutoff strategy is optimal w.r.t. the expected runtime for some restart time. In other words, there are no other restart strategies performing better on average. Their claim is originally for discrete runtime distributions. For continuous distributions, it was first conjectured [14] that the fixed-cutoff strategy (also called sharp restarts) is optimal; later, this was proven by Pal and Reuveni [15].

Theorem 1 [12, 15] *Let X be any random variable. Then there is a $t^* \in \mathbb{R}_+$ such that the fixed-cutoff strategy $(t^*, t^*, \dots)$ is an optimal strategy w.r.t. the expected value under restarts. Furthermore, the expected value $E[X_t]$ of X with fixed-cutoff restarts at $t \in \mathbb{R}_+$ is given by:*

$$E[X_t] = \frac{1 - F(t)}{F(t)}t + E[X \mid X < t]. \quad (1)$$

In the following, $\alpha_X^* = \inf_{t < \infty} E[X_t]$ denotes the expected runtime using the optimal fixed-cutoff strategy. It is worth noting that there are certain universal

principles regarding α_X^* applying for all distributions X . An example is that $\alpha_X^* = (1 - F_X(t^*)) / f_X(t^*)$ holds, where t^* is the optimal restart time [14, 16].

While the fixed-cutoff strategy is theoretically intriguing, it suffers from a significant drawback: Finding the optimal restart time t^* is a non-trivial task. In fact, nearly the complete runtime behavior has to be known a priori to find t^* . Moreover, even if the distribution is known, exactly determining t^* exactly is NP-hard [10]. In case the cumulative distribution function is known, there is an appropriate approximation algorithm [11]. However, the distribution is usually dependent on the problem instance, and the runtime behavior is therefore unknown. Thus, the fixed-cutoff strategy usually is not used in practice. Luby et al. [12] introduced a second strategy, referred to as Luby's strategy. This strategy does not use any knowledge of the distribution.

Definition 3 [12] Let $L = (t_1, t_2, \dots)$ be a universal restart strategy. If t_i is of the form

$$t_i = \begin{cases} 2^{k-1} & \text{if } i = 2^k - 1 \\ t_{i-2^{k-1}+1} & \text{if } 2^{k-1} \leq i < 2^k - 1 \end{cases}, \quad (2)$$

then L is called **Luby's strategy**.

We denote the expected value of a random variable X using Luby's strategy by $E[X_{\text{Luby}}]$. Luby et al. showed that the expected runtime of Luby's strategy can be bounded w.r.t. the best fixed-cutoff strategy.

Theorem 2 [12] For all discrete distributions X which are defined on $\mathbb{N}$

$$E[X_{\text{Luby}}] \leq 192\alpha^*(5 + \log_2 \alpha^*) \quad (3)$$

holds.

In their paper, Luby et al. introduced Luby's strategy for discrete runtime distributions. They also showed that their strategy is asymptotically optimal in the worst case. This universal strategy is also highly relevant in practice [7].

2.2 Properties of the Weibull Distribution

Throughout this work, we often use the Weibull distribution in proofs. Here, we define this distribution and show some of its properties.

Definition 4 [13] Let X be a positive real-valued random variable. The random variable X is **Weibull distributed** if and only if its cumulative distribution function F_X is given by

$$F_X(x) = 1 - e^{-(\frac{x}{\lambda})^k}, \quad (4)$$

where $\lambda \in \mathbb{R}_+$ is called scale, and $k \in \mathbb{R}_+$ is called shape.

It is known that for some Weibull distributions, the expected value with restarts approaches zero.

Lemma 1 [9] *Let X be a Weibull distribution with shape $k < 1$ and arbitrary scale $\lambda \in \mathbb{R}_+$. Then,*

$$\lim_{t \rightarrow 0} E[X_t] = 0 \quad (5)$$

holds. Here, $E[X_t]$ is the expected value with fixed-cutoff restarts after t time units.

Furthermore, we show that the expected values with restarts are strictly monotone for Weibull distributions.

Lemma 2 *Let X be a Weibull distributed random variable with arbitrary scale $\lambda > 0$. Then, the expected value $E[X_t]$ with restarts at t is strictly monotonically increasing for $t > 0$ if and only if the shape parameter k is less than one. If the shape parameter k is greater than one, then $E[X_t]$ is strictly monotonically decreasing.*

Proof The expected value $E[X_t]$ with restarts after t time units is implicitly given in [9].

$$E[X_t] = \frac{1}{1 - \exp(-(\lambda t)^k)} \left(t \exp(-(\lambda t)^k) + \frac{1}{\lambda} \gamma\left(1 + \frac{1}{k}, (\lambda t)^k\right) \right), \quad (6)$$

where $\gamma(s, x) = \int_0^x t^{s-1} e^{-t} dt$ is the lower incomplete gamma function. To determine the derivative of $E[X_t]$, the derivative of the lower incomplete gamma function $\gamma(1 + \frac{1}{k}, (\lambda t)^k)$ is considered. Using the definition of γ and the chain rule yields:

$$\frac{d}{dt} \gamma\left(1 + \frac{1}{k}, (\lambda t)^k\right) = k \cdot \lambda \cdot \exp(-(\lambda t)^k) \cdot \left((\lambda t)^k\right)^{\frac{1}{k}} \cdot (\lambda t)^{k-1} \quad (7)$$

$$= k \cdot \lambda \cdot \exp(-(\lambda t)^k) \cdot (\lambda t)^k \quad (8)$$

We now calculate the derivative of $E[X_t]$ w.r.t. t and denote it by $g(t)$. The derivative $g(t)$ is then given by:

$$\frac{\exp(-(\lambda t)^k)}{1 - \exp(-(\lambda t)^k)} \left(1 - \frac{\exp(-(\lambda t)^k) k (\lambda t)^k + k (\lambda t)^{k-1} \gamma\left(1 + \frac{1}{k}, (\lambda t)^k\right)}{1 - \exp(-(\lambda t)^k)} \right).$$

We show that $g(t) > 0$ for all $t > 0$. This inequality can be transformed to:

$$\frac{1 - \exp(-(\lambda t)^k)}{k} - (\lambda t)^k \exp(-(\lambda t)^k) - (\lambda t)^{k-1} \gamma\left(1 + \frac{1}{k}, (\lambda t)^k\right) > 0. \quad (9)$$

First, this inequality is observed for the limit $t \rightarrow 0$. The limit of the term $-(\lambda t)^{k-1} \gamma(1 + \frac{1}{k}, (\lambda t)^k)$ has to be examined closely for $k < 1$. The limit of all other terms is clearly zero. Due to L'Hospital's rule, we have:

$$\lim_{t \rightarrow 0} -\frac{\gamma(1 + \frac{1}{k}, (\lambda t)^k)}{(\lambda t)^{1-k}} = \lim_{t \rightarrow 0} -\frac{\exp(-(\lambda t)^k) k (\lambda t)^{2k}}{1 - k} = 0.$$

Thus, the left side of inequality (9) approaches zero in the limit. We now show that the derivative of the left side of inequality (9) is positive for all $t > 0$. This inequality can be transformed to

$$\frac{1 - \exp(-(\lambda t)^k)}{k (\lambda t)^{k-1}} - \exp(-(\lambda t)^k) \lambda t - \gamma\left(1 + \frac{1}{k}, (\lambda t)^k\right) > 0. \quad (10)$$

The derivative $h(t)$ of the left side of inequality (10) is given by:

$$h(t) = \frac{(1 - \exp(-(\lambda t)^k))(1 - k)\lambda}{k(\lambda t)^k}$$

It can be easily verified that for $t > 0$, $h(t)$ evaluates to:

$$h(t) \begin{cases} > 0, & \text{for } k < 1, \\ = 0, & \text{for } k = 1, \\ < 0, & \text{for } k > 1. \end{cases} \quad (11)$$

Thus, for $k < 1$ we have $h(t) > 0$ which implies $g(t) > 0$ for all t . This property means that $E[X_t]$ is strictly monotonically increasing. On the other hand, $k > 1$ implies that $E[X_t]$ is strictly monotonically decreasing. $\square$

3 Continuous Universal Strategies

Theorem 2 presents a useful performance guarantee for Luby's strategy w.r.t. the optimal fixed-cutoff strategy. However, the bounds are obtained for discrete runtime distributions. This raises the question of whether the performance guarantee w.r.t. the optimal fixed-cutoff strategy also applies to continuous runtime distributions.

3.1 Nonexistence Results

Here, we prove that, generally, there are no optimal universal strategies in a continuous setting. First, we show a useful property of the fixed-cutoff strategy: The expected value of any strategy can be expressed as a convex combination of expected values of fixed-cutoff strategies.

Lemma 3 *Let X be any random variable and let $L = (t_1, t_2, \dots)$ be an arbitrary restart strategy. Let $E[X_{t_i}]$ denote the expected value of X with the fixed-cutoff strategy with restart time t_i . Then, $E[X_L]$ can be expressed by*

$$E[X_L] = \sum_{i=1}^{\infty} a_i E[X_{t_i}], \quad (12)$$

where $(a_1, a_2, \dots)$ is a sequence with $a_i \in \mathbb{R}_+$ and $\sum_{i=1}^{\infty} a_i = 1$.

Proof Luby et al. [12] showed this property for discrete distributions. We prove the continuous case.

Let $L = (t_1, t_2, \dots)$ be an arbitrary restart strategy. Define the cumulative restart times $T_i := \sum_{k=1}^i t_k$ and set $T_0 := 0$ by convention. Let $Y := X_L$ be a random variable describing the process of restarting X with strategy L . The functions f_Y and F_Y denote the pdf and, respectively, the cumulative distribution function of Y . Furthermore, $S_Y(x)$ is used as abbreviation for $\Pr(Y > x) = 1 - F_Y(x)$. The function

S_Y may be specified with respect to F_X . Consider the probability $S_Y(T_i + x)$ with $x < t_{i+1}$ and $i \in \mathbb{N}$. Only if the first i runs failed and the first x time units of the $(i + 1)$ -th run are not successful, Y is greater than $T_i + x$. By definition, the runs are mutually independent. Therefore, the probability $S_Y(T_i + x)$ is:

$$S_Y(T_i + x) = (1 - F_X(x)) \prod_{k=1}^i (1 - F_X(t_k)). \quad (13)$$

The pdf f_Y of Y can easily be determined using the appropriate derivative:

$$f_Y(T_i + x) = f_X(x) \prod_{k=1}^i (1 - F_X(t_k)) = f_X(x) S_Y(T_i) \quad (14)$$

Next, the following auxiliary identity is shown:

$$\sum_{i=0}^{\infty} T_i S_Y(T_i) F_X(t_{i+1}) = \sum_{i=1}^{\infty} t_i S_Y(T_i). \quad (15)$$

The first term of the left sum can be omitted as its value is zero. Furthermore, it should be noted that $S_Y(T_i) F_X(t_{i+1})$ is the probability $\Pr(T_i < Y \leq T_{i+1})$. The result is then obtained using the following transformations.

$$\sum_{i=0}^{\infty} T_i S_Y(T_i) F_X(t_{i+1}) = \sum_{i=1}^{\infty} T_i S_Y(T_i) F_X(t_{i+1}) \quad (16)$$

$$= \sum_{i=1}^{\infty} \sum_{k=1}^i t_k S_Y(T_i) F_X(t_{i+1}) = \sum_{k=1}^{\infty} t_k \sum_{i=k}^{\infty} S_Y(T_i) F_X(t_{i+1}) \quad (17)$$

$$= \sum_{k=1}^{\infty} t_k \sum_{i=k}^{\infty} \Pr(T_i < Y \leq T_{i+1}) = \sum_{k=1}^{\infty} t_k S_Y(T_k) \quad (18)$$

We now proceed to determine the expected value $E[Y]$:

$$E[Y] = \int_0^{\infty} x f_Y(x) dx = \sum_{i=0}^{\infty} \int_{T_i}^{T_{i+1}} x f_Y(x) dx \quad (19)$$

$$= \sum_{i=0}^{\infty} \int_0^{t_{i+1}} (T_i + x) f_Y(T_i + x) dx \quad (20)$$

$$= \sum_{i=0}^{\infty} T_i S_Y(T_i) \int_0^{t_{i+1}} f_X(x) dx + S_Y(T_i) \int_0^{t_{i+1}} x f_X(x) dx \quad (21)$$

$$= \sum_{i=0}^{\infty} T_i S_Y(T_i) F_X(t_{i+1}) + S_Y(T_i) \int_0^{t_{i+1}} x f_X(x) dx. \quad (22)$$

The left part of (22) can be transformed with (15). We use Theorem 1 and the definition of the conditional expected value to simplify the right part:

$$\int_0^{t_i} x f_X(x) dx = F_X(t_i) E[X_{t_i}] - t_i (1 - F_X(t_i)), \quad (23)$$

where $E[X_{t_i}]$ is the expected value of X utilizing the fixed-cutoff strategy $(t_i, t_i, \dots)$. Using this equation and additional transformations result in

$$E[Y] = \sum_{i=1}^{\infty} t_i S_Y(T_i) + S_Y(T_{i-1}) (F_X(t_i) E[X_{t_i}] - t_i (1 - F_X(t_i))) \quad (24)$$

$$= \sum_{i=1}^{\infty} t_i S_Y(T_i) + S_Y(T_{i-1}) F_X(t_i) E[X_{t_i}] - t_i S_Y(T_i) \quad (25)$$

$$= \sum_{i=1}^{\infty} S_Y(T_{i-1}) F_X(t_i) E[X_{t_i}] = \sum_{i=1}^{\infty} \Pr(T_{i-1} < Y \leq T_i) E[X_{t_i}]. \quad (26)$$

For obtaining (25), (13) is used. The expected value $E[Y]$ can, therefore, be expressed as a combination of the expected values of fixed-cutoff strategies. Evidently $\Pr(T_{i-1} < Y \leq T_i)$ is non-negative, and the sum $\sum_{i=1}^{\infty} \Pr(T_{i-1} < Y \leq T_i)$ equals one. $\square$

We will now proceed and show that there is no universal strategy in a continuous setting that allows for any reasonable bound in contrast to the discrete case.

Theorem 3 *There is no universal strategy $L = (t_1, t_2, \dots)$ such that there is a function $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ with $\lim_{x \rightarrow 0} f(x) < \infty$ and $\exists x \in \mathbb{R}_+ : f(x) < \infty$ with*

$$E[X_L] \leq \alpha^* \cdot f(\alpha^*) \quad (27)$$

for all runtime distributions X .

Proof We consider two cases: Either the restart times of the universal strategy are all equal to zero or there is at least one restart time t_j which is not equal to zero. First, consider the strategy $L_1 = (t_1, t_2, \dots)$ such that there is a $j \in \mathbb{N}$ with $t_j > 0$. Define the cumulative restart times $T_i := \sum_{k=1}^i t_k$ and set $T_0 := 0$ by convention. Let Z be a Weibull distribution with shape $k < 1.0$ and arbitrary scale $\lambda \in \mathbb{R}_+$. According to Lemma 3, $\Pr(T_{j-1} < Z_{L_1} \leq T_j) E[Z_{t_j}]$ bounds the expectation of the universal strategy $E[Z_{L_1}]$ from below.

Furthermore, the expected value $E[Z_{t_j}]$ is greater than zero due to Lemma 1 and Lemma 2. In addition, $\Pr(T_{j-1} < Z_{L_1} \leq T_j)$ is also strictly greater than zero. On the other hand, the minimal expected value α^* is zero, as stated in Lemma 1. Let $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ be any function with $\lim_{x \rightarrow 0} f(x) < \infty$. Due to this property, $\alpha^* \cdot f(\alpha^*) = 0$ must be valid. Hence, we conclude that

$$E[X_{L_1}] \geq \Pr(T_{j-1} < Z_{L_1} \leq T_j) \cdot E[Z_{t_j}] > 0 = \alpha^* \cdot f(\alpha^*) \quad (28)$$

holds. Put differently, Inequality (27) is not valid.

In the second case, the universal strategy $L_2 = (0, 0, \dots)$ is considered. However, it is easy to construct a distribution such that the expected value with restarts using L_2 diverges to infinity. To this end, we consider the limiting value of the expected value when the restart times approach 0. First, we use (1) for the expected value and modify slightly its terms. The following derivation is well-known, compare, for example, Wolter [16].

$$\lim_{t \rightarrow 0} E[X_t] = \lim_{t \rightarrow 0} \frac{1 - F(t)}{F(t)} t + E[X \mid X < t] \quad (29)$$

$$= \lim_{t \rightarrow 0} \frac{t}{F(t)} - t + \frac{\int_0^t x \cdot f(x) dx}{F(t)} = \lim_{t \rightarrow 0} \frac{1}{f(t)} \quad (30)$$

In (30), L'Hospital's rule is used. We now define a new Weibull distributed random variable U with shape $k > 1.0$. The scale parameter λ is discussed later. As can be easily checked by derivation of (4), the pdf f_U is given by

$$f_U(x) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} \exp\left[-\left(\frac{x}{\lambda}\right)^k\right], \quad (31)$$

for $x \geq 0$. And since $k > 1$, it follows that $f(0) = 0$ is valid. Then (30) states that the expected value $E[U_{L_2}]$ diverges towards infinity. On the other hand, restarts have a negative effect on the expected value of U , according to Wolter [16]. In this case, this indicates $\alpha^* = E[U]$. The expected value $E[U]$ is given by $\lambda \cdot \Gamma(1 + \frac{1}{k})$, where $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ is the gamma function [13].

By assumption there is a $x \in \mathbb{R}_+$ with $f(x) < \infty$. Therefore, choose λ such that $\lambda \cdot \Gamma(1 + \frac{1}{k}) = x$ holds. Then the following observation is valid

$$\alpha^* \cdot f(\alpha^*) = x \cdot f(x) < \infty. \quad (32)$$

As discussed above, $E[U_{L_2}]$ approaches infinity; it follows that Inequality (27) does not hold in this case as well. $\square$

Theorem 3 can be applied to Luby's strategy. This is because the logarithm approaches minus infinity as its argument approaches zero. This means that the bound in Theorem 2 does not hold in a continuous setting. Of course, this statement also applies to all other bounds, according to Theorem 3. Given that Luby's strategy is of great importance in practice, we present this finding as a corollary.

Corollary 1 *There is no function $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ with $\lim_{x \rightarrow 0} f(x) < \infty$ such that*

$$E[X_{\text{Luby}}] \leq \alpha^* \cdot f(\alpha^*) \quad (33)$$

for all runtime distributions X .

Proof This statement immediately follows from Theorem 3. $\square$

Since computers are inherently discrete systems, not all real-valued restart times and, accordingly, not all restart strategies can be represented on a computer. Since computers are inherently discrete systems, not all real-valued restart times and, accordingly, not all restart strategies can be represented on a computer. Then, of

course, a bound of the form given in Theorem 3 is too strong for the strategies that can actually be represented. This, however, raises the question of whether a suitable bound can be found by introducing an error term depending on the machine precision. As shall be seen, the answer to this question is also negative.

To formalize this result, we use $\mathcal{R} \subset \mathbb{R}$ to denote a subset of the real numbers that are representable by a computer. For example, $\mathcal{R}$ could be floating-point numbers. However, we will not impose any further requirements on $\mathcal{R}$, so $\mathcal{R}$ can be defined arbitrarily. Instead, we use a constant $c_{\mathcal{R}} \in \mathbb{R}_+$, which is supposed to describe a precision error with respect to $\mathcal{R}$. We call a restart strategy $(t_1, t_2, \dots)$ representable with respect to $\mathcal{R}$ if $t_i \in \mathcal{R}$ holds for all $i \in \mathbb{N}$.

Corollary 2 *Let $\mathcal{R} \subset \mathbb{R}$ be a set of representable numbers and let $c_{\mathcal{R}} \in \mathbb{R}_+$ be an associated constant. There is no universal strategy $L = (t_1, t_2, \dots)$ representable with respect to $\mathcal{R}$ such that there is a function $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ with $\lim_{x \rightarrow 0} f(x) < \infty$ and $\exists x \in \mathbb{R}_+ : f(x) < \infty$ with*

$$E[X_L] \leq \alpha^* \cdot f(\alpha^*) + c_{\mathcal{R}} \quad (34)$$

for all runtime distributions X .

Proof Once more, we have to distinguish whether the restart strategy L is constantly zero or not. In case L is constantly zero, the corresponding arguments from the proof of Theorem 3 can be applied. We, therefore, focus on the other situation.

Let X be a Weibull distributed random variable with shape parameter $k < 1$. In the following, we assume that $x \in \mathcal{R}$ is the smallest positive representable number. Since L is a representable strategy, $x \leq t, \forall t \in L$ holds. As a consequence of Lemma 2 and 3, the following statement holds true:

$$E[X_L] \geq E[X_x]. \quad (35)$$

The expected value $E[X_x]$ can be bounded from below by applying Theorem 1:

$$E[X_x] = \frac{1 - F(x)}{F(x)}x + E[X \mid X < x] \geq \frac{1 - F(x)}{F(x)}x. \quad (36)$$

We now use the cdf of X , which is given in Definition 4:

$$\frac{1 - F(x)}{F(x)}x = \frac{\exp\left(-\left(\frac{x}{\lambda}\right)^k\right)}{1 - \exp\left(-\left(\frac{x}{\lambda}\right)^k\right)}x. \quad (37)$$

If the scale parameter λ approaches infinity, the following limiting value is obtained:

$$\lim_{\lambda \rightarrow \infty} \exp\left(-\left(\frac{x}{\lambda}\right)^k\right) = 1. \quad (38)$$

This means that (37) approaches infinity as λ approaches infinity. Therefore, there is a λ , such that $\frac{1 - F(x)}{F(x)}x > c_{\mathcal{R}}$ holds. Finally, according to Lemma 1, $\alpha^* = 0$ for every

Weibull distribution with $k < 1$, so $\alpha^* \cdot f(\alpha^*) + c_{\mathcal{R}} = c_{\mathcal{R}}$. By bringing all these findings together, we obtain:

$$E[X_L] \geq E[X_x] \geq \frac{1 - F(x)}{F(x)} x > \alpha^* \cdot f(\alpha^*) + c_{\mathcal{R}} = c_{\mathcal{R}}. \quad (39)$$

This completes the proof. $\square$

3.2 Existence Results

In this part, we show that there are optimal universal strategies if a more restricted class of distributions is considered. The proofs so far used the fact that if the runtime distribution is Weibull distributed, then both the optimal restart time and the expected runtime approach zero. Generally, the optimal expected value can be bounded from above. Our rationale is based on the reasoning presented in [8] and is adjusted to the continuous case. Primarily, we provide an enhanced analysis of the expected value in Theorem 4 and thus derive, to the best of our knowledge, the currently best known upper bound for the expected value when applying Luby's strategy.

In the following, β_X^* denotes $\min_{t \in \mathbb{R}_+} \frac{t}{F(t)}$ and $k^* \in \mathbb{R}_+$ is a real number with $\frac{k^*}{F(k^*)} = \beta_X^*$.

Lemma 4 [8, 12] *The inequality*

$$\beta_X^* \leq 4\alpha_X^* \quad (40)$$

holds for all distributions X .

Proof Luby et al. [12] state this property for discrete probability distributions. However, the proof is omitted in their work. A slightly modified version of the proof presented in [8] shows this property for continuous distributions.

Let $t \in \mathbb{R}_+$ be an optimal restart time, i.e., $E[X_t] = \alpha^*$. Two cases are distinguished: First, it is assumed that $E[X_t] < \frac{t}{2F(t)}$ holds. In this case, $\Pr(X \leq 2E[X_t]F(t)) \geq 0.5$ follows. An auxiliary random variable $Y := \min(X, t)$ is defined to show this. It can be seen that the expected values $E[Y]$ and $E[X_t]$ are related:

$$E[Y] = F(t) \cdot E[X | X < t] + (1 - F(t)) \cdot t = F(t) \cdot E[X_t]. \quad (41)$$

We analyze the probability $\Pr(X > 2E[X_t]F(t))$ in more detail. The following inequality can be derived by applying Markov's inequality and the assumption $2E[X_t]F(t) < t$:

$$\begin{aligned} \Pr(X > 2E[X_t]F(t)) &= \Pr(Y > 2E[X_t]F(t)) \\ &= \Pr(Y > 2E[X_t]F(t)) = \Pr(Y > 2E[Y]) \leq \frac{1}{2}. \end{aligned} \quad (42)$$

From this inequality, $\Pr(X \leq 2E[X_t]F(t)) \geq 1/2$ also follows. And thus we can estimate β^* as follows:

$$\beta^* = \min_x \frac{x}{F(x)} \leq \frac{2E[X_t]F(t)}{\Pr(X \leq 2E[X_t]F(t))} \leq 4E[X_t]F(t) \leq 4E[X_t] = 4\alpha^*.$$

Secondly, we now assume that $2E[X_t]F(t) > t$ is valid. With this assumption, α^* can be estimated as follows:

$$\alpha^* = E[X_t] > \frac{t}{2F(t)} \geq \frac{1}{2} \min_x \frac{x}{F(x)} = \frac{\beta^*}{2}.$$

In conclusion, $\beta^* \leq 4\alpha^*$ is true for all distributions. $\square$

This Lemma can be used to show that distributions from a more restricted class admit an upper bound on the expected value of Luby's strategy. That is, only runtime distributions are considered such that the optimal restart time does not approach zero.

Theorem 4 *Let X be a positive real-valued distribution such that there is a k^* with $\frac{k^*}{F(k^*)} = \beta^*$ and $k^* > 2^{-d-1}$ for some $d \in \mathbb{N}_0$. Then,*

$$E[X_{\text{Luby}}] \leq 2^{d+7} \alpha^* (d + 4 + \log_2 \alpha^*), \quad (44)$$

holds, where α^* is the optimal expected runtime under restart.

Proof The presentation of the proof follows [8].

Let $Y := X_{\text{Luby}}$ denote a random variable describing the process of restarting X with Luby's strategy. The number $z := 2^{\lceil \log_2 k^* \rceil + d}$ is the smallest element of Luby's strategy guaranteed to be at least as large as k^* . The elements of Luby's strategy are those natural numbers being powers of two. This is because all elements in Luby's strategy are natural numbers, which are powers of two. Furthermore, $k^* > 2^{-d-1}$ is valid by assumption, so $z \geq k^*$ and $z \in \mathbb{N}$ is ensured. Two auxiliary values $j_0 := \lceil \log_2 k^* \rceil + d$ and $n_0 := \lceil -\log_2 F(k^*) \rceil$ are used in the following.

First, the case $n_0 = 0$ is examined. This case implies $F(k^*) = 1$. In addition, let $m \in \mathbb{N}$ and $j \in \mathbb{N}_0$ be two natural numbers with $j < m$. As can be verified using Definition 3 there are at least 2^{m-1-j} runs with length of 2^j if $Y > m2^{m-1}$. Accordingly, for $Y > (j_0+1)2^{j_0}$ there is at least one run of length 2^{j_0} . Since $F(k^*) = 1$ is valid, $\Pr(Y > (j_0+1)2^{j_0}) = 0$ ensues. This allows the expected value $E[Y]$ to be bounded from above:

$$E[Y] \leq (\lceil \log_2 k^* \rceil + d + 1) 2^{\lceil \log_2 k^* \rceil + d} \leq k^* (\log_2 k^* + d + 2) 2^{d+1} \quad (45)$$

$$= \beta^* (\log_2 \beta^* + d + 2) 2^{d+1} \leq \alpha^* (\log_2 \alpha^* + d + 4) 2^{d+3} \quad (46)$$

In (46), we initially apply $\beta^* = k^*/F(k^*)$ and $F(k^*) = 1$, then we use Lemma 4. The last inequality conforms to the statement of the theorem.

We may now assume $n_0 \geq 1$. Furthermore, $j_0 \geq 0$ holds due to the definition of j_0 . Since, by definition, $j_0 \geq 0$, we may also use $j_0 + n_0 \geq 1$. Above, we have already indicated that if $Y > m2^{m-1}$, then there are at least 2^{m-1-j} runs with a length of 2^j . This fact can be used to bound the probability $\Pr(Y > m2^{m-1})$ from above:

$$\Pr(Y > m2^{m-1}) \leq \Pr(X > 2^j)^{2^{m-1-j}}. \quad (47)$$

The tail probability of Y is estimated by means of Inequality (47). The following holds for $\ell \in \mathbb{N}_0$:

$$\Pr(Y > (j_0 + n_0 + \ell + 1)2^{j_0+n_0+\ell}) \leq \Pr(X > 2^{j_0})^{2^{n_0+\ell}} \quad (48)$$

$$\leq (1 - F(k^*))^{\frac{2^\ell}{F(k^*)}} \leq \exp(-2^\ell). \quad (49)$$

Let $d(\ell) := (j_0 + n_0 + \ell + 1)2^{j_0+n_0+\ell}$ for $\ell \geq 0$ and $d(-1) := 0$. The expected value $E[Y]$ can be bounded from above by using Inequality (49).

$$\begin{aligned} E[Y] &= \sum_{\ell=0}^{\infty} \int_{d(\ell-1)}^{d(\ell)} (1 - F_Y(x)) dx \\ &\leq (j_0 + n_0 + 1)2^{j_0+n_0} + \sum_{\ell=0}^{\infty} (j_0 + n_0 + \ell + 3)2^{j_0+n_0+\ell} \exp(-2^\ell) \\ &\leq 2(j_0 + n_0)2^{j_0+n_0} + \sum_{\ell=0}^{\infty} (j_0 + n_0)(\ell + 4)2^{j_0+n_0+\ell} \exp(-2^\ell) \quad (50) \end{aligned}$$

$$\leq (j_0 + n_0)2^{j_0+n_0} \left(2 + \sum_{\ell=0}^{\infty} (\ell + 4) \left(\frac{2}{\exp(2)} \right)^\ell \right) \quad (51)$$

In Inequality (50), the assumption $j_0 + n_0 \geq 1$ is implicitly utilized. Next, we examine the series $\sum_{\ell=0}^{\infty} (\ell + 4) \left(\frac{2}{\exp(2)} \right)^\ell$.

$$\sum_{\ell=0}^{\infty} (\ell + 4) \left(\frac{2}{\exp(2)} \right)^\ell = \sum_{\ell=1}^{\infty} \ell \left(\frac{2}{\exp(2)} \right)^{\ell-1} + 3 \sum_{\ell=0}^{\infty} \left(\frac{2}{\exp(2)} \right)^\ell \quad (52)$$

$$= \left(\frac{e^2}{e^2 - 2} \right)^2 + 3 \frac{e^2}{e^2 - 2} \leq 6 \quad (53)$$

The estimation of the expected value $E[Y]$ is simplified by inserting Inequality (53) into Inequality (51).

$$E[Y] \leq 8(j_0 + n_0)2^{j_0+n_0}. \quad (54)$$

Using the definition of j_0 and n_0 results in:

$$2^{j_0+n_0} \leq 2^{2+d+\log_2 k^* - \log_2 F(k^*)} = 2^{d+2} \frac{k^*}{F(k^*)} = 2^{d+2} \beta^*. \quad (55)$$

In the final step, we insert Inequality (55) into Inequality (54) and apply Lemma 4.

$$E[Y] \leq 2^{d+5} \beta^* (d + 2 + \log_2 \beta^*) \leq 2^{d+7} \alpha^* (d + 4 + \log_2 \alpha^*). \quad (56)$$

Inequality (56) describes the desired property, and therefore the proof is complete. $\square$

Note that for $d = 0$, the upper bound of the expected runtime is $128\alpha^*(4 + \log_2 \alpha^*)$, which is, to the best of our knowledge, the best known bound for Luby's

strategy. This proof can also be transferred to discrete distributions, and thus this bound can be achieved in these distributions as well.

Furthermore, Theorem 7 in [12] bounds the performance of any universal strategy from below. While the proof is for discrete probability distributions, the arguments can be directly transferred to the case discussed here. This implies that Luby's strategy is asymptotically optimal on the class of distributions examined here.

Theorem 4 limits the minimal restart time to 2^{-d-1} . However, for some distributions, the optimal restart time approaches zero, and the reasoning from Theorem 4 cannot be applied. Theorem 3 states that for these distributions, there is no guarantee of the form $E[X_L] \leq \alpha_X^* \cdot g(\alpha_X^*)$ with $\lim_{t \rightarrow 0} g(t) < \infty$. The only remaining option is to provide a runtime guarantee for universal strategies with respect to the optimal fixed-cutoff strategy is using a bound function g with $\lim_{t \rightarrow 0} g(t) = \infty$. It is known that the expected runtime $E[X_x]$ approaches $1/f(x)$ as x approaches zero, where f is the density function. One of the critical distributions is the Weibull distribution, which has the property $f(x) \sim x^{k-1}$ for $x \rightarrow 0$ and $k \in (0, 1)$. Thus, every possible bound function $g(x)$ has to approach infinity at least as fast as $1/x$ for $x \rightarrow 0$. It is possible, though, that there are distribution types for which their density function approaches infinity faster than the Weibull distribution in the neighborhood of zero.

4 Simulations

So far, the Weibull distribution has been used to show the negative result. In this section, we take other distributions into account. More specifically, the quality of Luby's strategy in terms of the optimal expected value under restarts is examined. For this purpose, we consider the Lévy-Smirnov, the generalized Pareto, and the lognormal distribution. These distributions cover three possibilities: The Lévy-Smirnov distribution considered here will be shown to fulfill the conditions of Theorem 4. For the generalized Pareto distribution, it turns out that Theorem 4 is not applicable, but Luby's strategy often yields good results. Finally, in the case of the lognormal distribution, Theorem 4 can be applied to any given parameter combination. In the limit, however, Theorem 4 is not applicable. Consequently, the results of Luby's strategy become increasingly worse as the parameters approach the limit.

4.1 Setup of the Simulations

The following simulations are conducted as follows. For each random variable X , we use 1000 samples from X_{Luby} . In other words, we apply Luby's strategy 1000 times to X . Using the collected data, we calculate the sample mean $\bar{x}$ as an estimate for $E[X_{\text{Luby}}]$. Additionally, we provide the 95% confidence interval in each case. The results are graphically illustrated and compared with α^* . The data and the source code used to obtain the results are freely available online at https://github.com/jahehalo/restart_universal_strategies.

4.2 The Lévy-Smirnov Distribution

We start by defining the Lévy-Smirnov distribution.

Definition 5 [4] Let X be a positive, real-valued random variable. The random variable X is **Lévy-Smirnov distributed** if and only if its cumulative distribution function F_X is given by

$$F_X(x) = \operatorname{erfc}\left(\sqrt{\frac{c}{2x}}\right), \quad (57)$$

where $c \in \mathbb{R}_+$ is called scale, and erfc is the complementary error function with $\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-u^2} du$.

To examine the effect of the scale parameter c , we will test different values of c . The values of c range between 10^{-3} and 10^2 . In order to estimate the quality of the Luby strategy, we first determine α^* . For this purpose, we use the methodology presented in [9]. This method involves finding the optimal restart quantile $p^* \in [0, 1]$ by solving the following equation:

$$(p-1) \cdot Q(p) + (1-p) \cdot p \cdot Q'(p) - \Omega(p) + \Omega(0) = 0, \quad (58)$$

where Q is the quantile function of X , Q' is the derivative of Q and Ω is the antiderivative of Q . In the case of the Lévy-Smirnov distribution, these three functions are given by:

$$Q(p) = \frac{c}{2 \cdot \operatorname{erfc}^{-1}(p)^2}, \quad (59)$$

$$Q'(p) = c \cdot \frac{\sqrt{\pi} \cdot \exp\left[\operatorname{erfc}^{-1}(p)^2\right]}{2 \cdot \operatorname{erfc}^{-1}(p)^3}, \quad (60)$$

$$\Omega(p) = c \cdot \left(1 - p + \frac{\exp\left[-\operatorname{erfc}^{-1}(p)^2\right]}{\sqrt{\pi} \cdot \operatorname{erfc}^{-1}(p)}\right). \quad (61)$$

Here, erfc^{-1} is the inverse of the complementary error function erfc . As shown in [9], the optimal restart quantile p^* is invariant to the choice of the scale parameter c . Thus, we fix c to 1 and solve (58) numerically.

This approach leads to $p^* \approx 0.2963$. Using p^* the optimal restart time $t^* = Q(p^*)$ can be determined. This way, the optimal restart time $t^* \approx 0.9170 \cdot c$ is obtained, resulting in $\alpha^* \approx 2.6718 \cdot c$.

We also examine whether the Lévy-Smirnov distribution satisfies the conditions of Theorem 4. Therefore k^* has to be determined. For this purpose we have to minimize $t/F(t)$. By substituting $t = Q(p)$, we can alternatively consider $Q(p)/p$, because Q is the quantile function and therefore $F(Q(p)) = p$. Numerical calculations show that $Q(p)/p$ assumes its minimum at $p' \approx 0.23381$. Resubstituting yields $k^* = Q(p') \approx 0.7055 \cdot c$. Thus, the conditions for Theorem 4 are met for each $c > 0$.

Therefore we can safely assume, that $E[X_{\text{Luby}}]$ performs reasonably well if the value of c is not too small.

The results of the simulation are shown in Fig. 1. The blue line shows α^* while the orange dots represent the sample mean $\bar{x}$ of Luby's strategy. The particularly good performance of the Luby strategy at $c = 1$ is notable here. In this case, the sample mean is $\bar{x} \approx 2.7835$ with a 95% confidence interval of (2.6043, 2.9626). As can be seen on the log-log plot, Luby's strategy is also well suited for the other values of the scale parameter. Partly based on these results, we suspect that for many distributions, a much tighter bound than the one given in Theorem 4 can be achieved.

4.3 The Generalized Pareto Distribution

Next, we will focus on the generalized Pareto distribution.

Definition 6 [13] Let X be a positive, real-valued random variable. The random variable X is **generalized Pareto (GP) distributed** if and only if its cumulative distribution function F_X is given by

$$F_X(x) = 1 - \left(1 + \frac{kx}{\sigma}\right)^{-1/k}, \quad (62)$$

where $\sigma \in \mathbb{R}_+$ is called scale, and $k \in \mathbb{R}$ is called shape.

For the generalized Pareto distribution it is known that restarts have a beneficial effect if and only if $k > 0$. Therefore we focus on the case $k > 0$. In this case, it

Fig. 1 This plot represents the expected values under restarts of the Lévy-Smirnov distribution. The scale parameter c is displayed on the x-axis. The resulting expected values are on the y-axis. The blue line corresponds to α^* . The orange dots correspond to the sample mean of Luby's strategy from 1000 runs. In other words, the closer the orange dots are to the blue line, the better is Luby's strategy. The whiskers represent the 95% confidence interval. If the whiskers are not visible, this indicates that the confidence interval is smaller than the dot size

is known [9] that the optimal restart time is zero, and the associated expected value is $\alpha^* = \sigma$. It is also well known that $E[X_t] \leq t/F(t)$ holds for all t . To see this, consider the following:

$$E[X_t] = \frac{1 - F(t)}{F(t)}t + E[X \mid X < t] \leq \frac{1 - F(t)}{F(t)}t + t \leq \frac{t}{F(t)}. \quad (63)$$

Only the fact $E[X \mid X < t] \leq t$ has been used here.

We consider the limit of $t/F(t)$ for $t \rightarrow 0$. According to L'Hospital's rule, the limit approaches $1/f(0)$, where $f(t) = \frac{1}{\sigma}(1 + \frac{k \cdot t}{\sigma})^{-\frac{1}{k}-1}$ and therefore $1/f(0) = \sigma = \alpha^*$. In [9] it was implicitly shown that $E[X_t] > \sigma$ is valid for all $t > 0$. Together with Inequality (63), it follows that $k^* = 0$ and $\beta^* = \sigma$. This makes the GP distribution an interesting case: Since $k^* = 0$, Theorem 4 is not applicable. On the other hand, due to $\alpha^* > 0$, the proof of Theorem 3 is also not directly applicable either. In other words, neither of the two main results apply directly. Instead, we can only rely on numerical simulations to clarify how well suited Luby's strategy is for the GP distribution.

We particularly want to clarify the role of the two parameters σ and k . For this purpose, in Fig. 2, the scale parameter σ is set to 10, and different values of k are tested. The orange dots indicate the sample mean of Luby's strategy, while the blue line indicates α^* . As can be seen from the diagram, Luby's strategy seems to be quite suitable for this kind of parameter combination. As k increases, the expected value of Luby's strategy appears to increase approximately linearly. In this scenario, the difference between α^* and $E[X_{\text{Luby}}]$ is limited. On the other hand, for $k \rightarrow 0$,

Fig. 2 This plot represents the expected values under restarts of the generalized Pareto distribution. The scale parameter σ is fixed to 10, while the shape parameter k is displayed on the x-axis. The resulting expected values are on the y-axis. The blue line is $\alpha^* = \sigma$, which has been derived in [9]. The orange dots correspond to the sample mean of Luby's strategy from 1000 runs. In other words, the closer the orange dots are to the blue line, the better is Luby's strategy. The whiskers represent the 95% confidence interval

it seems that α^* and $E[X_{\text{Luby}}]$ approach the same value. This is consistent with the fact that in the limit for $k \rightarrow 0$, the GP distribution becomes an exponential distribution. For exponential distributions, it is known that restarts do not affect the expected value.

Next, we consider the case where the shape parameter k is fixed, and the scale parameter σ is a variable. Figure 3 shows the case with $k = 5$. As can be seen here, although the absolute difference between α^* and $E[X_{\text{Luby}}]$ grows with increasing σ , the log-log plot shows that they have the same order of magnitude for high values of σ . On the other hand, the expected value of Luby's strategy decreases considerably slower than α^* for $\sigma \rightarrow 0$.

4.4 The Lognormal Distribution

Finally, we consider the lognormal distribution.

Definition 7 [4] Let X be a positive, real-valued random variable. The random variable X is **lognormal distributed** if and only if its cumulative distribution function F_X is given by

$$F_X(x) = \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left(-\frac{\ln(x) - \mu}{\sqrt{2} \cdot \sigma} \right), \quad (64)$$

where erf is the error function $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2} du$. The parameter $\sigma \in \mathbb{R}_+$ is called shape, and the parameter $\mu \in \mathbb{R}$ is called scale.

Fig. 3 This plot represents the expected values under restarts of the generalized Pareto distribution. The shape parameter k is fixed to 5, while the scale parameter σ is displayed on the x-axis. The resulting expected values are on the y-axis. The blue line is $\alpha^* = \sigma$, which has been derived in [9]. The orange dots correspond to the sample mean of Luby's strategy from 1000 runs. In other words, the closer the orange dots are to the blue line, the better is Luby's strategy. The whiskers represent the 95% confidence interval. If the whiskers are not visible, this indicates that the confidence interval is smaller than the dot size. The x-axis of the left plot is scaled logarithmically, while in the right plot, both axes have a logarithmic scale

The lognormal distribution has the interesting property that for each parameter combination (μ, σ) , k^* is greater than zero, and thus Theorem 4 is applicable. However, in the limit for $\sigma \rightarrow \infty$ both, k^* and α^* , approach zero. This is, to the best of our knowledge, a previously unknown property.

Theorem 5 *Let $X_{\mu,\sigma}$ denote a lognormal distributed random variable with scale μ and shape σ . Then the following holds for the optimal expected value under restarts:*

$$\lim_{\sigma \rightarrow \infty} \alpha_{X_{\mu,\sigma}}^* = 0. \quad (65)$$

Proof In the following, we set the restart time t to $t = \exp(\mu - \sigma^2)$. It has already been explained in Inequality (63) that $E[X_t] \leq t/F(t)$ is valid. We, therefore, examine the limit of $t/F(t)$ for $\sigma \rightarrow \infty$ more closely. Two technical details should be noted in advance. First, $\lim_{x \rightarrow -\infty} \operatorname{erf}(x) = -1$ and also the derivative of the error function given by $\frac{d}{dx} \operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \exp(-x^2)$. The limiting value can then be determined by applying L'Hospital's rule twice.

$$\lim_{\sigma \rightarrow \infty} \frac{t}{F(t)} = \lim_{\sigma \rightarrow \infty} \frac{2 \exp(\mu - \sigma^2)}{1 + \operatorname{erf}\left(-\frac{\sigma}{\sqrt{2}}\right)} = \lim_{\sigma \rightarrow \infty} \frac{-4\sigma \exp(\mu - \sigma^2)}{-\exp\left(-\frac{\sigma^2}{2}\right)} \sqrt{\frac{\pi}{2}} \quad (66)$$

$$= \lim_{\sigma \rightarrow \infty} \frac{\sqrt{8\pi} \exp(\mu)\sigma}{\exp\left(\frac{\sigma^2}{2}\right)} = \lim_{\sigma \rightarrow \infty} \frac{\sqrt{8\pi} \exp(\mu)}{\sigma \exp\left(\frac{\sigma^2}{2}\right)} = 0 \quad (67)$$

Thus, since $t/F(t)$ tends towards zero, we can conclude that $E[X_t]$ also approaches zero. Due to the fact that $\alpha^* = \inf_x E[X_x] \leq E[X_t]$ holds, the statement follows. $\square$

This not only shows that the optimal expected value α^* approaches zero. As Theorem 5 is shown by using $t/F(t)$ as an upper bound for $E[X_t]$, this also implies that k^* approaches zero. Furthermore, the proof also indicates that α^* is approaching zero super exponentially.

As α^* tends to zero under the discussed circumstances, the reasoning from the proof of Theorem 3 is applicable. It is, therefore, reasonable to assume that the expected values of the optimal restart strategy and Luby's strategy are significantly different for large σ values. This behavior can be observed in Fig. 4. Note that the optimal restart time was not used on this occasion, but $t = \exp(\mu - \sigma^2)$, as in the proof for Theorem 5.

For the sake of completeness, the case in which the shape parameter σ is fixed and the scale parameter μ is a variable illustrated as well. In Fig. 5, σ is 3. In this scenario, it seems that the difference between $E[X_{\text{Luby}}]$ and α^* only increases slightly. The relative difference decreases as μ increases.

The main findings of this section are that Luby's strategy is, in many circumstances, a good alternative to the optimal fixed-cutoff strategy. Especially if the exact probability distribution is unknown, however, it is worth checking if one of the boundary cases is present. Then the relative difference of the expected values may

Fig. 4 These plots represent the expected values under restarts of the lognormal distribution. The scale parameter μ is fixed to 5, while the shape parameter σ is displayed on the x-axis. The resulting expected values are on the y-axis. Here, the fixed-cutoff strategy uses the restart time $t = e^{\mu - \sigma^2}$, which is, in general, not an optimal strategy. The orange dots correspond to the sample mean of Luby's strategy from 1000 runs. In other words, the closer the orange dots are to the blue dots, the better is Luby's strategy. The whiskers represent the 95% confidence interval. In the left plot, both axes are scaled linearly, while in the right plot, the y-axis is scaled logarithmically

Fig. 5 These plots represent the expected values under restarts of the lognormal distribution. The shape parameter σ is fixed to 3, while the scale parameter μ is displayed on the x-axis. The resulting expected values are on the y-axis. Here, the fixed-cutoff strategy uses the optimal restart time, as obtained in [9]. The orange dots correspond to the sample mean of Luby's strategy from 1000 runs. In other words, the closer the orange dots are to the blue line, the better is Luby's strategy. The whiskers represent the 95% confidence interval. If the whiskers are not visible, this indicates that the confidence interval is smaller than the dot size. In the left plot, both axes are scaled linearly, while in the right plot, the y-axis is scaled logarithmically

be significant, as has been illustrated by the GP distribution for low σ values and the lognormal distribution for high σ values.

5 Conclusion

Many established algorithms (e.g., [5, 6]) use restarts to optimize their respective expected runtimes. Among those algorithms, the use of Luby's universal strategy is especially prevalent. Luby's strategy consists of a slowly exponentially growing sequence of restart times. In their work, Luby et al. [12] showed that their universal strategy is asymptotically optimal in the worst case.

These properties have only been shown for discrete distributions. In this work, Luby's strategy and universal strategies are studied in a continuous setting. We show that Luby's strategy does not permit any bounds in a continuous setting.

Theorem 3 shows that in a continuous setting, there are no universal strategies with reasonable worst-case bounds. In other words, every universal strategy, including Luby's strategy, is arbitrarily worse than an optimal strategy for some distributions.

Universal strategies are attractive to practitioners because knowledge of the underlying distribution is not necessary. Nonetheless, some information about the distribution is often available. Therefore, we study under which circumstances there are worst-case bounds for Luby's strategy. Theorem 4 shows that if the optimal restart time is not too small, there are useful worst-case bounds. In particular, we improved the original bounds from Luby et al. [12] if the optimal restart time is at least 0.5. The results are also demonstrated with simulations. As can be observed, the results are often significantly better than the worst-case bound of Theorem 4.

In conclusion, there are two extreme cases: On the one hand, if the underlying distribution is completely known, then an optimal restart time can be calculated, resulting in an optimal strategy. On the other hand, if the distribution is entirely unknown, then only universal strategies are available. For future work, it would be interesting whether partial knowledge of the distribution can be exploited to efficiently create an intermediate restart strategy that performs better than Luby's strategy.

In addition, the generalized Pareto distribution does not satisfy the conditions of Theorem 4 if its parameters are chosen appropriately. On the other hand, the simulations have shown that Luby's strategy often yields good results in this case. Therefore the question arises whether the statement of Theorem 4 can be extended to cover cases like the generalized Pareto distribution.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Biere, A., Heule, M., van Maaren, H.: Handbook of Satisfiability. IOS Press (2009)
2. Eliazar, I., Koren, T., Klafter, J.: Searching circular DNA strands. *J. Phys.: Condens. Matter* **19**(6), 065140 (2007)
3. Evans, M., Majumdar, S.: Diffusion with stochastic resetting. *Phys. Rev. Lett.* **106**(16), 160601 (2011)
4. Florescu, I., Tudor, C.A.: Handbook of Probability. Wiley, New York (2013)
5. Friedrich, T., Kötzing, T., Quinzan, F., Sutton, A.: Improving the run time of the $(1 + 1)$ evolutionary algorithm with Luby sequences. In: Proceedings of the Genetic and Evolutionary Computation Conference, GECCO '18, pp. 301–308. ACM, New York (2018)
6. Gomes, C., Selman, B., Crato, N.: Heavy-tailed distributions in combinatorial search. In: CP'97, pp. 121–135. Springer (1997)
7. Huang, J.: The effect of restarts on the efficiency of clause learning. In: IJCAI, vol. 7, pp. 2318–2323 (2007)
8. König, W.: Stochastische Algorithmen. Universität zu Köln, lecture notes (in German). Available at <http://www.wias-berlin.de/people/koenig/www/AlgStoch.pdf> (2003)
9. Lorenz, J.: Runtime distributions and criteria for restarts. In: SOFSEM'18, pp. 493–507. Springer International Publishing, Cham (2018)
10. Lorenz, J.H.: On the complexity of restarting. In: International Computer Science Symposium in Russia, pp. 250–261. Springer (2019)
11. Lorenz, J.H., Schöning, U.: Promise problems on probability distributions. In: Complexity and Approximation, pp. 57–66. Springer (2020)
12. Luby, M., Sinclair, A., Zuckerman, D.: Optimal speedup of Las Vegas algorithms. *Inf. Process. Lett.* **47**(4), 173–180 (1993)
13. Norman, L., Kotz, S., Balakrishnan, N.: Continuous Univariate Distributions, vol 1 of Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics (1994)
14. Pal, A., Kundu, A., Evans, M.R.: Diffusion under time-dependent resetting. *J. Phys. A: Math. Theor.* **49**(22), 225001 (2016)
15. Pal, A., Reuveni, S.: First passage under restart. *Phys. Rev. Lett.* **118**(3), 030603 (2017)
16. Wolter, K.: Stochastic Models for Fault Tolerance: Restart, Rejuvenation and Checkpointing. Springer Science & Business Media, New York (2010)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.