Probabilistic Analysis of the $(1+1)$-Evolutionary Algorithm

Hsien-Kuei Hwang[1] hkhwang@stat.sinica.edu.tw
Institute of Statistical Science & Institute of Information Science, Academia Sinica, Taipei 115, Taiwan

Alois Panholzer[2] Alois.Panholzer@tuwien.ac.at
Institut für Diskrete Mathematik und Geometrie, Technische Universität Wien, Wiedner Hauptstraße 8-10/104, 1040 Wien, Austria

Nicolas Rolin nicolasjrolin@gmail.com
LIPN, Institut Galilée, Université Paris 13, 93430, Villetaneuse, France

Tsung-Hsi Tsai chonghi@stat.sinica.edu.tw
Institute of Statistical Science, Academia Sinica, Taipei 115, Taiwan

Wei-Mei Chen[3] wmchen@mail.ntust.edu.tw
Department of Electronic and Computer Engineering, National Taiwan University of Science and Technology, Taipei 106, Taiwan

Abstract

We give a detailed analysis of the optimization time of the $(1+1)$-Evolutionary Algorithm under two simple fitness functions (ONEMAX and LEADINGONES). The problem has been approached in the evolutionary algorithm literature in various ways and with different degrees of rigor. Our asymptotic approximations for the mean and the variance represent the strongest of their kind. The approach we develop is based on an asymptotic resolution of the underlying recurrences and can also be extended to characterize the corresponding limiting distributions. While most of our approximations can be derived by simple heuristic calculations based on the idea of matched asymptotics, the rigorous justifications are challenging and require a delicate error analysis.

Keywords

$(1+1)$-evolutionary algorithm, probabilistic analysis, OneMax function, LeadingOnes function, asymptotic approximations, error analysis, limit laws, recurrences.

1 Introduction

The last two decades or so have seen an explosion of application areas of evolutionary algorithms (EAs) in diverse scientific or engineering disciplines. An EA is a random search heuristic, using evolutionary mechanisms such as crossover and mutation, for finding a solution that often aims at optimizing an objective function. EAs proved to be extremely useful for combinatorial optimization problems because they can find good solutions for complicated problems using only basic mathematical modeling and simple operators with reasonable efficiency; see Coello Coello (2006); Deb (2001); Horn (1997) for more information. Although EAs have been widely

[1]Partially supported by a grant from the Franco-Taiwan Orchid Program (2013–2014).
[2]Supported by the Austrian Science Foundation FWF under the Grant P25337-N23.
[3]Partially supported by MOST under the Grant 103-2221-E-011-113.

applied in solving practical problems, the analysis of their performance and efficiency, which often provides better modeling prediction for potential uses in practice, is much less developed, and only computer simulation results are available for most of the EAs in use; see, for example, Beyer et al. (2002); Droste et al. (1998); Garnier et al. (1999); He and Yao (2001, 2002). We are concerned in this paper with a precise probabilistic analysis of a simple algorithm called $(1 + 1)$-EA.

A typical EA comprises several ingredients: the coding of solution, the population of individuals, the selection for reproduction, the operations for generating new individuals, and the fitness function to evaluate the new individual, and the mathematical analysis of the time complexity is often challenging mostly because the stochastic dynamics is difficult to capture. It proves more insightful to look instead at simplified versions of the algorithm, seeking for a compromise between mathematical tractability and general predictability. Such a consideration was first attempted in Bäck (1992) and Mühlenbein (1992) in the early 1990's for the $(1 + 1)$-EA, using only one individual with a single mutation operator at each stage. An outline of the procedure is as follows.

Algorithm $(1 + 1)$-EA

1. Choose an initial string $\mathbf{x} \in \{0, 1\}^n$ uniformly at random
2. Repeat until a terminating condition is reached
   - Create $\mathbf{y}$ by flipping each bit of $\mathbf{x}$ (with probability p of flipping), each bit independently of the others
   - Replace $\mathbf{x}$ by $\mathbf{y}$ iff $f(\mathbf{y}) \geqslant f(\mathbf{x})$

Step 1 is often realized by tossing a fair coin for each of the n bits, one independently of the others, and the terminating condition is either exhausting the number of assigned iterations or reaching a state when no further improvement has been observed for a given amount of time.

Mühlenbein (1992) considered in detail the complexity of $(1+1)$-EA under the fitness function OneMax, which counts the number of 1s, namely, $f(\mathbf{x}) = \sum_{1\leqslant j\leqslant n} x_j$. More precisely, let X_n denote the time needed to reach the optimum value (often referred to as the *optimization time* of OneMax). Then the expected time $\mathbb{E}(X_n)$ was argued to be of order $n \log n$, indicating the efficiency of the $(1 + 1)$-EA. Bäck (1992) derived expressions for the success, failure and stagnation probabilities of mutation. A finer asymptotic approximation of the form

$$\mathbb{E}(X_n) = en \log n + c_1 n + o(n), \tag{1}$$

was derived by Garnier et al. (1999), where $c_1 \approx -1.9$ when the mutation rate p equals $\frac{1}{n}$. They went further by characterizing the limiting distribution of $\frac{X_n - en\log n}{en}$ in terms of a log-exponential distribution (which is indeed a double exponential or a Gumbel distribution). However, some of their proofs, notably the error analysis, seem incomplete (as indicated in their paper). Thus a precise result such as (1) has remained obscure in the EA literature.

While the $(1+1)$-EA under simple fitness functions may seem too simplified to be of much practical value, the study of its complexity continues to attract the attention in the literature (see, for example, Auger and Doerr (2011); Neumann and Witt (2010) for more recent developments) for several reasons. First, the $(1 + 1)$-EA (under some fitness functions) represents one of the simplest models whose behavior is *mathematically tractable*. Second, the stochastic behaviors under such a simple formulation often have, although hard to prove, a wider range of applicability or predictability, either for more general models or for other meta-heuristics. Such a *complexity robustness* can on the other hand be checked by simulations, and in such a case, the theoretical results are useful in directing good guesses. For example, Neumann and Witt (2009) showed that 1-ANT behaves identically to the (1 + 1)-EA in some situations. Also Sudholt and Witt (2010) showed a similar behavior translation into Particle Swarm Optimization algorithms. Third, although tractable, most of the analyses of the (1 + 1)-EAs are far from being trivial, and different mathematical tools have been introduced or developed for such a purpose, leading to more general *methodological developments*, which may also be useful for the analysis of other EAs or heuristics. Fourth, from a mathematical point of view, few models can be solved precisely, and those that can be often exhibit additional *structural properties* that are fundamental and may be of interest for further investigation. Finally, understanding the expected complexity of algorithms may help in identifying hard inputs and in *improving the efficiency* of the algorithm; see, for example, Doerr and Doerr (2014); Doerr et al. (2013).

The expected optimization time required by (1 + 1)-EA has undergone successive improvements, yet none of them reached the precision of Garnier et al.'s result (1); we summarize in Figure 1 some recent findings; a brief account of earlier results can be found in Garnier et al. (1999).

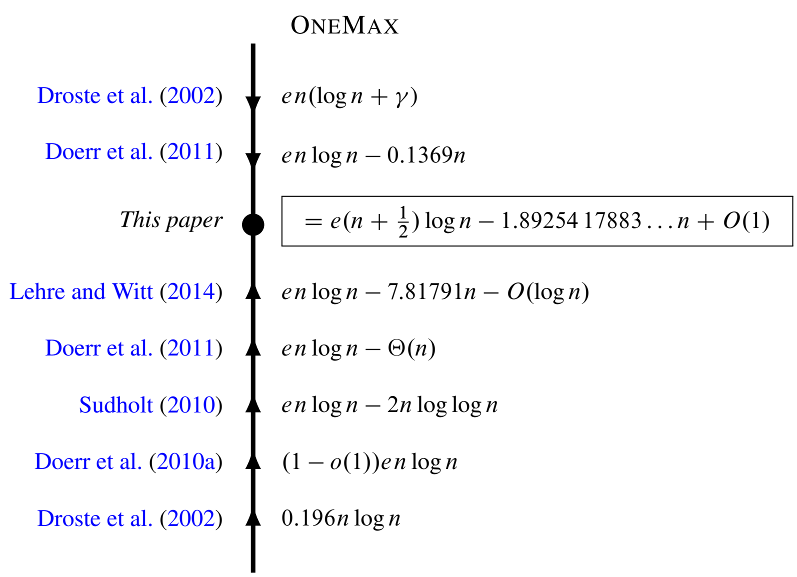


Figure 1: *Some known lower (indicated by an up-arrow) and upper (marked by a down-arrow) bounds for the optimization time of the* (1 + 1)*-EA under* OneMax.

Note that, by a result of Doerr et al. (2010b, 2012), which showed that OneMax is the easiest function among all functions with a unique optimum (particularly, among all linear functions), any lower bound for OneMax provides also a lower bound for linear functions. Thus the precise asymptotic bounds we derive in this paper may also extend as effective lower bounds for other fitness functions. On the other hand, Sudholt (2013) established that the (1 + 1)-EA is, for OneMax, the fastest non-adaptive EA that only uses standard bit mutations to create new offspring, starting with a single search point.

In this paper we focus on the mutation rate $p = \frac{1}{n}$ and prove that the expected number of steps taken by the (1 + 1)-EA to reach the optimum of OneMax function satisfies

$$\mathbb{E}(X_n) = en \log n + c_1 n + \tfrac{1}{2} e \log n + c_2 + O\big(n^{-1} \log n\big), \tag{2}$$

where c_1 and c_2 are explicitly computable constants. More precisely,

$$c_1 = -e\left(\log 2 - \gamma - \phi_1\left(\tfrac{1}{2}\right)\right) \approx -1.89254\,17883\,44686\,82302\,25714\ldots,$$

where $\gamma = \int_1^\infty \left(\frac{1}{\lfloor t\rfloor} - \frac{1}{t}\right) \mathrm{d}t \approx 0.57721\,56649$ is Euler's constant, and

$$\phi_1(z) := \int_0^z \left(\frac{1}{S_1(t)} - \frac{1}{t}\right) \mathrm{d}t, \tag{3}$$

with $S_1(z)$ an entire function defined by

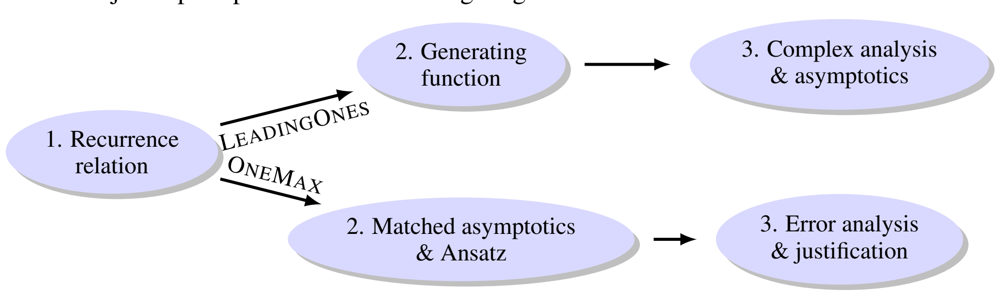
$$S_1(z) := \sum_{\ell \geqslant 1} \frac{z^\ell}{\ell!} \sum_{0 \leqslant j < \ell} (\ell - j) \frac{(1-z)^j}{j!}.$$

See (40) for an analytic expression and numerical value for c_2. Note that from an algorithmic point of view, a mutation rate $\gg \frac{1}{n}$ leads to a complexity $\Omega(n \log n)$; see Witt (2013) for more information.

These expressions, as well as the numerical values, are consistent with those given in Garnier et al. (1999). From the expression of c_1, it is clear that its characterization lies much deeper than the dominant term $en \log n$. Numerically, such a characterization is also important because $\log n$ is close to being a constant for moderate values of n, so that the overshoot (c_1 being negative) from the leading term $en \log n$ is not small in such a situation.

Finer properties such as more precise expansions for $\mathbb{E}(X_n)$, the variance and limiting distribution will also be established. In particular, the study of the variance provides a measure of spread of the asymptotic distribution, and is in line with recent research on tail probabilities; see, for example, Witt (2014); Zhou et al. (2012). The extension to $p = \frac{c}{n}$ does not lead to additional new phenomena as already discussed in Garnier et al. (1999); it is thus omitted in this paper.

Our approach relies essentially on the asymptotic resolution of the underlying recurrence relation for the optimization time, and the method of proof is different from all previous approaches (including Markov chains, coupon collection, coupling, drift analysis, etc.). It consists of three major steps depicted in the following diagram.



Briefly, due to the recursive nature of the algorithm, we first derive the corresponding recurrence relation satisfied by the random variables that capture the remaining optimization time from different states of the algorithm. In case when the recurrence can be solved by techniques from *analytic combinatorics* through the use of generating functions, the corresponding asymptotic approximations can often be obtained by suitable complex-analytic tools such as singularity analysis and saddle-point method; see the authoritative book Flajolet and Sedgewick (2009) for more information. The analysis of (1 + 1)-EA under LeadingOnes belongs to such a case; see Section 7 for details. On the other hand, when such a generating function-based approach fails to provide more manageable forms (in terms of functional or differential equations), a different route through "matched asymptotics" may be attempted, which is the one we adopt for the analysis of the (1 + 1)-EA under ONEMAX. Roughly, we identify terms with the largest contribution on the right-hand side and guess the right form (the *Ansatz*) by matching the asymptotic expansions on both sides of the recurrence. The Ansatz is, once postulated, often easily checked by direct numerical calculations. The final stage is to justify the Ansatz by a proper *error analysis*, which often involves a delicate asymptotic analysis; see Wong (2014) for a recent survey of techniques for recurrences of linear type. Our recurrences are, however, of a nonlinear nature, and involve two parameters. On the other hand, these two approaches are not exclusive, but instead complementary in many cases. For example, we rely on generating functions and complex-analytic tools for the proof of several auxiliary results in this paper.

More precisely, we consider $f(\mathbf{x}) = \sum_{1 \leqslant j \leqslant n} x_j$ and study the random variables $X_{n,m}$, which counts the number of steps taken by $(1+1)$-EA before reaching the optimum state $f(\mathbf{x}) = n$ when starting with $n-m$ 1s (namely, $f(\mathbf{x}) = n-m$). We will derive very precise asymptotic approximations for each $X_{n,m}$, $1 \leqslant m \leqslant n$. In particular, the distribution of $X_{n,m}$ is for large n well approximated by a sum of m exponential distributions, and this in turn implies a Gumbel limit law when $m \to \infty$. Then the time for X_n to reach the optimum state by the (1 + 1)-EA when starting with a random initial configuration (every bit being Bernoulli($\frac{1}{2}$)) can be readily characterized because the binomial distribution is highly concentrated near the mean; see Table 1 for a summary of our major results.

In addition to its own methodological merit of obtaining stronger asymptotic approximations and potential use in other problems in similar EAs, our approach, to the best of our knowledge, provides the first rigorous justification of the far-reaching results of Garnier et al. (1999) more than seventeen years ago. It also sheds new light on further potential use of similar techniques to related problems of a recursive nature.

This paper is organized as follows. We begin with deriving the recurrence relation satisfied by the random variables $X_{n,m}$ (when the initial configuration is not random). From this recurrence, it is straightforward to characterize inductively the distribution of $X_{n,m}$ for small $1 \leqslant m = O(1)$. The hard case when $m \to \infty, m \leqslant n$ requires the development of more asymptotic tools, which we elaborate in Section 3. Asymptotics of the mean values of $X_{n,m}$ and X_n are presented in Section 4 with a complete error analysis. Section 5 then addresses the asymptotics of the variance. Limit laws are established in Section 6 by an inductive argument and fine error analysis. Finally, we consider briefly in Section 7 the optimization time of the $(1+1)$-EA for the LEADINGONES problem. Denote the corresponding optimization time by Y_n. We summarize the major results in Table 1. Note that all results for Y_n have previously been obtained in Ladret (2005) and we will sketch a different self-contained method of proof for them.

Some technical material is collected in Appendices A–F.

Notation. *Throughout this paper, all O-terms are with respect to $n \to \infty$ unless otherwise stated. We say that a quantity $X = O(f(n,m))$ uniformly for $m = O(1)$ as $n \to \infty$, or in words "$X = f(n,m)$ uniformly for bounded m and large n", if there exists a $C > 0$ such that for any $m \geqslant 0$ there is an $n_0 > 0$ such that $|X| \leqslant Cf(n,m)$ for all $n \geqslant n_0$. Here n_0 may depend on C and m. The definition extends similarly when $X = O(f(n,m,t))$ holds uniformly for $m, t = O(1)$.*

2 Recurrence and the limit laws of $X_{n,m}$ when $m = O(1)$

In this section, we derive first a recurrence relation satisfied by the probability generating function $P_{n,m}(t) := \mathbb{E}\big(t^{X_{n,m}}\big)$ of $X_{n,m}$, where $X_{n,m}$ denotes the number of steps taken by $(1+1)$-EA to reach $f(\mathbf{x}) = n$ for the first time when starting from the initial state $f(\mathbf{x}) = n-m$. From this recurrence and starting with $P_{n,0}(t) = 1$, we can then get closed-form expressions one

Properties \ Fitness	OneMax (X_n)	LeadingOnes (Y_n)
Mean $\sim$	$en\log n + c_1 n$	$\frac{e-1}{2}\,n^2$
Variance $\sim$	$\frac{\pi^2}{6}(en)^2 - (2e+1)en\log n$	$\frac{e^2-1}{8}\,n^3$
Limit law	Gumbel distribution $\mathbb{P}\left(\frac{X_n}{en} - \log\frac{n}{2} - \phi_1(\frac12) \leqslant x\right)$ $\to e^{-e^{-x}}$	Gaussian distribution $\mathbb{P}\left(\frac{Y_n - \frac{e-1}{2}n^2}{\sqrt{\frac{e^2-1}{8}n^3}} \leqslant x\right)$ $\to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,\mathrm{d}t$
Approach	Ansatz & error analysis	Analytic combinatorics

Table 1: *A summary of our findings for the time complexity of the* $(1+1)$*-EA under* OneMax *(X_n) and* LeadingOnes *(Y_n) fitness function, respectively, when starting from a random initial state and when the mutation probability is* $\frac{1}{n}$*. The constant* c_1 *is defined in* (2) *and the function* ϕ_1 *in* (3)*. The symbol* $a_n \sim b_n$ *here means that the ratio* $\frac{a_n}{b_n}$ *tends to* 1 *as* $n \to \infty$*.*

after another by iterating the recurrence, but the expressions soon become too cumbersome. We then use a simple inductive argument to derive the corresponding limit laws when m remains bounded, together with an asymptotic approximation to the mean and one to the variance. These results not only reveal the complexity of the analytic problem when viewed from a generating function perspective but also serve to introduce the prototype forms of the mean and variance asymptotics, respectively, which we will examine in more detail later.

2.1 Recurrence for $P_{n,m}(t)$

Lemma 1. *The probability generating function* $P_{n,m}(t)$ *satisfies the recurrence*

$$P_{n,m}(t) = \frac{t\sum_{1\leqslant \ell\leqslant m}\lambda_{n,m,\ell}P_{n,m-\ell}(t)}{1-\left(1-\sum_{1\leqslant\ell\leqslant m}\lambda_{n,m,\ell}\right)t} \qquad (1\leqslant m\leqslant n), \tag{4}$$

with $P_{n,0}(t) = 1$*, where*

$$\lambda_{n,m,\ell} := \left(1-\frac{1}{n}\right)^n (n-1)^{-\ell}\sum_{0\leqslant j\leqslant \min\{n-m,m-\ell\}}\binom{n-m}{j}\binom{m}{j+\ell}(n-1)^{-2j}. \tag{5}$$

The leading factor $\left(1-\frac{1}{n}\right)^n$ in (5) is the origin of the pervasive presence of "e" in our asymptotic approximations.

Proof. Start from the state $f(\mathbf{x}) = n-m$ and run the two steps inside the loop of Algorithm $(1+1)$-EA. The new state becomes $\mathbf{y}$ with $f(\mathbf{y}) = n-m+\ell$ if j bits in the group $\{x_i = 1\}$ and $j+\ell$ bits in the other group $\{x_i = 0\}$ toggled their values, where $0\leqslant j\leqslant\min\{n-m,m-\ell\}$ and $\ell > 0$. Thus, the probability from state $\mathbf{x}$ to $\mathbf{y}$ is given by

$$\lambda_{n,m,\ell} = \sum_{0\leqslant j\leqslant\min\{n-m,m-\ell\}}\binom{n-m}{j}\left(\frac{1}{n}\right)^j\left(1-\frac{1}{n}\right)^{n-m-j}\binom{m}{j+\ell}\left(\frac{1}{n}\right)^{j+\ell}\left(1-\frac{1}{n}\right)^{m-j-\ell},$$

which is identical to (5). Since (see Figure 2.1)

$$X_{n,m} \stackrel{d}{=} \begin{cases} 1 + X_{n,m-\ell}, & \text{with probability } \lambda_{n,m,\ell},\ 1\leqslant\ell\leqslant m;\\ 1 + X_{n,m}, & \text{with probability } 1-\sum_{1\leqslant\ell\leqslant m}\lambda_{n,m,\ell},\end{cases}$$

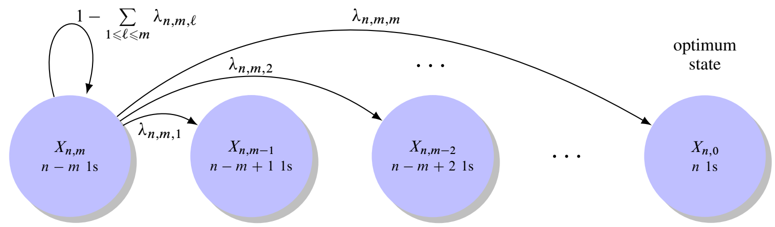


Figure 2: *The transition of states and their probabilities.*

where the symbol $\stackrel{d}{=}$ denotes distributional equivalence, we see that

$$P_{n,m}(t) = t \sum_{1\leqslant \ell \leqslant m} \lambda_{n,m,\ell} P_{n,m-\ell}(t) + \left(1 - \sum_{1\leqslant \ell \leqslant m} \lambda_{n,m,\ell}\right) t P_{n,m}(t), \tag{6}$$

and this proves the lemma. ▮

While this simple recurrence relation is not new in the EA literature (see, for example, Bäck (1992); Garnier et al. (1999); He and Yao (2003)), tools have been lacking for a direct asymptotic resolution, which we will develop in detail in this paper.

From a computational point of view (notably for higher moments), it is often preferable to use the following recurrence because fewer terms depending on t are involved.

Corollary 1. *For* $1 \leqslant m \leqslant n$

$$\sum_{1\leqslant \ell \leqslant m} \lambda_{n,m,\ell} \left(P_{n,m}(t) - P_{n,m-\ell}(t)\right) = \left(1 - t^{-1}\right) P_{n,m}(t). \tag{7}$$

Proof. This follows from dividing both sides of (6) by t and then rearranging terms there. ▮

For convenience, define

$$\Lambda_{n,m} := \sum_{1\leqslant \ell \leqslant m} \lambda_{n,m,\ell},$$

which can be interpreted as the probability that a mutation is successful in increasing the fitness (objective function value).

2.2 $X_{n,1}$: from geometric to exponential

As the first nontrivial case beyond $X_{n,0} = 0$, here we study in detail $X_{n,1}$ or, equivalently, $P_{n,1}(t)$. In this case,

$$\Lambda_{n,1} = \lambda_{n,1,1} = \frac{1}{n}\left(1 - \frac{1}{n}\right)^{n-1},$$

so that, by (4),

$$P_{n,1}(t) = \frac{\frac{1}{n}\left(1-\frac{1}{n}\right)^{n-1} t}{1 - \left(1 - \frac{1}{n}\left(1 - \frac{1}{n}\right)^{n-1}\right) t}.$$

This is a standard geometric distribution Geo(p') (assuming only positive integer values) with probability

$$p' = \frac{1}{n}\left(1 - \frac{1}{n}\right)^{n-1} = \frac{1}{en}\left(1 + O\left(\frac{1}{n}\right)\right).$$

This implies that $\mathbb{E}(X_{n,1}) = \frac{1}{p'} = en\big(1 + O(\frac{1}{n})\big)$, and it is natural to consider the normalized random variable $\frac{X_{n,1}}{en}$ for which we have, by taking $t = e^{\frac{i\theta}{en}}$ ($\theta \in \mathbb{R}$) and by using the expansions $e^{iu} = 1 + O(|u|)$ and $e^{iu} = 1 + iu + O(u^2)$ for bounded $u \in \mathbb{R}$,

$$\begin{aligned} P_{n,1}\big(e^{\frac{i\theta}{en}}\big) &= \frac{\frac{1}{en}\big(1 + O(\frac{1}{n})\big)\big(1 + O(\frac{|\theta|}{n})\big)}{1 - \big(1 - \frac{1}{en}\big(1 + O(\frac{1}{n})\big)\big)\big(1 + \frac{i\theta}{en} + O(\frac{\theta^2}{n^2})\big)} \\ &= \frac{\frac{1}{en}\big(1 + O\big(\frac{1+|\theta|}{n}\big)\big)}{\frac{1}{en}(1 - i\theta) + O\big(\frac{1+\theta^2}{n^2}\big)} \\ &\to \frac{1}{1 - i\theta}, \end{aligned} \tag{8}$$

as $n \to \infty$, for bounded real θ. Note that the error term for the last convergence is of the form $O\big(\frac{1+\theta^2}{n|1-i\theta|^2}\big)$, which holds uniformly for bounded θ, but we do not need this uniform estimate. Also observe that $\frac{1}{1-i\theta}$ is the characteristic function of an exponential distribution with parameter 1. The passage from the convergence of a sequence of characteristic functions to that of the corresponding distribution functions can be justified by Lévy's continuity theorem (van der Vaart, 1998, §2.3) or (Flajolet and Sedgewick, 2009, §IX 4.2):

> *If the sequence of characteristic functions $\{\varphi_n(\theta)\}$ of the random variables $\{X_n\}$ converges pointwise to $\varphi(\theta)$ as $n \to \infty$ for $\theta \in \mathbb{R}$, and $\varphi(\theta)$ is continuous at zero, then X_n converges in distribution to X whose characteristic function is $\varphi(\theta)$.*

This and (8) imply the convergence in distribution

$$\frac{X_{n,1}}{en} \xrightarrow{d} \mathrm{Exp}(1),$$

where $\mathrm{Exp}(c)$ denotes an exponential distribution with parameter c. Equivalently, this can be rewritten as

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{X_{n,1}}{en} \leqslant x\right) = 1 - e^{-x},$$

for $x > 0$. Such a limit law indeed extends to the case when $m = O(1)$, which we formulate in the next subsection.

2.3 The distribution of $X_{n,m}$ when $m = O(1)$

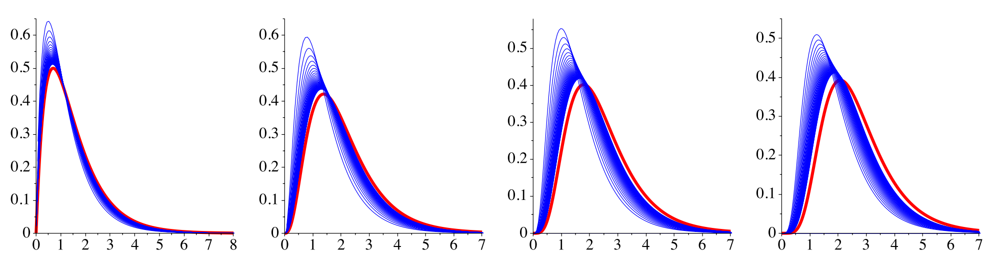


Figure 3: *Histograms of $\frac{X_{n,2j}}{en}$ for $j = 1, \ldots, 4$ (in left to right order) and $n = 5, \ldots, 50$ (in topdown order when starting from the peak in each figure), and their corresponding limit laws.*

Let $H_m^{(r)} = \sum_{1\leqslant j\leqslant m} \frac{1}{j^r}$ denote the r-th order harmonic numbers and $H_m = H_m^{(1)}$. For convenience, we define $H_0^{(r)} = 0$.

Theorem 1. *Assume $1 \leqslant m = O(1)$. Then the time used by $(1+1)$-EA to reach the optimum state $f(\mathbf{x}) = n$, when starting from $f(\mathbf{x}) = n-m$, converges (after normalized by en) to a sum of m exponential random variables*

$$\frac{X_{n,m}}{en} \xrightarrow{d} \sum_{1 \leqslant r \leqslant m} \operatorname{Exp}(r); \tag{9}$$

moreover, the mean of $X_{n,m}$ is asymptotic to enH_m and the variance to $(en)^2 H_m^{(2)}$.

From the integral representation (by induction) for the moment generating function of $\sum_{1 \leqslant r \leqslant m} \operatorname{Exp}(r)$:

$$\prod_{1 \leqslant r \leqslant m} \frac{1}{1 - \frac{s}{r}} = \int_0^\infty m(1-e^{-x})^{m-1} e^{x(s-1)} \,\mathrm{d}x \qquad (s < 1; m \geqslant 1),$$

we see that the convergence in distribution (9) can alternatively be expressed in the more transparent form

$$\lim_{n \to \infty} \mathbb{P}\left(\frac{X_{n,m}}{en} \leqslant x\right) = (1-e^{-x})^m \qquad (x > 0);$$

see Figure 3 for some plots for the density functions of $X_{n,m}$.

Before proving this theorem, we derive a simple estimate for $\lambda_{n,m,\ell}$, which will be useful in our analysis below.

Lemma 2. *Assume $1 \leqslant m = O(1)$. Then*

$$\lambda_{n,m,\ell} = \binom{m}{\ell} e^{-1} n^{-\ell} \left(1 + O\left(\frac{m-\ell}{n(\ell+1)} + \frac{\ell}{n}\right)\right) \qquad (1 \leqslant \ell \leqslant m), \tag{10}$$

where the O-term holds uniformly in ℓ and m.

Proof. When $\ell = m$

$$\lambda_{n,m,m} = \left(1 - \frac{1}{n}\right)^n n^{-m} = \binom{m}{m} e^{-1} n^{-m}\big(1 + O\big(n^{-1}\big)\big) \qquad (m \geqslant 1),$$

so (10) holds. Assume now $1 \leqslant \ell < m$. By the sum definition (5) of $\lambda_{n,m,\ell}$, we see that when $m = O(1)$, the binomial factor $\binom{m}{j+\ell}$ is also bounded, and the other factors $\binom{n-m}{j}(n-1)^{-2j}$ decrease with j (for fixed n and m), which means that the largest term comes from $j = 0$, all other terms being of a smaller order. More precisely, the term with $j = 0$ equals

$$\left(1 - \frac{1}{n}\right)^n \binom{m}{\ell} (n-1)^{-\ell} = \binom{m}{\ell} e^{-1} n^{-\ell} \left(1 + O\left(\frac{\ell}{n}\right)\right),$$

and the contribution from the remaining terms (with $j \geqslant 1$) is bounded above by

$$\begin{aligned}
\sum_{j \geqslant 1} \binom{m}{j+\ell}\binom{n-m}{j}(n-1)^{-2j-\ell} &\leqslant \sum_{j \geqslant 1} \binom{m}{j+\ell} \frac{(n-1)^{-j-\ell}}{j!} \\
&= \binom{m}{\ell+1} \sum_{j \geqslant 1} \frac{(n-1)^{-j-\ell}}{j!} \cdot \frac{\binom{m}{j+\ell}}{\binom{m}{\ell+1}} \\
&= \binom{m}{\ell+1} (n-1)^{-\ell} \sum_{j \geqslant 1} \frac{(n-1)^{-j}}{j!} \prod_{1 \leqslant r < j} \frac{m-\ell-r}{\ell+1+r} \\
&< \binom{m}{\ell+1} (n-1)^{-\ell} \sum_{j \geqslant 1} \frac{m^{j-1}}{j!(n-1)^j \ell^{j-1}}.
\end{aligned}$$

The last series then satisfies

$$\sum_{j\geqslant 1}\frac{m^{j-1}}{j!(n-1)^j\ell^{j-1}} = O\bigg(\frac{1}{n}\sum_{j\geqslant 0}\frac{1}{j!}\Big(\frac{m}{n\ell}\Big)^j\bigg) = O\big(n^{-1}\big),$$

for $1 \leqslant \ell < m$ and $m = O(1)$. So we have for m, ℓ in the same range

$$\sum_{j\geqslant 1}\binom{m}{j+\ell}\binom{n-m}{j}(n-1)^{-2j-\ell} = O\left(\binom{m}{\ell+1}n^{-\ell-1}\right) = O\left(\frac{m-\ell}{n(\ell+1)}\binom{m}{\ell}n^{-\ell}\right).$$

Collecting these estimates, we then obtain (10). ∎

Corollary 2. *For $1 \leqslant m = O(1)$, the probability that a mutation succeeds in increasing the fitness is asymptotic to*

$$\sum_{1\leqslant\ell\leqslant m}\lambda_{n,m,\ell} = \frac{m}{en}\left(1 + O\left(\frac{m}{n}\right)\right), \tag{11}$$

where the O-term holds uniformly in m.

Proof. By (10)

$$\begin{aligned}\sum_{1\leqslant\ell\leqslant m}\lambda_{n,m,\ell} &= e^{-1}\sum_{1\leqslant\ell\leqslant m}\binom{m}{\ell}n^{-\ell}\left(1+O\left(\frac{m-\ell}{n(\ell+1)}+\frac{\ell}{n}\right)\right)\\
&= \frac{m}{en} + O\left(\sum_{2\leqslant\ell\leqslant m}\binom{m}{\ell}n^{-\ell} + \sum_{1\leqslant\ell\leqslant m}\binom{m}{\ell+1}n^{-\ell-1} + \sum_{2\leqslant\ell\leqslant m}\binom{m}{\ell}\ell n^{-\ell-1}\right)\\
&= \frac{m}{en} + O\left(\sum_{\ell\geqslant 2}\frac{1}{\ell!}\Big(\frac{m}{n}\Big)^\ell + \frac{1}{n}\sum_{\ell\geqslant 1}\frac{\ell}{\ell!}\Big(\frac{m}{n}\Big)^\ell\right)\\
&= \frac{m}{en} + O\left(\frac{m^2}{n^2}\right),\end{aligned}$$

from which (11) follows. ∎

Proof of Theorem 1: Limit laws. Using the fact that $\big|P_{n,m}(e^{it})\big| \leqslant 1$ for $t \in \mathbb{R}$ (being a characteristic function), we then deduce, by (4), (10) and (11), that

$$\begin{aligned}P_{n,m}\big(e^{\frac{i\theta}{en}}\big) &= \frac{\lambda_{n,m,1}e^{\frac{i\theta}{en}}P_{n,m-1}\big(e^{\frac{i\theta}{en}}\big) + O\left(\sum_{2\leqslant\ell\leqslant m}\lambda_{n,m,\ell}\right)}{1-\left(1-\frac{m}{en}\left(1+O\left(\frac{m}{n}\right)\right)\right)e^{\frac{i\theta}{en}}}\\
&= \frac{\frac{m}{en}\big(1+O\big(\frac{m}{n}\big)\big)\big(1+O\big(\frac{|\theta|}{n}\big)\big)P_{n,m-1}\big(e^{\frac{i\theta}{en}}\big) + O\big(\frac{m^2}{n^2}\big)}{\frac{m-i\theta}{en} + O\big(\frac{m^2+\theta^2}{n^2}\big)}\\
&= \frac{P_{n,m-1}\big(e^{\frac{i\theta}{en}}\big)}{1-\frac{i\theta}{m}}\left(1+O\left(\frac{m}{n|1-\frac{i\theta}{m}|}+\frac{m+|\theta|}{n}\right)\right) + O\left(\frac{m}{n|1-\frac{i\theta}{m}|}\right).\end{aligned}$$

Since $P_{n,m}(s)$ is an analytic function of s near unity and $P_{n,m}(1) = 1$, there exists an interval in which $|P(e^{it})| \geqslant c$ for some $c > 0$. Thus we can rewrite the above relation as

$$P_{n,m}\big(e^{\frac{i\theta}{en}}\big) = \frac{P_{n,m-1}\big(e^{\frac{i\theta}{en}}\big)}{1-\frac{i\theta}{m}}(1+\varepsilon_m(\theta)),$$

where

$$\varepsilon_m(\theta) = O\left(\frac{m + |\theta|}{n}\right),$$

the O-term holding uniformly for bounded $|\theta|$ and $1 \leqslant m = O(1)$. By iterating this relation m times using $P_{n,0}(t) = 1$, we obtain

$$\begin{aligned}
P_{n,m}\left(e^{\frac{i\theta}{en}}\right) &= \left(\prod_{1\leqslant r\leqslant m} \frac{1}{1 - \frac{i\theta}{r}}\right)\left(\prod_{1\leqslant \ell\leqslant m} (1 + \varepsilon_\ell(\theta))\right) \\
&= \left(\prod_{1\leqslant r\leqslant m} \frac{1}{1 - \frac{i\theta}{r}}\right)\left(\prod_{1\leqslant \ell\leqslant m} \left(1 + O\left(\frac{\ell + |\theta|}{n}\right)\right)\right) \\
&= \left(\prod_{1\leqslant r\leqslant m} \frac{1}{1 - \frac{i\theta}{r}}\right)\left(1 + O\left(\frac{m^2 + m|\theta|}{n}\right)\right), \qquad (12)
\end{aligned}$$

uniformly for bounded θ and $m = O(1)$. This and Lévy's continuity theorem (van der Vaart, 1998, §2.3) imply (9).

Note that if we compute formally the Taylor expansion at $\theta = 0$ of the right-hand side of the uniform asymptotic estimate (12), we obtain

$$\prod_{1\leqslant r\leqslant m} \frac{1}{1 - \frac{i\theta}{r}} = \exp\left(H_m i\theta - \tfrac{1}{2} H_m^{(2)}\theta^2 + O\left(|\theta|^3\right)\right),$$

so we expect that the mean and the variance of $\frac{X_{n,m}}{en}$ will be asymptotic to H_m and $H_m^{(2)}$, respectively, which is true and in consistency with the results obtained by the Quasi-power framework (see Hwang (1998) or (Flajolet and Sedgewick, 2009, Sec. IX.9)). For self-containedness and to pave the way for more refined arguments later, we will instead prove these two estimates by a direct, independent approach.

Proof of Theorem 1: Mean value. We now turn to the mean $\mu_{n,m} := \mathbb{E}(X_{n,m})$, which satisfies, by taking derivative with respect to t and then substituting $t = 1$ in (4), the recurrence

$$\sum_{1\leqslant \ell\leqslant m} \lambda_{n,m,\ell}\left(\mu_{n,m} - \mu_{n,m-\ell}\right) = 1. \qquad (13)$$

By the same reasoning used above for $P_{n,m}(t)$ based on (10), we see that the largest contribution on the left-hand side comes from terms with $\ell = 1$, and we expect the estimate

$$\frac{m}{en}(\mu_{n,m} - \mu_{n,m-1}) \sim 1,$$

or

$$\mu_{n,m} \sim \mu_{n,m-1} + \frac{en}{m},$$

which, by iteration, yields $\mu_{n,m} \sim enH_m$. The error analysis to justify this is not difficult but less interesting. Isolating first the terms corresponding to $\ell = 1$ and then rearranging all other terms, we get, again by (10),

$$\begin{aligned}\mu_{n,m} - \mu_{n,m-1} &= \frac{1}{\lambda_{n,m,1}}\left(1 - \sum_{2\leqslant\ell\leqslant m}\lambda_{n,m,\ell}(\mu_{n,m}-\mu_{n,m-\ell})\right)\\ &= \frac{en}{m}\left(1+O\left(\frac{m}{n}\right)\right)\left(1+O\left(\mu_{n,m}\sum_{2\leqslant\ell\leqslant m}\binom{m}{\ell}n^{-\ell}\right)\right)\\ &= \frac{en}{m}\left(1+O\left(\frac{m}{n}\right)\right)+O\left(\frac{m}{n}\,\mu_{n,m}\right).\end{aligned} \tag{14}$$

Thus we can rewrite this as

$$\mu_{n,m}(1+\varepsilon_m) = \mu_{n,m-1} + \frac{en}{m}\,(1+\delta_m)\,,$$

where both $\varepsilon_m, \delta_m = O(\frac{m}{n})$. Note that ε_m and δ_m may be negative. Rearranging this recurrence as

$$\mu_{n,m} = \frac{\mu_{n,m-1}}{1+\varepsilon_m} + \frac{en}{m}\cdot\frac{1+\delta_m}{1+\varepsilon_m}.$$

A direct iteration gives, by $\mu_{n,0} = 0$,

$$\begin{aligned}\mu_{n,m} &= en\sum_{1\leqslant\ell\leqslant m}\frac{1}{\ell}(1+\delta_\ell)\prod_{\ell\leqslant j\leqslant m}\frac{1}{1+\varepsilon_j}\\ &= en\sum_{1\leqslant\ell\leqslant m}\frac{1}{\ell}\left(1+O\left(\frac{\ell}{n}\right)\right)\left(1+O\left(\frac{m^2}{n}\right)\right)\\ &= enH_m + O(m^2H_m).\end{aligned}$$

This proves the required estimate for the mean when $m = O(1)$.

Proof of Theorem 1: Variance of $X_{n,m}$. To compute the variance, one may start with the second moment and then consider the difference with the square of the mean; however, it is computationally more advantageous to study directly the recurrence satisfied by the variances themselves.

For that purpose, we begin with (7) and substitute $t = e^s$, obtaining

$$\sum_{1\leqslant\ell\leqslant m}\lambda_{n,m,\ell}\big(P_{n,m}(e^s) - P_{n,m-\ell}(e^s)\big) = (1-e^{-s})P_{n,m}(e^s).$$

To derive a recurrence satisfied by the variance, we consider the moment generating function for the centered random variables $X_{n,m} - \mu_{n,m}$

$$\bar{P}_{n,m}(s) := P_n(e^s)e^{-s\mu_{n,m}},$$

which then satisfies the recurrence

$$\sum_{1\leqslant\ell\leqslant m}\lambda_{n,m,\ell}\big(\bar{P}_{n,m}(s) - \bar{P}_{n,m-\ell}(s)e^{-s(\mu_{n,m}-\mu_{n,m-\ell})}\big) = (1-e^{-s})\bar{P}_{n,m}(s),$$

for $1 \leqslant m \leqslant n$ with $\bar{P}_{n,0}(s) = 1$. Let $\sigma^2_{n,m} = \mathbb{V}(X_{n,m}) = \bar{P}''_{n,m}(0)$ be the variance of $X_{n,m}$. Then $\sigma^2_{n,m}$ satisfies the recurrence

$$\sum_{1\leqslant\ell\leqslant m}\lambda_{n,m,\ell}\left(\sigma^2_{n,m} - \sigma^2_{n,m-\ell}\right) = -1 + \sum_{1\leqslant\ell\leqslant m}\lambda_{n,m,\ell}\,\left(\mu_{n,m}-\mu_{n,m-\ell}\right)^2. \tag{15}$$

Heuristically, the dominant terms on both sides come from $\ell = 1$ when $m = O(1)$, and we expect that ($H_m - H_{m-1} = \frac{1}{m}$)

$$\frac{m}{en}\left(\sigma_{n,m}^2 - \sigma_{n,m-1}^2\right) \sim \frac{m}{en} \times \frac{(en)^2}{m^2},$$

or

$$\sigma_{n,m}^2 \sim \sigma_{n,m-1}^2 + \frac{(en)^2}{m^2};$$

iterating this recurrence $m-1$ times leads to the required estimate $\sigma_{n,m}^2 \sim e^2 H_m^{(2)} n^2$ for the variance.

For the justification, we follow the same argument used above for $\mu_{n,m}$. First, we rearrange the terms in (15) as

$$\begin{aligned}\sigma_{n,m}^2 - \sigma_{n,m-1}^2 &= (\mu_{n,m} - \mu_{n,m-1})^2 - \frac{1}{\lambda_{n,m,1}} \\ &\quad - \sum_{2\leqslant \ell \leqslant m} \frac{\lambda_{n,m,\ell}}{\lambda_{n,m,1}} \left(\sigma_{n,m}^2 - \sigma_{n,m-\ell}^2\right) + \sum_{2\leqslant \ell \leqslant m} \frac{\lambda_{n,m,\ell}}{\lambda_{n,m,1}} \left(\mu_{n,m} - \mu_{n,m-\ell}\right)^2.\end{aligned}$$

Then by the estimates (10) and (14), we obtain

$$\begin{aligned}\sigma_{n,m}^2 - \sigma_{n,m-1}^2 &= \left(\frac{en}{m} + O(1 + mH_m)\right)^2 + O\left(\frac{n}{m}\right) + O\left(\frac{m}{n}\sigma_{n,m}^2\right) + O\left(mnH_m^2\right) \\ &= \left(\frac{en}{m}\right)^2 + O\left(mnH_m^2\right) + O\left(\frac{m}{n}\sigma_{n,m}^2\right),\end{aligned}$$

and we are led to the form

$$\sigma_{n,m}^2(1+\varepsilon_m) = \sigma_{n,m-1}^2 + \frac{(en)^2}{m^2}(1+\delta_m),$$

where $\varepsilon_m = O\left(\frac{m}{n}\right)$ and $\delta_m = O\left(\frac{m^3}{n} H_m^2\right)$ when $m = O(1)$. By a direct iteration, we obtain

$$\begin{aligned}\sigma_{n,m}^2 &= (en)^2 \sum_{1\leqslant \ell \leqslant m} \frac{1}{\ell^2}(1+\delta_\ell) \prod_{\ell \leqslant j \leqslant m} \frac{1}{1+\varepsilon_j} \\ &= (en)^2 \sum_{1\leqslant \ell \leqslant m} \frac{1}{\ell^2}\left(1 + O\left(\frac{\ell^3 H_\ell^2}{n}\right)\right)\left(1 + O\left(\frac{m^2}{n}\right)\right) \\ &= (en)^2 H_m^{(2)} + O\left(nm^2 H_m^2\right).\end{aligned}$$

This implies the asymptotic estimate $\sigma_{n,m}^2 \sim (en)^2 H_m^{(2)}$ when m is bounded, and completes the proof of Theorem 1. ∎

To summarize, we saw that both the mean and the variance satisfy the same type of recurrence

$$\sum_{1\leqslant \ell \leqslant m} \lambda_{n,m,\ell}\left(a_{n,m} - a_{n,m-\ell}\right) = b_{n,m},$$

with suitable initial conditions, for some given sequences $b_{n,m}$. We also observe the transfer between the asymptotics of $b_{n,m}$ and that of $a_{n,m}$

$$m = O(1): \begin{cases} b_{n,m} \sim 1 & \Longrightarrow \quad a_{n,m} \sim enH_m \\ b_{n,m} \sim \frac{(en)^2}{m^2} & \Longrightarrow \quad a_{n,m} \sim (en)^2 H_m^{(2)}. \end{cases}$$

Such a correspondence will be useful in guiding our guess of the right Ansatz to be explored below when m lies in the most interesting range $m \asymp n$ (the symbol means that the left-hand side is of the same growth order as the right-hand side).

The simple inductive argument we used here extends to a wider range than $m = O(1)$ (as obvious from the error terms established) but fails when, say $m \gg \sqrt{n}/\log n$. In order to cover the whole range $1 \leqslant m \leqslant n$, we will need more refined uniform estimates for the error terms, which will be dealt with in Section 4. Some of the tools needed are developed in the next section.

2.4 Asymptotic expansions and Ansätze for $\mathbb{E}(X_{n,m})$

Normalizing the mean values. Let $\mu_{n,m} := \mathbb{E}(X_{n,m}) = P'_{n,m}(1)$. For simplicity, we consider

$$\mu^*_{n,m} := \frac{e_n}{n}\,\mu_{n+1,m},$$

where $e_n := \left(1 - \frac{1}{n+1}\right)^{n+1}$, so that $\mu^*_{n,m}$ satisfies, by (13), the simpler-looking recurrence

$$\sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell}\left(\mu^*_{n,m} - \mu^*_{n,m-\ell}\right) = \frac{1}{n} \qquad (1 \leqslant m \leqslant n), \tag{16}$$

with $\mu^*_{n,0} = 0$, where

$$\lambda^*_{n,m,\ell} := \frac{\lambda_{n+1,m,\ell}}{e_n} = \sum_{0\leqslant j\leqslant \min\{n+1-m,m-\ell\}} \binom{m}{j+\ell}\binom{n+1-m}{j} n^{-\ell-2j}. \tag{17}$$

From these relations, we obtain $\mu^*_{n,1} = 1$, and

$$\begin{aligned} \mu^*_{n,2} &= \frac{3n^2+n-1}{2n^2+2n-1},\\ \mu^*_{n,3} &= \frac{22n^6+40n^5-19n^4-42n^3+14n^2+15n-6}{(2n^2+2n-1)(6n^4+12n^3-7n^2-9n+6)}. \end{aligned} \tag{18}$$

In general, the $\mu^*_{n,m}$'s are all rational functions of n but their expressions become long as m increases. We thus turn to asymptotic approximation.

Asymptotic expansions for $\mu^*_{n,m}$. Our uniform asymptotic approximation to $\mu^*_{n,m}$ was largely motivated by intensive symbolic computations for small m. We briefly summarize them here, which will also be crucial in specifying the initial conditions for the differential equations satisfied by the functions $(\phi_1, \phi_2, \dots)$ involved in the full asymptotic expansion of $\mu^*_{n,m}$; see (86).

Starting from the closed-form expressions (18), we readily obtain $\mu^*_{n,0} = 0$, $\mu^*_{n,1} = 1$, and

$$\begin{aligned} \mu^*_{n,2} &= \tfrac{3}{2} - n^{-1} + \tfrac{5}{4}n^{-2} - \tfrac{7}{4}n^{-3} + \tfrac{19}{8}n^{-4} - \tfrac{13}{4}n^{-5} + O\big(n^{-6}\big),\\ \mu^*_{n,3} &= \tfrac{11}{6} - \tfrac{13}{6}n^{-1} + \tfrac{155}{36}n^{-2} - \tfrac{323}{36}n^{-3} + \tfrac{4007}{216}n^{-4} - \tfrac{2783}{72}n^{-5} + O\big(n^{-6}\big). \end{aligned}$$

Similarly, we have

$$\begin{aligned} \mu^*_{n,4} &= \tfrac{25}{12} - \tfrac{41}{12}n^{-1} + \tfrac{329}{36}n^{-2} - \tfrac{917}{36}n^{-3} + \tfrac{61841}{864}n^{-4} - \tfrac{19501}{96}n^{-5} + O\big(n^{-6}\big),\\ \mu^*_{n,5} &= \tfrac{137}{60} - \tfrac{283}{60}n^{-1} + \tfrac{2839}{180}n^{-2} - \tfrac{19859}{360}n^{-3} + \tfrac{848761}{4320}n^{-4} - \tfrac{5107063}{7200}n^{-5} + O\big(n^{-6}\big). \end{aligned}$$

From these expansions, we first observe that the leading sequence is exactly H_m ($H_0 := 0$)

$$\{H_m\}_{m\geqslant 0} = \left\{0, 1, \tfrac{3}{2}, \tfrac{11}{6}, \tfrac{25}{12}, \tfrac{137}{60}, \tfrac{49}{20}, \cdots\right\}.$$

An Ansatz for small m. These also suggest the Ansatz

$$\mu^*_{n,m} \approx \sum_{k\geqslant 0} \frac{d_k(m)}{n^k},$$

for some functions $d_k(m)$ of m. Using this form and the above expansions to match the undetermined coefficients of the polynomials (in m), we obtain successively

$$\begin{aligned}
d_0(m) &= H_m \quad (m \geqslant 0),\\
d_1(m) &= H_m + \tfrac{1}{2} - \tfrac{3}{2}m \quad (m \geqslant 1),\\
d_2(m) &= \tfrac{2}{3}H_m + \tfrac{1}{12} - \tfrac{7}{4}m + \tfrac{11}{12}m^2 \quad (m \geqslant 2),\\
d_3(m) &= \tfrac{1}{2}H_m + \tfrac{7}{24} - \tfrac{575}{432}m + \tfrac{23}{18}m^2 - \tfrac{283}{432}m^3, \quad (m \geqslant 2),\\
d_4(m) &= \tfrac{5}{18}H_m - \tfrac{59}{720} - \tfrac{3439}{3456}m + \tfrac{15101}{11520}m^2 - \tfrac{19951}{17280}m^3 + \tfrac{5759}{11520}m^4, \quad (m \geqslant 4).
\end{aligned}$$

So we observe the general pattern

$$\mu^*_{n,m} \approx \sum_{k\geqslant 0} \frac{1}{n^k}\left(b_k H_m + \sum_{0\leqslant j\leqslant k} \varpi_{k,j} m^j\right),$$

for some explicitly computable sequence b_k and coefficients $\varpi_{k,j}$. A crucial complication arises here: the general form of each $d_k(m)$ holds only for $m \geqslant 2\lfloor \frac{k}{2} \rfloor$, and correction terms are needed for smaller m. For example,

$$\begin{aligned}
d_1(m) &= H_m + \tfrac{1}{2} - \tfrac{3}{2}m - \tfrac{1}{2}[\![m=0]\!], \quad (m \geqslant 0)\\
d_2(m) &= \tfrac{2}{3}H_m + \tfrac{1}{12} - \tfrac{7}{4}m + \tfrac{11}{12}m^2 - \tfrac{1}{12}[\![m=0]\!] + \tfrac{1}{12}[\![m=1]\!], \quad (m \geqslant 0),
\end{aligned}$$

where we use the Iverson bracket notation $[\![A]\!] = 1$ if A holds, and 0, otherwise. It is such a complication that makes the determination of smaller-order terms more involved.

An Ansatz for large m. All the expansions here hold only for small m. When m grows, we write $m = \alpha n$ and see that

$$n^{-k} \sum_{0\leqslant j\leqslant k} \varpi_{k,j} m^j = \varpi_{k,k}\alpha^k + \varpi_{k,k-1}\frac{\alpha^{k-1}}{n} + \text{smaller order terms},$$

and it is exactly this form that motivated naturally our choice of the Ansatz

$$\mu^*_{n,m} \sim H_m + \phi(\alpha), \tag{19}$$

for some function $\phi(\alpha)$. This will be seen to be equivalent to the approximation $\mu_{n,m} \sim en(H_m + \phi(\alpha))$. Note that the omnipresence of the harmonic numbers H_m may be traced to the asymptotic estimate (10); see also Lemma 4.

Formal calculations. The next formal question then is how to guess this function ϕ (before proving all assumptions)? Here is the quick sketch of our ideas.

Substituting formally (19) into (16) using the expansions $H_m - H_{m-\ell} \sim \frac{\ell}{m}$ (see Corollary 6) and $\phi(\frac{m}{n}) - \phi(\frac{m-\ell}{n}) \sim \phi'(\alpha)\frac{\ell}{n}$ (see Lemma 5), we expect that

$$\begin{aligned}
\frac{1}{n} = \sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell}\left(\mu^*_{n,m} - \mu^*_{n,m-\ell}\right) &\approx \sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell}\left(\frac{\ell}{m} + \phi'(\alpha)\frac{\ell}{n}\right)\\
&= \frac{1}{n}\left(\frac{1}{\alpha} + \phi'(\alpha)\right)\sum_{1\leqslant \ell\leqslant m} \ell\lambda^*_{n,m,\ell}.
\end{aligned} \tag{20}$$

We are then led to the study of sums of the form $\sum_{1\leqslant\ell\leqslant m} a_\ell \lambda^*_{n,m,\ell}$ for a given sequence a_ℓ. As we will see in the next section $\sum_{1\leqslant\ell\leqslant m} \ell\lambda^*_{n,m,\ell} \sim S_1(\alpha)$ (defined in the introduction), so that ϕ has to satisfy the differential equation

$$\phi'(z) = \frac{1}{S_1(z)} - \frac{1}{z}.$$

Choosing properly the initial condition, we then conclude that $\phi = \phi_1$, as defined in (3). The justification of all these estimates, as well as more refined expansions, turns out to be highly nontrivial and requires several asymptotic tools that will be developed in the following two sections.

3 Asymptotics of sums of the form $\sum_{1\leqslant\ell\leqslant m} a_\ell \lambda^*_{n,m,\ell}$

Sums of the form

$$A^*_{n,m} := \sum_{1\leqslant\ell\leqslant m} a_\ell \lambda^*_{n,m,\ell}$$

will appear frequently in our analysis. We thus digress in this section to develop tools for deriving the asymptotic behaviors of such sums. We consider first general a_ℓ and then specialize the discussion to the cases when $a_\ell = \ell^r$ for $r \in \mathbb{Z}^+$.

Throughout this paper, we use the abbreviation

$$\alpha := \frac{m}{n}.$$

3.1 Asymptotics of $A^*_{n,m}$

Observe that, by (10), we see that most contribution to $A^*_{n,m}$ comes from small ℓ, say $\ell = o(m)$, provided that a_ℓ does not grow too fast. Indeed, we expect more precisely that

$$\begin{aligned} A^*_{n,m} &= \sum_{j\geqslant 0} \binom{n-m+1}{j} n^{-j} \sum_{j<\ell\leqslant m} a_{\ell-j} \binom{m}{\ell} n^{-\ell} \\ &= \sum_{j\geqslant 0} \frac{(n-m+1)(n-m)\cdots(n-m-j+2)}{j!\,n^j} \sum_{j<\ell\leqslant m} \frac{a_{\ell-j} m \cdots (m-\ell+1)}{\ell!\,n^\ell} \\ &= \sum_{j\geqslant 0} \frac{1}{j!} \prod_{0\leqslant k<j} \left(1-\alpha-\frac{k-1}{n}\right) \sum_{j<\ell\leqslant m} \frac{a_{\ell-j}}{\ell!} \prod_{0\leqslant k<\ell} \left(\alpha - \frac{k}{n}\right) \\ &\approx \sum_{j\geqslant 0} \frac{(1-\alpha)^j}{j!} \sum_{\ell>j} a_{\ell-j} \frac{\alpha^\ell}{\ell!}. \end{aligned} \tag{21}$$

The last step can be justified by bounding all errors involved, but the calculations become messy, especially when one needs more terms in the expansion. We use instead a more elegant approach via generating functions.

Lemma 3. *Let $\{a_\ell\}_{\ell\geqslant 1}$ be a given sequence whose generating function $A(z) = \sum_{\ell\geqslant 1} a_\ell z^{\ell-1}$ has a nonzero radius of convergence in the complex z-plane. Then*

$$A^*_{n,m} = \tilde{A}_0(\alpha) + \frac{\tilde{A}_1(\alpha)}{n} + O\left(\alpha n^{-2}\right), \tag{22}$$

where $\tilde{A}_0(\alpha)$ *and* $\tilde{A}_1(\alpha)$ *are entire functions of* α *defined by*

$$\tilde{A}_0(\alpha) := \sum_{\ell \geqslant 1} \frac{\alpha^\ell}{\ell!} \sum_{0 \leqslant j < \ell} a_{\ell-j} \frac{(1-\alpha)^j}{j!}, \tag{23}$$

and ($a_0 := 0$)

$$\tilde{A}_1(\alpha) := -\frac{1}{2} \sum_{\ell \geqslant 1} \frac{\alpha^\ell}{\ell!} \sum_{0 \leqslant j < \ell} \frac{(1-\alpha)^j}{j!} \big((\ell - j) a_{\ell+1-j} - (\ell + 2 - j) a_{\ell-1-j} + a_{\ell-j} \big). \tag{24}$$

Proof. Observe that the sum on the left-hand side of (17) is itself a Cauchy product, namely,

$$\lambda^*_{n,m,\ell} = \sum_{0 \leqslant j \leqslant m-\ell} \binom{m}{m-\ell-j} n^{-\ell-j} \times \binom{n+1-m}{j} n^{-j}.$$

Let $[z^n]f(z)$ denote the coefficients of z^n in the Taylor expansion of $f(z)$. Then the right-hand side equals

$$[z^{m-\ell}]\Big(z + \frac{1}{n}\Big)^m \Big(1 + \frac{z}{n}\Big)^{n+1-m} = [z^{-\ell}]\Big(1 + \frac{1}{nz}\Big)^m \Big(1 + \frac{z}{n}\Big)^{n+1-m}.$$

Our analytic proof then starts from the relation (Cauchy's integral representation)

$$\lambda^*_{n,m,\ell} = \frac{1}{2\pi i} \oint_{|z|=c} z^{\ell-1} \Big(1 + \frac{1}{nz}\Big)^m \Big(1 + \frac{z}{n}\Big)^{n+1-m} \mathrm{d}z, \tag{25}$$

where $c > 0$. The relation (25) holds *a priori* for $1 \leqslant \ell \leqslant m$, but the right-hand side becomes zero for $\ell > m$. It follows, by multiplying both sides by a_ℓ and summing over $1 \leqslant \ell \leqslant m$ for the left-hand side and over all $\ell \geqslant 1$ for the right-hand side, that

$$A^*_{n,m} = \frac{1}{2\pi i} \oint_{|z|=c} A(z) \Big(1 + \frac{1}{nz}\Big)^m \Big(1 + \frac{z}{n}\Big)^{n+1-m} \mathrm{d}z,$$

where $0 < c < \varrho$, ϱ being the radius of convergence of A. By the asymptotic expansion

$$\begin{aligned}
\Big(1 + \frac{1}{nz}\Big)^m \Big(1 + \frac{z}{n}\Big)^{n+1-m} &= \exp\Big(\alpha n \log\Big(1 + \frac{1}{nz}\Big) + (1-\alpha) n \log\Big(1 + \frac{z}{n}\Big)\Big) \Big(1 + \frac{z}{n}\Big) \\
&= e^{\frac{\alpha}{z} + (1-\alpha)z} \Big(1 - \frac{1}{2n}\Big((1-\alpha)z^2 - 2z + \frac{\alpha}{z^2}\Big) + O\Big(\frac{(1-\alpha)^2 |z|^4 + 1 + |z|^{-4}}{n^2}\Big)\Big),
\end{aligned}$$

for large n and bounded $|z|$ and $\frac{1}{|z|}$, where the O-term holds uniformly for z on the integration path, and the integral representations

$$\begin{aligned}
\tilde{A}_0(\alpha) &= \frac{1}{2\pi i} \oint_{|z|=c} A(z) e^{\frac{\alpha}{z} + (1-\alpha)z} \mathrm{d}z, \\
\tilde{A}_1(\alpha) &= -\frac{1}{4\pi i} \oint_{|z|=c} A(z) \Big((1-\alpha)z^2 - 2z + \frac{\alpha}{z^2}\Big) e^{\frac{\alpha}{z} + (1-\alpha)z} \mathrm{d}z,
\end{aligned} \tag{26}$$

we deduce (22). The expression (23) is then obtained by first expanding $e^{\frac{\alpha}{z}}$ (in decreasing powers of z) and then by integrating term-by-term:

$$\tilde{A}_0(\alpha) = \sum_{\ell \geqslant 0} \frac{\alpha^\ell}{\ell!} \cdot \frac{1}{2\pi i} \oint_{|z|=c} z^{-\ell-1} A(z) e^{(1-\alpha)z} \mathrm{d}z = \sum_{\ell \geqslant 0} \frac{\alpha^\ell}{\ell!} \cdot [z^\ell] A(z) e^{(1-\alpha)z}.$$

Note that $a_0 = 0$. For (24), we apply an integration by parts using the relation

$$\frac{\mathrm{d}}{\mathrm{d}z}\, e^{\frac{\alpha}{z}+(1-\alpha)z} = \left(-\frac{\alpha}{z^2}+1-\alpha\right) e^{\frac{\alpha}{z}+(1-\alpha)z},$$

and the decomposition

$$(1-\alpha)z^2 - 2z + \frac{\alpha}{z^2} = -(1-z^2)\left(-\frac{\alpha}{z^2}+1-\alpha\right)+1+2z,$$

giving

$$\tilde{A}_1(\alpha) = -\frac{1}{4\pi i}\oint_{|z|=c}\Big((1-z^2)A'(z)+(1-4z)A(z)\Big)\, e^{\frac{\alpha}{z}+(1-\alpha)z}\,\mathrm{d}z. \tag{27}$$

Substituting the series expansion $A(z) = \sum_{\ell\geqslant 1} a_\ell z^{\ell-1}$ and then integrating term by term, we get (24). ∎

When α tends to the two boundaries 0 and 1, we have

$$A^*_{n,m} \sim \tilde{A}_0(\alpha) \sim \begin{cases} \dfrac{a_k}{k!}\,\alpha^k, & \text{as } \alpha \to 0^+,\\ \displaystyle\sum_{\ell\geqslant 1}\frac{a_\ell}{\ell!}, & \text{as } \alpha \to 1^-, \end{cases}$$

where k is the smallest integer such that $a_k \neq 0$.

3.2 Asymptotics of $\sum_{1\leqslant \ell\leqslant m} \ell^r \lambda^*_{n,m,\ell}$

We now discuss special sums of the form (when $a_\ell = \ell^r$) and define

$$\Lambda^{(r)}_{n,m} := \sum_{1\leqslant \ell\leqslant m} \ell^r \lambda_{n,m,\ell} \qquad (r \geqslant 0),$$

which will be repeatedly encountered below. Define also $\Lambda^{*(r)}_{n,m} := \sum_{1\leqslant \ell\leqslant m} \ell^r\lambda^*_{n,m,\ell}$, so that $\Lambda^{(r)}_{n,m} = e_n \Lambda^{*(r)}_{n-1,m}$. For convenience, we write

$$\Lambda^*_{n,m} := \Lambda^{*(0)}_{n,m} = \sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell}. \tag{28}$$

Let I_k denote the modified Bessel functions

$$I_k(2z) := \sum_{j\geqslant 0} \frac{z^{2j+k}}{j!(j+k)!} \qquad (k \in \mathbb{Z}).$$

We now show that $\tilde{A}_0(\alpha) = S_r(\alpha)$ and $\tilde{A}_1(\alpha)$ in (22) can be expressed in terms of linear combination of S_r and I_k.

Corollary 3. *Uniformly for* $1 \leqslant m \leqslant n$

$$\Lambda^{*(r)}_{n,m} = S_r(\alpha) + \frac{U_r(\alpha)}{n} + O\left(\alpha n^{-2}\right), \tag{29}$$

for $r = 0, 1, \dots,$ *where both* S_r *and* U_r *are entire functions given by*

$$S_r(z) = \sum_{\ell\geqslant 1}\frac{z^\ell}{\ell!}\sum_{0\leqslant j<\ell}(\ell-j)^r\,\frac{(1-z)^j}{j!},$$

and

$$U_r(\alpha) = \begin{cases} \dfrac{S_0(\alpha)}{2} - \dfrac{3}{2}\sqrt{\dfrac{\alpha}{1-\alpha}}\, I_1\left(2\sqrt{\alpha(1-\alpha)}\right), & \text{if } r=0\\ -\dfrac{1}{2}\left((2r-1)S_r(\alpha) + \displaystyle\sum_{0\leqslant j<r}\binom{r}{j}\frac{j-(-1)^{r-j}(2r+2-3j)}{r+1-j}S_j(\alpha)\right), & \text{if } r\geqslant 1. \end{cases} \tag{30}$$

In particular,

$$\begin{aligned} U_1(\alpha) &= -S_0(\alpha) - \tfrac{1}{2}S_1(\alpha)\\ U_2(\alpha) &= \quad S_0(\alpha) - 2S_1(\alpha) - \tfrac{3}{2}S_2(\alpha)\\ U_3(\alpha) &= -S_0(\alpha) + 2S_1(\alpha) - 3S_2(\alpha) - \tfrac{5}{2}S_3(\alpha). \end{aligned} \tag{31}$$

These are sufficient for our uses.

Proof. By applying (22), (23), and (26) with $a_\ell = \ell^r$, we see that

$$\tilde{A}_0(\alpha) = S_r(\alpha) = \frac{1}{2\pi i}\oint_{|z|=c} E_r(z)e^{\frac{\alpha}{z}+(1-\alpha)z}\,\mathrm{d}z \qquad (r=0,1,\dots), \tag{32}$$

where $E_r(z) := \sum_{\ell\geqslant 1}\ell^r z^{\ell-1}$. It remains to simplify $\tilde{A}_1(\alpha)$. To that purpose, we start with the integral representation (see (27))

$$U_r(\alpha) = -\frac{1}{4\pi i}\oint_{|z|=c}\left((1-z^2)E_r'(z) + (1-4z)E_r(z)\right)e^{\frac{\alpha}{z}+(1-\alpha)z}\,\mathrm{d}z, \tag{33}$$

When $r=0$, we have $E_0(z) = (1-z)^{-1}$. Thus

$$\begin{aligned} U_0(\alpha) &= \frac{1}{2\pi i}\oint_{|z|=c}\left(\frac{1}{2(1-z)} - \frac{3}{2}\right)e^{\frac{\alpha}{z}+(1-\alpha)z}\,\mathrm{d}z\\ &= \frac{S_0(\alpha)}{2} - \frac{3}{2}\sum_{\ell\geqslant 1}\frac{\alpha^\ell(1-\alpha)^{\ell-1}}{\ell!(\ell-1)!}, \end{aligned}$$

which proves (30) for $r=0$. For $r\geqslant 1$, we have

$$\begin{aligned} &(1-z^2)E_r'(z) + (1-4z)E_r(z)\\ &\quad= (1-z^2)\sum_{\ell\geqslant 2}\ell^r(\ell-1)z^{\ell-2} + (1-4z)\sum_{\ell\geqslant 1}\ell^r z^{\ell-1}\\ &\quad= \sum_{\ell\geqslant 1}\ell(\ell+1)^r z^{\ell-1} - \sum_{\ell\geqslant 2}(\ell+2)(\ell-1)^r z^{\ell-1} + E_r(z)\\ &\quad= \sum_{0\leqslant j\leqslant r}\binom{r}{j}E_{j+1}(z) - \sum_{0\leqslant j\leqslant r}\binom{r}{j}(-1)^{r-j}\left(E_{j+1}(z) + 2E_j(z)\right) + E_r(z). \end{aligned}$$

From this and the relation (33), we obtain (30). Note that the coefficient of $E_{r+1}(z)$ is zero. ∎

The Corollary implies specially that

$$\Lambda_{n,m}^{(r)} = e^{-1}S_r(\alpha)\left(1 + O\left(n^{-1}\right)\right), \tag{34}$$

uniformly for $1 \leqslant m \leqslant n$. Since $S_r(z) = z + O(|z|^2)$ as $|z| \to 0$, we have the uniform bound

$$\Lambda_{n,m}^{*\,(r)} \asymp S_r(\alpha) \asymp \alpha \qquad (1 \leqslant m \leqslant n), \tag{35}$$

meaning that the ratio of $\frac{\Lambda_{n,m}^{*\,(r)}}{\alpha}$ remains bounded away from zero and infinity for all m in the specified range.

The following expansions for $S_r(z)$ and $U_r(z)$ as $z \to 0$ will be used later

$$\begin{aligned} S_r(z) &= z + \frac{2^r + 1}{2} z^2 + O(z^3), \\ U_r(z) &= -\frac{2^r + 1}{2} z + O(z^2), \end{aligned} \tag{36}$$

for $r = 0, 1, \dots$. See also Appendix A for other properties of $S_r(\alpha)$.

4 The expected values of $X_{n,m}$ and their asymptotics

We will derive in this section a more precise expansion for the mean $\mu_{n,m} := \mathbb{E}(X_{n,m})$.

Theorem 2. *The expected optimization time of the* $(1+1)$*-EA when starting with* $n - m$ 1*s satisfies the asymptotic approximation*

$$\frac{\mathbb{E}(X_{n,m})}{en} = H_m + \phi_1(\alpha) + \frac{H_m - \phi_1(\alpha) + 2\phi_2(\alpha) + 2\alpha\phi_1'(\alpha)}{2n} + O\left(n^{-2} H_m\right), \tag{37}$$

uniformly for $1 \leqslant m \leqslant n$*, where* ϕ_1 *is defined in* (3) *and* ϕ_2 *is an analytic function defined by*

$$\phi_2(\alpha) = \frac{1}{2} - \int_0^\alpha \left(\frac{S_2(x) S_1'(x)}{2S_1(x)^3} - \frac{S_0(x)}{S_1(x)^2} - \frac{1}{2S_1(x)} - \frac{1}{2x^2} + \frac{1}{x} \right) \mathrm{d}x, \tag{38}$$

the integrand having a removable singularity at $x = 0$.

As discussed above, we consider for simplicity $\mu_{n,m}^* := \frac{e_n}{n} \mu_{n+1,m}$, and we will prove the slightly simpler expansion

$$\mu_{n,m}^* = H_m + \phi_1(\alpha) + \frac{H_m + \phi_2(\alpha)}{n} + O\left(n^{-2} H_m\right), \tag{39}$$

for $1 \leqslant m \leqslant n$, which is identical to (37) by using the relation $\mu_{n,m} = \frac{n-1}{e_{n-1}} \mu_{n-1,m}^*$ and the asymptotic expansions

$$\phi_1\left(\frac{m}{n-1}\right) = \phi_1(\alpha) + \frac{\alpha \phi_1'(\alpha)}{n} + O\big(n^{-2}\big),$$

and

$$e^{-1}\left(1 - \frac{1}{n}\right)^{-n} = 1 + \frac{1}{2n} + O\big(n^{-2}\big).$$

See Figure 4 for graphic renderings. More figures are collected in Appendix C.

Our analysis will be based on the recurrence (16) for $\mu_{n,m}^*$ and use the idea of successive asymptotic iteration (or bootstrapping; see de Bruijn (1981) or Flajolet and Sedgewick (2009)), which proceeds as follows. We consider first the difference $\mu_{n,m}^* - H_m - \phi_1(\alpha)$, which satisfies itself a recurrence of the same type but with a different non-homogeneous part. We bound this difference by Lemma 4 and a transfer technique, which deduces a uniform bound for the difference from that of the non-homogeneous part. Then we repeat the same procedure by

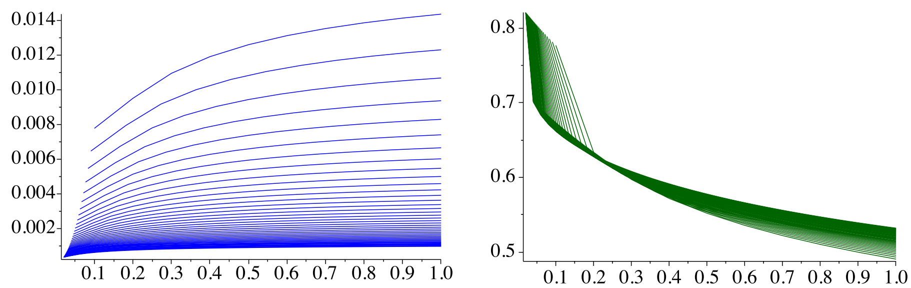


Figure 4: *Left: the differences* $\mu^*_{n,m} - (H_m + \phi_1(\alpha) + \frac{H_m+\phi_2(\alpha)}{n})$ *for* $1 \leqslant m \leqslant n$ *(normalized to the unit interval) and* $n = 10, \ldots, 50$ *(in top-down order); Right: the normalized differences* $\frac{n^2}{H_m}(\mu^*_{n,m} - (H_m + \phi_1(\alpha) + \frac{H_m+\phi_2(\alpha)}{n}))$ *for* $n = 10, \ldots, 50$ *(in bottom-up order when viewing from the point* $(1, 0)$.

subtracting more terms and get a refined expansion. This same procedure can then be extended and yields a more precise expansion; see Appendix D.

Instead of starting from a state with a fixed number of 1s, the first step of the Algorithm $(1+1)$-EA described in the introduction corresponds to the situation when the initial state $f(\mathbf{x})$ (the number of 1s) is not fixed but random. Assume that this input follows a binomial distribution of parameter $1-\rho \in (0,1)$ (each bit being 1 with probability $1-\rho$ and 0 with probability ρ). Denote by X_n the number of steps taken by $(1+1)$-EA to reach the optimum state. The following result describes precisely the asymptotic behavior of the expected optimization time.

Theorem 3. *The expected value of* X_n *satisfies*

$$\frac{\mathbb{E}(X_n)}{en} = \log \rho n + \gamma + \phi_1(\rho) + \frac{\log \rho n + \gamma + c_3}{2n} + O\left(\frac{\log n}{n^2}\right),$$

where $c_3 := 1 - \phi_1(\rho) + 2\rho\phi_1'(\rho) + \rho(1-\rho)\phi_1''(\rho) + 2\phi_2(\rho)$.

Note that $e(\log \rho + \gamma + \phi_1(\rho))$ is an increasing function of ρ, which is consistent with the intuition that it takes less steps to reach the final state if we start with more 1s (small ρ means $1-\rho$ closer to 1, or 1 occurring with higher probability). Also

$$1 + 2\rho\phi_1'(\rho) + \rho(1-\rho)\phi_1''(\rho) = -2 + \frac{1}{\rho} + \frac{2\rho}{S_1(\rho)} - \rho(1-\rho)\frac{S_1'(\rho)}{S_1(\rho)^2}.$$

The constant c_2 in (2) can now be computed and has the value ($\rho = \frac{1}{2}$)

$$c_2 = \frac{e}{2}\left(-\log 2 + \gamma - \phi_1(\tfrac{1}{2}) + 2\phi_2(\tfrac{1}{2}) + \frac{1}{S_1(\frac{1}{2})} - \frac{S_1'(\frac{1}{2})}{4S_1(\frac{1}{2})^2}\right) \approx 0.59789875\ldots. \tag{40}$$

Numerically, to compute the value of $\phi_1(\alpha)$ for $\alpha \in (0,1]$, the most natural way consists in using the Taylor expansion

$$\frac{1}{S_1(x)} - \frac{1}{x} = \sum_{j\geqslant 0} \sigma_j x^j,$$

and after a term-by-term integration

$$\phi_1(\alpha) = \sum_{j\geqslant 0} \frac{\sigma_j}{j+1}\alpha^{j+1}.$$

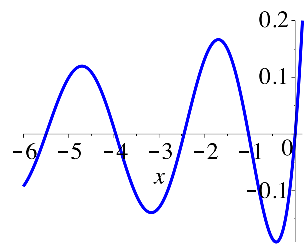


Figure 5: $S_1(x)$ *has an infinite number of zeros on* $\mathbb{R}^-$.

While $S_1(x)$ is an entire function with rapidly decreasing coefficients, such an expansion converges slowly when α is close to 1, the main reason being that the smallest nonzero $|x|$ for which $S_1(x) = 0$ occurs when $x \approx -1.0288$, implying that the radius of convergence of this series is only slightly larger than unity. Note that $S_1(0) = 0$ but the simple pole is removed by subtracting $\frac{1}{x}$. A better idea is then expanding $\frac{1}{S_1(x)} - \frac{1}{x}$ at $x = 1$ and then integrating term-by-term, giving

$$\phi_1(\alpha) = \sum_{j\geqslant 0} \frac{\sigma_j'}{j+1}\left(1-(1-\alpha)^{j+1}\right) \quad \text{where} \quad \frac{1}{S_1(1-x)} - \frac{1}{1-x} = \sum_{j\geqslant 0} \sigma_j' x^j.$$

This expansion is numerically more efficient and stable because of better convergence when $\alpha \in [0, 1]$. The same technique also applies to the calculation of ϕ_2 and other functions in this paper.

A direct consequence of the precise estimates we derived is the following asymptotic approximation measuring the difference between $\mathbb{E}(X_n)$ (random input) and $\mathbb{E}(X_{n,\rho n+o(n)})$ (fixed input), which improves the $O(1)$-bound for $\rho = \frac{1}{2}$ derived in the recent paper Doerr and Doerr (2014).

Corollary 4. *The difference between $\mathbb{E}(X_n)$ and $\mathbb{E}(X_{n,\rho n+o(n)})$ satisfies*

$$\begin{aligned}\mathbb{E}(X_n) - \mu_{n,\rho n-\vartheta_n} &= \frac{e}{2\rho}\left(2(1+\rho\phi_1'(\rho))\vartheta_n + \rho^2(1-\rho)\phi_1''(\rho) - 1 + \rho\right) \\ &\quad + \frac{e}{2n}\left(\tau_2\vartheta_n^2 + \tau_1\vartheta_n + \tau_0\right) + O\left(\frac{|\vartheta_n|^3+1}{n^2}\right),\end{aligned} \tag{41}$$

uniformly for $\vartheta_n = o(n)$, where

$$\tau_2 := \frac{1-\rho^2\phi_1''(\rho)}{\rho^2}, \quad \tau_1 := \frac{2\rho^3\phi_1''(\rho) + \rho^2\phi_1'(\rho) + 2\rho^2\phi_2'(\rho) - 1 + \rho}{\rho^2}$$

$$\tau_0 = (1-\rho)\left(\frac{\rho^2(1+\rho)\phi_1^{(4)}(\rho)}{4} + \frac{\rho(1+\rho)\phi_1'''(\rho)}{3} + \frac{3\rho\phi_1''(\rho)}{2} + \rho\phi_2''(\rho) + \frac{1-2\rho}{6\rho^2}\right).$$

Note that the dominant term on the right-hand side is bounded when ϑ_n is so. Also the coefficients here can be completely written in terms of the $S_r(\rho)$'s; for example

$$\frac{1}{2\rho}\left(2(1+\rho\phi_1'(\rho))\vartheta_n + \rho^2(1-\rho)\phi_1''(\rho) - 1 + \rho\right) = \frac{\vartheta_n}{S_1(\rho)} - \frac{\rho(1-\rho)S_1'(\rho)}{2S_1(\rho)^2}.$$

Since the proof is straightforward either from the expansions in Theorems 2 and 3 or by the same method of proof of Theorem 3, we omit the details, which can be readily manipulated by standard symbolic computation tools.

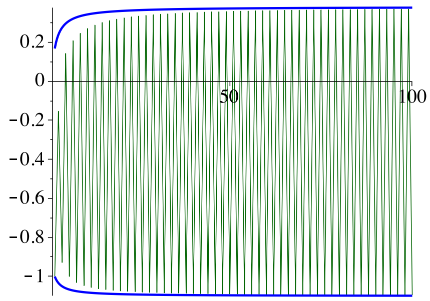
4.1 More asymptotic tools

We develop here some other asymptotic tools that will be used in proving Theorem 2.

The following lemma is very helpful in obtaining error estimates to be addressed below. It also sheds new light on the occurrence of the harmonic numbers H_m in (39).

Lemma 4. *Consider the recurrence*

$$\sum_{1\leqslant \ell\leqslant m} \lambda_{n,m,\ell}^*(a_{n,m} - a_{n,m-\ell}) = b_{n,m} \qquad (m \geqslant 1),$$



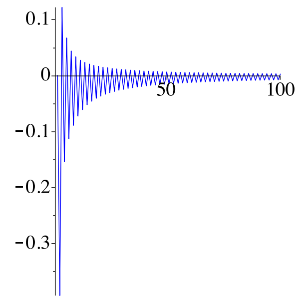


Figure 6: *The differences* $\mathbb{E}(X_n) - \mathbb{E}(X_{n,\lfloor \frac{n}{2} \rfloor})$ *when* $\rho = \frac{1}{2}$ *for* $n = 2, \dots, 100$ *(left) and the* O*-term of* (41) *(right). The blue curves (left) correspond to the right-hand side of* (41) *without the* O*-term. The periodic fluctuations come from* $\vartheta_n = \{\frac{n}{2}\}$*, where* $\{x\}$ *denotes the fractional part of* x*. Numerically, when* n *is even (odd), the first term on the right-hand side is approximately* -1.00875 *(*0.47564*). Note here that* $X_{n,\lfloor \frac{n}{2} \rfloor}$ *starts with* $\lceil \frac{n}{2} \rceil$ *1s.*

where $b_{n,m}$ *is defined for* $1 \leqslant m \leqslant n$ *and* $n \geqslant 1$*. Assume that* $|a_{n,0}| \leqslant d$ *for* $n \geqslant 1$*, where* $d \geqslant 0$*. If* $|b_{n,m}| \leqslant \frac{c}{n}$ *holds uniformly for* $1 \leqslant m \leqslant n$ *and* $n \geqslant 1$*, where* $c > 0$*, then*

$$|a_{n,m}| \leqslant cH_m + d \qquad (0 \leqslant m \leqslant n).$$

Proof. The result is true for $m = 0$. For $m \geqslant 1$, we start from the simple inequality

$$\Lambda^*_{n,m} = \sum_{1 \leqslant \ell \leqslant m} \lambda^*_{n,m,\ell} \geqslant \frac{m}{n} \qquad (1 \leqslant m \leqslant n),$$

because all terms in the sum expression (17) are positive and taking only one term ($j = 0$ and $\ell = 1$) gives the lower bound. Then, by the induction hypothesis,

$$\begin{aligned} |a_{n,m}| &\leqslant \frac{|b_{n,m}|}{\Lambda^*_{n,m}} + |a_{n,m-1}| \\ &\leqslant \frac{c}{n} \cdot \frac{n}{m} + cH_{m-1} + d \\ &= cH_m + d, \end{aligned}$$

proving the lemma. ∎

Applying this lemma to the recurrence (16), we then get a simple upper bound for $\mu^*_{n,m}$.

Corollary 5. *For* $0 \leqslant m \leqslant n$*, the inequality* $\mu^*_{n,m} \leqslant H_m$ *holds.*

Lemma 5. *If* ϕ *is a* $C^2[0,1]$*-function (twice continuously differentiable in the unit interval), then*

$$\sum_{1 \leqslant \ell \leqslant m} \lambda^*_{n,m,\ell} \left(\phi\left(\frac{m}{n}\right) - \phi\left(\frac{m-\ell}{n}\right) \right) = \frac{\phi'(\alpha)}{n} \sum_{1 \leqslant \ell \leqslant m} \ell \lambda^*_{n,m,\ell} + O\big(n^{-2}\big),$$

uniformly for $1 \leqslant m \leqslant n$*.*

Uniformity of the estimate in the lemma and the asymptotic expansion (29) play a crucial rôle in our analysis.

Proof. A direct Taylor expansion with remainder gives

$$\phi(\alpha) - \phi\left(\alpha - \tfrac{\ell}{n}\right) = \phi'(\alpha)\tfrac{\ell}{n} + O\left(\ell^2 n^{-2}\right),$$

uniformly for $1 \leqslant \ell \leqslant m$, since $\phi''(t) = O(1)$ for $t \in [0,1]$. The lemma follows from the estimates (29). ∎

The approximation can be easily extended and refined if more smoothness properties of ϕ are known, which is the case for all functions appearing in our analysis (they are all $C^\infty[0,1]$).

Another standard technique we need is Stirling's formula for the factorials

$$\log n! = \log\Gamma(n+1) = \left(n+\tfrac{1}{2}\right)\log n - n + \tfrac{1}{2}\log(2\pi) + \tfrac{1}{12}n^{-1} + O\left(n^{-3}\right), \tag{42}$$

where Γ denotes Euler's Gamma function.

4.2 Proof of Theorem 2

Our method of proof consists in three steps: first a heuristic calculation to get the dominant term, then an error analysis to justify the dominant term with an explicit error term, and finally another refined analysis (of the same inductive argument) to complete the proof of Theorem 2. The main idea of the error analysis is to express the error term as another recurrence of the same type but with a different non-homogeneous part. Then showing the smallness of the non-homogeneous part will then lead to the required order estimate for the error.

We start with the following identity whose proof is straightforward.

Lemma 6. *For a given sequence $f(k)$, let ∇ denote the backward difference operator $\nabla f(k) = f(k) - f(k-1)$. Then for $0 \leqslant \ell \leqslant m$*

$$f(m) + f(m-1) + \cdots + f(m-\ell+1) = \sum_{1\leqslant k\leqslant m} \binom{\ell}{k}(-1)^{k-1}\nabla^{k-1} f(m).$$

Note that the sum vanishes for $k > \ell$. Take $f(k) = \frac{1}{k}$. Then we obtain the following identity, which is itself an asymptotic expansion for large m (and $\ell = o(m)$).

Corollary 6. *For $m \geqslant 1$ and $0 \leqslant \ell \leqslant m$,*

$$H_m - H_{m-\ell} = \sum_{1\leqslant k\leqslant m} \frac{\ell(\ell-1)\cdots(\ell-k+1)}{km(m-1)\cdots(m-k+1)}.$$

Formal calculations. Assuming the validity of (19), we can make the formal calculations in (20) more precise by (29), (39), Lemma 5, and Corollary 6, obtaining

$$\frac{1}{n} = \sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell}\left(\mu^*_{n,m} - \mu^*_{n,m-\ell}\right) \sim \frac{1}{n}\left(\frac{1}{\alpha} + \phi'(\alpha)\right) S_1(\alpha).$$

Thus we see that ϕ satisfies

$$\phi'(z) = \frac{1}{S_1(z)} - \frac{1}{z}.$$

We now specify the initial condition $\phi(0)$. Since the postulated form (19) holds for $1 \leqslant m \leqslant n$ (indeed also true for $m = 0$), we take $m = 1$ and see that $\phi(0) = 0$ because $\mu^*_{n,1} = 1$. This implies that $\phi = \phi_1$. The first few terms in the Taylor expansion of ϕ_1 read as follows.

$$\phi_1(z) = -\tfrac{3}{2}z + \tfrac{11}{12}z^2 - \tfrac{283}{432}z^3 + \tfrac{5759}{11520}z^4 - \tfrac{57137}{144000}z^5 + \tfrac{2353751}{7257600}z^6 + \cdots, \tag{43}$$

which can then be checked with the explicit expressions of $\mu^*_{n,m}$ for small m (see Section 2.4).

Error analysis. To justify the form (19) (with $\phi = \phi_1$), we consider the difference

$$\Delta^*_{n,m} := \mu^*_{n,m} - H_m - \phi_1(\alpha),$$

which satisfies, by (16), the recurrence

$$\sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell}\left(\Delta^*_{n,m} - \Delta^*_{n,m-\ell}\right) = E_1(n,m), \tag{44}$$

where

$$E_1(n,m) := \frac{1}{n} - \sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell}\left(H_m - H_{m-\ell} + \phi_1(\alpha) - \phi_1\left(\alpha - \tfrac{\ell}{n}\right)\right).$$

We first show that $E_1 = O\big(n^{-2}\big)$, and this will imply, by Lemma 4, the estimate $\Delta^*_{n,m} = O(n^{-1}H_m)$.

By the asymptotic relation (29) with $r = 1$ and the definition of ϕ_1, we have

$$\frac{1}{n} = \sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell}\left(\tfrac{\ell}{m} + \phi_1'(\alpha)\tfrac{\ell}{n}\right) + O\big(n^{-2}\big),$$

and thus

$$E_1(n,m) = -\sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell}\left(H_m - H_{m-\ell} - \tfrac{\ell}{m} + \phi_1(\alpha) - \phi_1\left(\alpha - \tfrac{\ell}{n}\right) - \phi_1'(\alpha)\tfrac{\ell}{n}\right) + O\big(n^{-2}\big)$$

By Lemma 5, we see that

$$\sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell}\left(\phi_1(\alpha) - \phi_1\left(\alpha - \tfrac{\ell}{n}\right) - \phi_1'(\alpha)\tfrac{\ell}{n}\right) = O\big(n^{-2}\big),$$

uniformly for $1 \leqslant m \leqslant n$. On the other hand, we have the upper bounds (see Corollary 6)

$$H_m - H_{m-\ell} - \tfrac{\ell}{m} = \begin{cases} O\left(\ell^2 m^{-2}\right), & \text{if } \ell = o(m), \\ O(H_m), & \text{for } 1 \leqslant \ell \leqslant m. \end{cases}$$

Note that the first estimate is only uniform for $1 \leqslant \ell = o(m)$. When ℓ is close to m, say $m - \ell = O(m^{1-\varepsilon})$, where $\varepsilon \in (0,1)$, the left-hand side blows up with m but the right-hand side $O\left(\ell^2 m^{-2}\right)$ remains bounded. Thus we split the sum at $\lceil\sqrt{m}\rceil$ and then obtain ($H_m - H_{m-\ell} - \frac{\ell}{m} = 0$ when $\ell = 1$)

$$\begin{aligned} &\sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell}\left(H_m - H_{m-\ell} - \tfrac{\ell}{m}\right) \\ &\qquad = O\left(m^{-2}\sum_{2\leqslant \ell\leqslant \lceil\sqrt{m}\rceil} \ell^2\lambda^*_{n,m,\ell} + H_m \sum_{\lceil\sqrt{m}\rceil+1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell}\right). \end{aligned} \tag{45}$$

Now, by (17) and an analysis similar to that in (21),

$$\begin{aligned} m^{-2}\sum_{2\leqslant \ell\leqslant \lceil\sqrt{m}\rceil} \ell^2\lambda^*_{n,m,\ell} &= O\left(m^{-2}\sum_{j\geqslant 0}\frac{(1-\alpha)^j}{j!}\sum_{j+2\leqslant \ell\leqslant m}\frac{(j+\ell)^2\alpha^\ell}{\ell!}\right) \\ &= O\left(m^{-2}\alpha^2\right) = O\big(n^{-2}\big); \end{aligned} \tag{46}$$

similarly,

$$\begin{aligned} H_m \sum_{\sqrt{m}+1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell} &= O\left(H_m \sum_{j\geqslant 0} \frac{(1-\alpha)^j}{j!} \sum_{j+\sqrt{m}+1\leqslant \ell\leqslant m} \frac{\alpha^\ell}{\ell!}\right) \\ &= O\left(H_m \sum_{\ell\geqslant\sqrt{m}+1} \frac{\alpha^\ell}{\ell!}\right). \end{aligned}$$

Let $M := \lceil \sqrt{m}\rceil + 1$. Then

$$H_m \sum_{\ell\geqslant\sqrt{m}+1} \frac{\alpha^\ell}{\ell!} = H_m \frac{\alpha^M}{M!} \sum_{\ell\geqslant 0} \frac{\alpha^\ell}{(M+1)\cdots(M+\ell)} = O\left(\frac{H_m \alpha^{\sqrt{m}+1}}{\Gamma(\sqrt{m}+2)}\right). \tag{47}$$

By Stirling's formula (42), the last O-term is of order

$$\frac{m^{\frac14}\, H_m}{n}\, e^{-\sqrt{m}(\log n - \frac12 \log m - 1)} = O\big(n^{-2}\big),$$

for $m \geqslant 1$. Combining these estimates, we then obtain

$$E_1(n,m) = O\big(n^{-2}\big),$$

uniformly for $1 \leqslant m \leqslant n$. Thus $\Delta_{n,m} := n\,\Delta^*_{n,m}$ satisfies, by (44), a recurrence of the form

$$\sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell} \big(\Delta_{n,m} - \Delta_{n,m-\ell}\big) = O\big(n^{-1}\big) \qquad (1 \leqslant m \leqslant n),$$

with $\Delta_{n,0} = 0$. It follows, by applying Lemma 4, that $\Delta_{n,m} = O(H_m)$, and we conclude that, uniformly for $0 \leqslant m \leqslant n$,

$$\mu^*_{n,m} = H_m + \phi_1(\alpha) + O\left(n^{-1} H_m\right).$$

This proves the first two terms of the asymptotic approximation to $\mu^*_{n,m}$ in (39). The more refined expansion is obtained by refining the same calculations and justification the main steps of which are carried out in Appendix B.

We summarize this proof by the following "asymptotic transfer" for the recurrence

$$\sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell} (a_{n,m} - a_{n,m-\ell}) = b_{n,m},$$

where $b_{n,m}$ is a given sequence:

$$1 \leqslant m \leqslant n : \left\{ \begin{array}{rcl} b_{n,m} = n^{-1} & \Longrightarrow & a_{n,m} \sim H_m + \phi_1(\alpha) \\ b_{n,m} = O\big(n^{-2}\big) & \Longrightarrow & a_{n,m} = O(n^{-2} H_m). \end{array}\right.$$

4.3 Proof of Theorem 3

We now prove Theorem 3 concerning the asymptotics of $\mathbb{E}(X_n)$, which, by linearity of expectation, satisfies

$$\mathbb{E}(X_n) = \sum_{0\leqslant m\leqslant n} \pi_{n,m}\, \mu_{n,m} = \frac{n-1}{\left(1-\frac{1}{n}\right)^n} \sum_{0\leqslant m\leqslant n} \pi_{n,m} \mu^*_{n-1,m},$$

where $\pi_{n,m} := \binom{n}{m}\rho^m\hat{\rho}^{n-m}$ ($\hat{\rho} := 1-\rho$). Roughly, since the binomial distribution is highly centered around its mean value ρn, we expect that the right-hand side behaves like

$$en\mu^*_{n,\lfloor\rho n\rfloor} \sim enH_{\lfloor\rho n\rfloor} \sim en\log\rho n.$$

To justify and refine this, the most straightforward means is to apply the local limit theorem of the binomial distribution and then approximate the sum by a Gaussian integral; all details are messy but can be readily assisted by a symbolic computation software.

Here, due to the simple form of the asymptotic expansion (39), we use instead a different, self-contained approach relying partly on the following identity.

Lemma 7. *For $n \geqslant 1$,*

$$\sum_{0\leqslant m\leqslant n}\pi_{n,m}H_m = H_n - \sum_{1\leqslant k\leqslant n}\frac{\hat{\rho}^k}{k}. \tag{48}$$

Note that

$$\sum_{1\leqslant k\leqslant n}\frac{\hat{\rho}^k}{k} = -\log\rho + O\left(n^{-1}\hat{\rho}^n\right), \tag{49}$$

the O-term being exponentially small. So we can replace the right-hand side by $H_n + \log\rho$, introducing only an asymptotically negligible error.

Proof. We begin with the identity

$$\sum_{0\leqslant m\leqslant n}\pi_{n,m}a_m = [z^n]\frac{1}{1-\hat{\rho}z}\mathcal{A}\left(\frac{\rho z}{1-\hat{\rho}z}\right),$$

for any given sequence a_m, where $\mathcal{A}(z) := \sum_{m\geqslant 0}a_m z^m$. The proof of this identity is as follows.

$$\frac{1}{1-\hat{\rho}z}\mathcal{A}\left(\frac{\rho z}{1-\hat{\rho}z}\right) = \sum_{m\geqslant 0}a_m\frac{(\rho z)^m}{(1-\hat{\rho}z)^{m+1}} = \sum_{m\geqslant 0}a_m(\rho z)^m\sum_{\ell\geqslant 0}\binom{m+\ell}{\ell}\hat{\rho}^\ell z^\ell,$$

and by collecting the coefficient of z^n, we obtain the identity. Now substituting $a_m = H_m$ so that $\mathcal{A}(z) = \frac{1}{1-z}\log\frac{1}{1-z}$ and we get

$$\sum_{0\leqslant m\leqslant n}\pi_{n,m}H_m = [z^n]\frac{1}{1-\hat{\rho}z}\frac{1}{1-\frac{\rho z}{1-\hat{\rho}z}}\log\frac{1}{1-\frac{\rho z}{1-\hat{\rho}z}} = [z^n]\frac{1}{1-z}\log\frac{1-\hat{\rho}z}{1-z},$$

which proves (48). ∎

On the other hand, we also need an estimate for the following binomial sum.

Lemma 8. *Let ϕ be a $C^4[0,1]$ function. Then*

$$\sum_{0\leqslant m\leqslant n}\pi_{n,m}\phi(\alpha) = \phi(\rho) + \frac{\rho\hat{\rho}\phi''(\rho)}{2n} + O\big(n^{-2}\big), \tag{50}$$

Proof. By a Taylor expansion of $\phi(\alpha) = \phi(\frac{m}{n}) = \phi(\rho + \frac{m-\rho n}{n})$ at ρ of order four and then by exchanging the summations, we obtain

$$\sum_{0\leqslant m\leqslant n} \pi_{n,m}\phi(\alpha) = \sum_{0\leqslant j\leqslant 3} \frac{\phi^{(j)}(\rho)}{j!} \sum_{0\leqslant m\leqslant n} \pi_{n,m} \frac{(m-\rho n)^j}{n^j} + O\left(\sum_{0\leqslant m\leqslant n} \pi_{n,m} \frac{(m-\rho n)^4}{n^4}\right).$$

Then (50) follows from known estimates for the central moments of a binomial distribution: if $\mathcal{B}_n$ is binomially distributed with parameters n and ρ, then its variance equals $\rho\hat{\rho}n$, its third central moment is linear, and its fourth quadratic. ∎

Now, by replacing n by $n-1$ in (39), we have the expansion

$$\mu^*_{n-1,m} = H_m + \phi_1(\alpha) + \frac{H_m + \phi_2(\alpha) + \alpha\phi_1'(\alpha)}{n} + O\left(n^{-2}H_m\right).$$

By applying (48) and (50) term by term (and using (49)), we obtain

$$\begin{aligned}\sum_{0\leqslant m\leqslant n} \pi_{n,m}\mu^*_{n-1,m} &= H_n + \log\rho + \phi_1(\rho) \\ &\quad + \frac{2(H_n + \log\rho) + \phi_2(\rho) + \rho\phi_1'(\rho) + \rho\hat{\rho}\phi_1''(\rho)}{2n} + O\left(n^{-2}H_n\right).\end{aligned}$$

Then Theorem 3 follows from this and the expansions

$$\begin{aligned}\frac{n-1}{(1-\frac{1}{n})^n} &= e\left(n - \frac{1}{2} - \frac{1}{24n}\right) + O\big(n^{-2}\big), \\ H_n &= \log n + \gamma + \frac{1}{2n} + O\big(n^{-2}\big).\end{aligned}$$

∎

5 Asymptotics of the variance of $X_{n,m}$

We prove in this section that the variance $\sigma^2_{n,m} := \mathbb{V}(X_{n,m}) = \mathbb{E}(X^2_{n,m}) - (\mathbb{E}(X_{n,m}))^2$ of $X_{n,m}$ is asymptotically quadratic.

Theorem 4. *For $1 \leqslant m \leqslant n$, the variance of $X_{n,m}$ satisfies*

$$\begin{aligned}\frac{\mathbb{V}(X_{n,m})}{en} &= eH_m^{(2)}n - (2e+1)H_m + eH_m^{(2)} + e\psi_1(\alpha) - \phi_1(\alpha) - \frac{11e+1}{2n}H_m \\ &\quad + \frac{5eH_m^{(2)} + 2e\psi_2(\alpha) - 2\phi_2(\alpha) + 2e\alpha\psi_1'(\alpha) - 2\alpha\phi_1'(\alpha) + \phi_1(\alpha)}{2n} + O\left(n^{-2}H_m\right),\end{aligned}$$

where

$$\psi_1(\alpha) = \int_0^\alpha \left(\frac{S_2(x)}{S_1(x)^3} - \frac{1}{x^2} + \frac{2}{x}\right) \mathrm{d}x, \tag{51}$$

and

$$\begin{aligned}\psi_2(\alpha) = \frac{7}{12} - \int_0^\alpha \Bigg(&\frac{5S_1'(x)S_2(x)^2}{2S_1(x)^5} - \frac{2S_1'(x)S_3(x) + S_2(x)S_2'(x) + 6S_0(x)S_2(x)}{2S_1(x)^4} \\ &- \frac{S_0(x)}{S_1(x)^3} + \frac{2}{S_1(x)^2} - \frac{1}{x^3} + \frac{3}{x^2} - \frac{11}{2x}\Bigg) \mathrm{d}x.\end{aligned}$$

Note that $x = 0$ is a removable singularity for both integrands inside the parentheses above although $S_1(x) = x + O(x^2)$ for $x \sim 0$.

Similar to the mean, we work on the sequence $V^*_{n,m} := e_n^2(\sigma^2_{n+1,m} + \mu_{n+1,m})n^{-2}$ and prove that (see Figure 7 and Appendix F)

$$V^*_{n,m} = H_m^{(2)} - \frac{2H_m - \psi_1(\alpha) - 2H_m^{(2)}}{n} - \frac{\frac{11}{2}H_m - \psi_2(\alpha) - \frac{7}{3}H_m^{(2)}}{n^2} + O(n^{-3}H_m) \qquad (2 \leqslant m \leqslant n), \tag{52}$$

which can be proved to be equivalent to the expansion of Theorem 4.

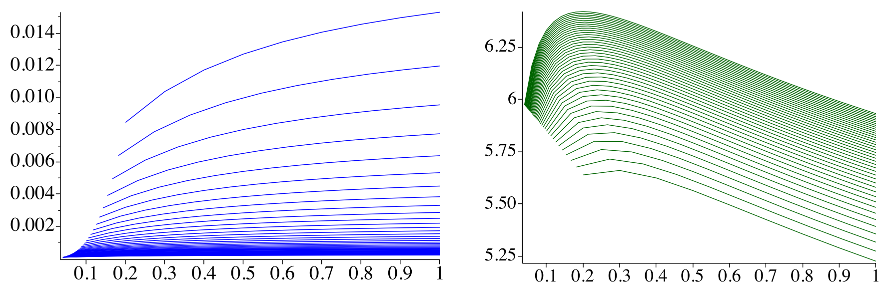


Figure 7: *The absolute differences* $|V^*_{n,m}-$ *RHS of* (52)$|$ *for* $2 \leqslant m \leqslant n$ *(normalized to the unit interval) and* $n = 10, \dots, 50$ *(left in top-down order), and the absolute normalized differences* $n^3 H_m^{-1}|V^*_{n,m}-$ *RHS of* (52)$|$ *for* $n = 10, \dots, 50$ *(right).*

The variance of X_n is computed by the relation

$$\mathbb{V}(X_n) = \sum_{0\leqslant m\leqslant n} \pi_{n,m}\left(\sigma^2_{n,m} + \mu^2_{n,m}\right) - \left(\sum_{0\leqslant m\leqslant n} \pi_{n,m}\mu_{n,m}\right)^2, \tag{53}$$

where $\pi_{n,m} = \binom{n}{m}\rho^m(1-\rho)^{n-m}$, $\mu_{n,m} := \mathbb{E}(X_{n,m})$ and $\sigma^2_{n,m} := \mathbb{V}(X_{n,m})$.

Theorem 5. *The variance of* X_n *satisfies asymptotically*

$$\frac{\mathbb{V}(X_n)}{en} = \frac{\pi^2}{6}en - (2e+1)(\log \rho n + \gamma) + v_1 - \frac{(11e+1)(\log\rho n + \gamma) - v_2}{2n} + O\left(n^{-2}\log n\right),$$

where ($\hat\rho := 1-\rho$) $v_1 := e\left(\frac{\pi^2}{6} - 1\right) - \phi_1(\rho) + e\psi_1(\rho) + 2e\hat\rho\phi_1'(\rho) + e\hat\rho\rho\phi_1'(\rho)^2$*, and*

$$\begin{aligned} v_2 = {} & e\hat\rho^2\rho^2\phi_1''(\rho)^2 + 2e\hat\rho^2\rho(1+\rho\phi_1'(\rho))\phi_1'''(\rho) + 2e\rho\hat\rho(1+\phi_1'(\rho))\phi_1''(\rho) \\ & + 4e\hat\rho(1+\rho\phi_2'(\rho))\phi_1'(\rho) + 2e\hat\rho\rho\phi_1'(\rho)^2 + 4e\hat\rho\phi_2'(\rho) + e\hat\rho\rho\psi_1''(\rho) - \hat\rho\rho\phi_1''(\rho) \\ & + 2e\psi_2(\rho) - 2\phi_2(\rho) + 2e\rho\psi_1'(\rho) - 2\rho\phi_1'(\rho) + \phi_1(\rho) + \tfrac{5}{6}e\pi^2 - 3e - 1. \end{aligned}$$

Recurrence satisfied by the variance. As in the case of $m = O(1)$, we begin our asymptotic analysis with the recurrence (15), which, in terms of $V^*_{n,m} := e_n^2(\sigma^2_{n+1,m} + \mu_{n+1,m})n^{-2}$, has the form: $V^*_{n,0} = 0$, and for $1 \leqslant m \leqslant n$

$$\sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell}\left(V^*_{n,m} - V^*_{n,m-\ell}\right) = T^*_{n,m}, \tag{54}$$

where

$$T^*_{n,m} := \sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell} \left(\mu^*_{n,m} - \mu^*_{n,m-\ell}\right)^2 . \tag{55}$$

In particular, this gives

$$\begin{aligned} V^*_{n,1} &= 1, \\ V^*_{n,2} &= \frac{5n^4 + 8n^3 - n^2 - 4n + 1}{(2n^2 + 2n - 1)^2}. \end{aligned}$$

The expressions become very lengthy as m increases. In Appendix E we give asymptotic expansions for $V^*_{n,m}$ for a few small m as $n \to \infty$. Based on these expansions, a suitable Ansatz for the asymptotic behavior of $V^*_{n,m}$ can be deduced (assisted again by a computer algebra system), which can then be proved by the same method of proof presented in Section 4 for the mean.

Proof of the first-order asymptotics with error analysis. By the same procedure used for $\mu^*_{n,m}$, we start from computing the asymptotic expansions for $V^*_{n,m}$ for small m. These expansions suggest the more uniform (for $1 \leqslant m \leqslant n$) asymptotic expansion

$$V^*_{n,m} \sim c_4 H^{(2)}_m + \frac{a_1 H_m + \psi_1(\alpha) + c_5 H^{(2)}_m}{n},$$

for some constants c_4, c_5 and a_1, and some function $\psi_1(z)$. Such an asymptotic form can be justified by the same approach we used above for $\mu^*_{n,m}$. More precisely, we now prove that

$$V^*_{n,m} = H^{(2)}_m - \frac{2H_m - \psi_1(\alpha) - 2H^{(2)}_m}{n} + O(n^{-2} H_m), \tag{56}$$

uniformly for $0 \leqslant m \leqslant n$ and $n \geqslant 1$, where $\psi_1(z)$ is given in (51). Our proof starts from considering the difference

$$\Delta^*_{n,m} := V^*_{n,m} - c_4 H^{(2)}_m - \frac{a_1 H_m + \psi_1(\alpha) + c_5 H^{(2)}_m}{n},$$

and specify the involved coefficients and $\psi_1(z)$ such that $\Delta^*_{n,m} = O(n^{-2} H_m)$. By (54), $\Delta^*_{n,m}$ satisfies, for $1 \leqslant m \leqslant n$, the recurrence

$$\sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell} \left(\Delta^*_{n,m} - \Delta^*_{n,m-\ell}\right) = \tilde{E}_1(n,m), \tag{57}$$

with the initial value $\Delta^*_{n,0} = -\frac{\psi_1(0)}{n}$, where ($T^*_{n,m}$ being defined in (55))

$$\begin{aligned} \tilde{E}_1(n,m) := T^*_{n,m} - \sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell} \Bigg\{ &\left(c_4 + \frac{c_1}{n}\right)\left(H^{(2)}_m - H^{(2)}_{m-\ell}\right) + \frac{a_1}{n}\left(H_m - H_{m-\ell}\right) \\ &+ \frac{\psi_1\left(\frac{m}{n}\right) - \psi_1\left(\frac{m-\ell}{n}\right)}{n} \Bigg\}. \end{aligned}$$

We will derive an asymptotic expansion for $\tilde{E}_1(n,m)$. For that purpose, we use the expansions (82), (83) as well as Theorem 2 in Section 4, and apply the same error analysis used for

$\mu^*_{n,m}$. A careful analysis then leads to

$$\sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell}\left(\mu^*_{n,m}-\mu^*_{n,m-\ell}\right)^2 = \frac{1}{mn}+\frac{1}{n^2}\left(-\frac{1}{\alpha}+\frac{(1+\alpha\phi_1'(\alpha))^2}{\alpha^2}S_2(\alpha)\right) + \frac{3}{2mn^2}+\frac{[\![m\geqslant 2]\!]}{2m(m-1)n^2}-\frac{[\![m=1]\!]}{n^2}+O\big(n^{-3}\big), \tag{58}$$

and

$$\sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell}\left(H_m^{(2)}-H_{m-\ell}^{(2)}\right) = \frac{1}{mn}+\frac{1}{n^2}\left(\frac{S_1(\alpha)-\alpha}{\alpha^2}\right)-\frac{1}{2mn^2}+\frac{[\![m\geqslant 2]\!]}{2m(m-1)n^2}-\frac{[\![m=1]\!]}{n^2}+O\big(n^{-3}\big), \tag{59}$$

both holding uniformly for $1\leqslant m\leqslant n$ as $n\to\infty$.

Collecting the expansions (82), (83), (58) and (59), we obtain

$$\begin{aligned}\tilde{E}_1(n,m) = \frac{1-c_4}{mn}+\frac{1}{n^2}\Big\{&-\frac{1}{\alpha}+\frac{(1+\alpha\phi_1'(\alpha))^2}{\alpha^2}S_2(\alpha)-\frac{c_4}{\alpha^2}\left(S_1(\alpha)-\alpha\right)\\ &-\left(\frac{a_1}{\alpha}+\psi_1'(\alpha)\right)S_1(\alpha)\Big\}+\frac{1}{mn^2}\left(\frac{3}{2}+\frac{c_4}{2}-c_5\right)\\ &-\frac{[\![m=1]\!](1-c_4)}{n^2}+\frac{[\![m\geqslant 2]\!](1-c_4)}{2m(m-1)n^2}+O\big(n^{-3}\big),\end{aligned}$$

uniformly for $1\leqslant m\leqslant n$.

We can now specify all the undetermined constants and $\psi_1(z)$ such that all terms except the last will vanish and $\tilde{E}_1(n,m)=O\big(n^{-3}\big)$. This entails first the choices $c_4=1$ and $c_1=2$.

It remains only the $\frac{1}{n^2}$-term. We consider the limit when α tends to zero using the Taylor expansions (36), and deduce that $a_1=-2$. These values give the equation satisfied by $\psi_1'(z)$

$$\psi_1'(z)S_1(z) = -\frac{S_1(z)}{z^2}+\frac{(1+z\phi_1'(z))^2}{z^2}S_2(z)+\frac{2S_1(z)}{z},$$

which in view of (3) leads to the differential equation

$$\psi_1'(z) = \frac{S_2(z)}{S_1^3(z)}-\frac{1}{z^2}+\frac{2}{z}. \tag{60}$$

Thus with the choices $c_4=1$, $a_1=-2$, $c_5=2$, and the function $\psi_1'(z)$ by (60), we get the bound $\tilde{E}_1(n,m)=O\big(n^{-3}\big)$ uniformly for $1\leqslant m\leqslant n$. Accordingly, by (57), the sequences $\Delta_{n,m}:=n^2\Delta^*_{n,m}$ satisfy the recurrence

$$\sum_{1\leqslant \ell\leqslant m}\lambda^*_{n,m,\ell}\left(\Delta_{n,m}-\Delta_{n,m-\ell}\right)=O\big(n^{-1}\big)\qquad(1\leqslant m\leqslant n),$$

with $\Delta_{n,0}=-\psi_1(0)n$. Choose now the initial value $\psi_1(0)=0$, so that $\Delta_{n,0}=0$ and Lemma 4 can be applied. This implies that $\Delta_{n,m}=O(H_m)$, and consequently $\Delta^*_{n,m}=O(n^{-2}H_m)$.

Also $\psi_1(z)$ is indeed given by (51). In particular, the first few terms in the Taylor expansion of $\psi_1(z)$ are given as follows.

$$\psi_1(z)=\tfrac{11}{4}z-\tfrac{49}{36}z^2+\tfrac{2473}{4320}z^3+\tfrac{1307}{14400}z^4-\tfrac{12743687}{18144000}z^5+\tfrac{194960323}{152409600}z^6+\cdots.$$

This completes the proof of the asymptotic expansion (56) for $V^*_{n,m}$. The more refined approximation (52) follows the same line of proof but with more detailed expansions; see Appendix E.

The variance of X_n. The method of proof used for characterizing the asymptotics of $\mathbb{E}(X_n)$ (namely, § 4.3) can be applied here but requires slight modifications because unlike (48) we do not have simple closed-form expressions for the sum $\sum_{0\leqslant m\leqslant n} \pi_{n,m}(H_m^{(2)} + H_m^2)$, which arises from the first sum on the right-hand side of (53). There are a few different ways to manipulate the corresponding asymptotic expansion. One simple idea is to apply Lemma 8 using the relations

$$H_m = \Psi(m+1) + \gamma, \qquad H_m^{(2)} = \frac{\pi^2}{6} - \Psi'(m+1),$$

where $\Psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}$ denotes the digamma function. Noting that $\Psi'(x) = \sum_{j\geqslant 0} \frac{1}{(j+x)^2}$, we see that successive derivatives of $\Psi^{(j)}(x)$ behave like $j!x^{-j}$ for large x. We then group terms of the same power and deduce Theorem 5.

6 Limit laws of $X_{n,m}$ when $m \to \infty$

We show in this section that the distribution of $X_{n,m}$, when properly normalized, tends to a Gumbel (or extreme-value or double exponential) distribution, as $m \to \infty$, $m \leqslant n$. The proof consists in showing that the result (9) when $m = O(1)$ extends to all $m \leqslant n$ but requires an additional correction term ϕ_1 coming from the linear part of the random variables, which complicates significantly the proof.

The standard Gumbel distribution $\mathscr{G}(1)$ (with mode zero, mean γ) is characterized by the distribution function $e^{-e^{-x}}$ and the characteristic function $\Gamma(1 - i\theta)$, respectively. Note that if $X \sim \text{Exp}(1)$, then $-\log X \sim \mathscr{G}(1)$, which was the description used in Garnier et al. (1999).

The genesis of the Gumbel distribution is easily seen as follows.

Lemma 9. *Let $W_m := \sum_{1\leqslant r\leqslant m} \text{Exp}(r)$, where the m exponential random variables are independent. Then $W_m - \log m$ converges in distribution to the Gumbel distribution*

$$\mathbb{P}(W_m - \log m \leqslant x) \to e^{-e^{-x}} \qquad (x \in \mathbb{R}; m \to \infty).$$

Proof. We have

$$\mathbb{E}\big(e^{(W_m - H_m)i\theta}\big) = \prod_{1\leqslant r\leqslant m} \frac{e^{-\frac{i\theta}{r}}}{1 - \frac{i\theta}{r}} \to \prod_{r\geqslant 1} \frac{e^{-\frac{i\theta}{r}}}{1 - \frac{i\theta}{r}} = e^{-\gamma i\theta}\Gamma(1 - i\theta),$$

for each bounded $\theta \in \mathbb{R}$. Here we used the infinite-product representation of the Gamma function

$$\Gamma(1+s) = e^{-\gamma s} \prod_{r\geqslant 1} \frac{e^{\frac{s}{r}}}{1 + \frac{s}{r}} \qquad (s \in \mathbb{C} \setminus \mathbb{Z}^-).$$

The lemma then follows from the asymptotic estimate

$$H_m = \log m + \gamma + O(m^{-1}) \qquad (m \to \infty),$$

and Lévy's continuity theorem (van der Vaart, 1998, §2.3). ∎

Unlike the case when $m = O(1)$, we need to subtract more terms to have the limiting distribution. *Throughout this section, let $\kappa \in (0,1)$ be a generic symbol whose value is independent of m, n and may differ from one occurrence to another.*

Proposition 1. *For $1 \leqslant m \leqslant n$, we have the uniform asymptotic approximation*

$$\mathbb{E}\big(e^{\frac{X_{n,m}}{en}s - (H_m + \phi_1(\frac{m}{n}))s}\big) = \left(1 + O\left(\frac{H_m}{n}\right)\right) \prod_{1\leqslant r\leqslant m} \frac{e^{-\frac{s}{r}}}{1 - \frac{s}{r}},$$

for $|s| \leqslant \kappa$, where ϕ_1 is defined in (3).

Note that $\phi_1(x) = O(x)$ as $x \to 0$, and thus $\phi_1(\frac{m}{n}) = o(1)$ when $m = O(1)$. In this case, the proposition re-proves Theorem 1 (with a better error term).

A combination of Lemma 9 and Proposition 1 leads to the limit law for $X_{n,m}$ in the remaining range.

Theorem 6. *If $m \to \infty$ with n and $m \leqslant n$, then the number $X_{n,m}$ of steps taken by the $(1+1)$-EA to reach the optimal state with all bits 1 (when starting from the initial state with $n-m$ 1s) satisfies*

$$\mathbb{P}\left(\frac{X_{n,m}}{en} - \log m - \phi_1(\tfrac{m}{n}) \leqslant x\right) \to e^{-e^{-x}} \qquad (x \in \mathbb{R}),$$

where ϕ_1 is defined in (3).

Theorem 7. *The number X_n of steps taken by the $(1+1)$-EA to reach the final state $f(\mathbf{x}) = n$, when starting from a random initial state where each bit assumes 1 with probability ρ, satisfies*

$$\mathbb{P}\left(\frac{X_n}{en} - \log \rho n - \phi_1(\rho) \leqslant x\right) \to e^{-e^{-x}} \qquad (x \in \mathbb{R}).$$

From Figure 8, we see the fast convergence of the distribution to the limit law.

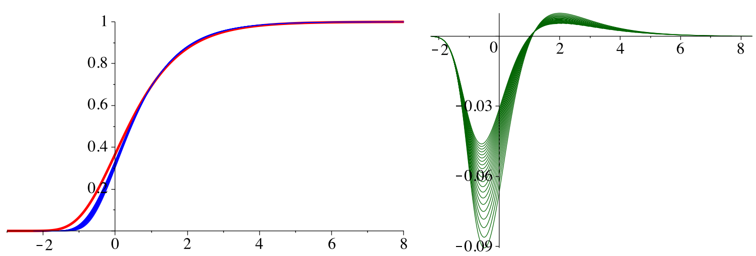


Figure 8: *Left: distributions of $X_n^* := \frac{X_n}{en} - \log n + \log 2 - \phi_1(\frac{1}{2})$ for $n = 15, \dots, 35$ (in blue), and the limiting Gumbel curve (in red); Right: the difference between the distribution function of X_n and that of a Gumbel.*

Outline of proofs. We focus on the proof of Proposition 1, which is the main hard part of all the proofs. Since we want to prove that the two moment generating functions are asymptotically close, we introduce the normalized function

$$F_{n,m}(s) := \frac{\mathbb{E}\big(e^{\frac{X_{n,m}}{en}s}\big)e^{-\phi(\frac{m}{n})s}}{\prod_{1\leqslant r\leqslant m}\frac{1}{1-\frac{s}{r}}} = \frac{P_{n,m}\big(e^{\frac{s}{en}}\big)e^{-H_m s-\phi(\frac{m}{n})s}}{\prod_{1\leqslant r\leqslant m}\frac{e^{-\frac{s}{r}}}{1-\frac{s}{r}}}.$$

If we assume ϕ to be a $C^2[0,1]$-function and choose $\phi(x) = \phi_1(x)$, where ϕ_1 (see (3)) appears as the second-order term in the asymptotic expansion of the mean (see (39)), then

$$F_{n,m}(s) \sim 1, \tag{61}$$

uniformly for all $1 \leqslant m \leqslant n$, $n \to \infty$, and $|s| \leqslant \kappa$, and this will prove Proposition 1. Indeed, our induction proof here does not rely on any information of the mean asymptotics and entails particularly the right choice of $\phi(x)$. This is why we specify ϕ only at a later stage.

The idea of our proof for (61) relies again on a similar inductive argument we used above but with more parameters (n, m, s) involved in the analysis, which adds to the technical complication, notably in controlling the uniformity of the error terms. To simplify the proof, we thus assume that s is real (implying that $e^s > 0$), so that all inequalities become easier to handle.

The recurrence satisfied by $F_{n,m}$. Rewriting $P_{n,m}$ in terms of $F_{n,m}$

$$P_{n,m}\big(e^{\frac{s}{en}}\big) = F_{n,m}(s)e^{H_m s+\phi(\frac{m}{n})s}\prod_{1\leqslant r\leqslant m}\frac{e^{-\frac{s}{r}}}{1-\frac{s}{r}},$$

and substituting $t = e^{\frac{s}{en}}$ in (4), we see that $F_{n,m}(s)$ satisfies the recurrence

$$F_{n,m}(s) = \frac{e^{\frac{s}{en}}\displaystyle\sum_{1\leqslant \ell\leqslant m}\lambda_{n,m,\ell}F_{n,m-\ell}(s)e^{-\left(\phi(\frac{m}{n})-\phi(\frac{m-\ell}{n})\right)s}\prod_{m-\ell+1\leqslant r\leqslant m}\big(1-\frac{s}{r}\big)}{1-\big(1-\Lambda_{n,m}\big)e^{\frac{s}{en}}},$$

for $1\leqslant m\leqslant n$, with $F_{n,0}(s) = 1$, where $\Lambda_{n,m} := \sum_{1\leqslant \ell\leqslant m}\lambda_{n,m,\ell}$.

An auxiliary sum. Since we expect $F_{n,m}(s)$ to be close to 1, we replace all occurrences of F on the right-hand side by 1 and consider the function

$$G_{n,m}(s) := \frac{e^{\frac{s}{en}}\displaystyle\sum_{1\leqslant \ell\leqslant m}\lambda_{n,m,\ell}e^{-\left(\phi(\frac{m}{n})-\phi(\frac{m-\ell}{n})\right)s}\prod_{m-\ell+1\leqslant r\leqslant m}\big(1-\frac{s}{r}\big)}{1-\big(1-\Lambda_{n,m}\big)e^{\frac{s}{en}}}. \tag{62}$$

We will show how to choose ϕ so that $G_{n,m}(s)$ will be very close to 1. The following lemma is the crucial step in our proof.

Lemma 10. *Let $\phi(x)$ be a C^2-function on the unit interval satisfying $\phi(0) = 0$. Then*

$$G_{n,m}(s) = \frac{1-\frac{s}{m}\left(1+\alpha\phi'(\alpha)\right)\frac{S_1(\alpha)}{S(\alpha)}+O\big(\frac{1}{mn}\big)}{1-\frac{s}{m}\cdot\frac{\alpha}{S(\alpha)}+O\big(\frac{1}{mn}\big)}, \tag{63}$$

where the O-terms hold uniformly for $1\leqslant m\leqslant n$ and $|s|\leqslant\kappa$.

Proof. The proof consists in a detailed inspection of all factors, using estimates (34) and (35) we derived earlier for $\Lambda_{n,m}^{(r)}$. We consider first the case when $m = O(1)$. In this case, $S(\alpha), S_1(\alpha) = \alpha + O(\alpha^2)$ (see (35)) and the numerator and the denominator of (62) both have the form

$$1-\frac{s}{m}+O\big(n^{-1}\big),$$

which can be readily checked by using the estimates (10) and

$$\Lambda_{n,m} = e^{-1}\alpha + O(\alpha^2).$$

From now on, we assume $m\geqslant m_0$, where m_0 is sufficiently large, say $m_0\geqslant 10$. Throughout the proof, all O-terms hold uniformly for $|s|\leqslant\kappa$ and $m_0\leqslant m\leqslant n$ and n large enough.

We begin with the denominator of $G_{n,m}(s)$, which satisfies

$$\begin{aligned}1-(1-\Lambda_{n,m})e^{\frac{s}{en}} &= \Lambda_{n,m}-\frac{s}{en}+\Lambda_{n,m}\frac{s}{en}+O\big(n^{-2}\big)\\ &= \Lambda_{n,m}\left(1+\frac{s}{en}\right)\left(1-\frac{s}{en\Lambda_{n,m}}+O\left((mn)^{-1}\right)\right),\end{aligned}$$

where we used the estimate $\Lambda_{n,m} = \Omega(\alpha)$; see (35). By (34) and (35), the second-order term on the right-hand side satisfies

$$\frac{s}{en\Lambda_{n,m}} = \frac{s}{nS(\alpha)(1+O\big(n^{-1}\big))} = \frac{s}{m}\cdot\frac{\alpha}{S(\alpha)}+O\left((mn)^{-1}\right).$$

Thus we obtain

$$1-(1-\Lambda_{n,m})e^{\frac{s}{en}} = \Lambda_{n,m}\left(1+\frac{s}{en}\right)\left(1-\frac{s}{m}\cdot\frac{\alpha}{S(\alpha)}+O\left((mn)^{-1}\right)\right). \tag{64}$$

Now we turn to the numerator of $G_{n,m}(s)$ and look first at the exponential term

$$\begin{aligned} e^{-\left(\phi(\frac{m}{n})-\phi(\frac{m-\ell}{n})\right)s} &= e^{-\frac{\ell}{n}\phi'(\alpha)s+O\left(\frac{\ell^2}{n^2}\right)} \\ &= \left(1-\frac{\ell}{n}\phi'(\alpha)s\right)\left(1+O\left(\frac{\ell^2}{n^2}\right)\right), \end{aligned}$$

uniformly for $1\leqslant\ell\leqslant m$, where we used the twice continuous differentiability of ϕ.

Consider now the finite product $\prod_{m-\ell+1\leqslant r\leqslant m}\left(1-\frac{s}{r}\right)$. Obviously, for $|s|\leqslant 1$, we have the uniform bound

$$\prod_{m-\ell+1\leqslant r\leqslant m}\left|1-\frac{s}{r}\right| \leqslant \prod_{m-\ell+1\leqslant r\leqslant m}\left(1+\frac{1}{r}\right)\leqslant e^{H_m} = O(m).$$

On the other hand, we also have the finer estimates

$$\begin{aligned} \prod_{m-\ell+1\leqslant r\leqslant m}\left(1-\frac{s}{r}\right) &= e^{-(H_m-H_{m-\ell})s}\left(1+O\left(\frac{\ell^2}{m^2}\right)\right) \\ &= e^{-\frac{\ell}{m}s+O\left(\frac{\ell^2}{m^2}\right)}\left(1+O\left(\frac{\ell^2}{m^2}\right)\right) \\ &= \left(1-\frac{\ell}{m}s\right)\left(1+O\left(\frac{\ell^2}{m^2}\right)\right), \end{aligned}$$

uniformly for $1\leqslant\ell=o(m)$.

Combining these two estimates, we obtain the approximation

$$\prod_{m-\ell+1\leqslant r\leqslant m}\left(1-\frac{s}{r}\right) = \left(1-\frac{\ell}{m}s\right)\left(1+[\![\ell\geqslant 2]\!]O\left(\frac{\ell^2}{m^2}\right)+[\![\ell>\lceil\sqrt{m}\rceil]\!]O(m)\right),$$

which holds uniformly for $1\leqslant\ell\leqslant m$. Thus the numerator, up to the factor $e^{\frac{s}{en}}$, satisfies

$$\begin{aligned} &\sum_{1\leqslant\ell\leqslant m}\lambda_{n,m,\ell}e^{-\left(\phi(\frac{m}{n})-\phi(\frac{m-\ell}{n})\right)s}\prod_{m-\ell+1\leqslant r\leqslant m}\left(1-\frac{s}{r}\right) \\ &\quad= \sum_{1\leqslant\ell\leqslant m}\lambda_{n,m,\ell}-\frac{s}{m}\left(1+\alpha\phi'(\alpha)\right)\sum_{1\leqslant\ell\leqslant m}\ell\lambda_{n,m,\ell} \\ &\qquad+O\left(\frac{1}{m^2}\sum_{2\leqslant\ell\leqslant m}\ell^2\lambda_{n,m,\ell}+\frac{1}{mn}\sum_{1\leqslant\ell\leqslant m}\ell^2\lambda_{n,m,\ell}+m\sum_{\lceil\sqrt{m}\rceil+1\leqslant\ell\leqslant m}\lambda_{n,m,\ell}\right). \end{aligned}$$

Each of the sums can be readily estimated as in (46) and (47), and we have

$$m^{-2}\sum_{2\leqslant\ell\leqslant m}\ell^2\lambda_{n,m,\ell} = O\left(n^{-2}\right).$$

Similarly,

$$(mn)^{-1} \sum_{1\leqslant \ell\leqslant m} \ell^2 \lambda_{n,m,\ell} = O\left((mn)^{-1}\alpha\right) = O\big(n^{-2}\big).$$

Finally, for $m \geqslant 1$,

$$\begin{aligned} m \sum_{\lceil\sqrt{m}\rceil+1\leqslant \ell\leqslant m} \lambda_{n,m,\ell} &= O\left(\frac{m\alpha^{\sqrt{m}+1}}{\Gamma(\sqrt{m}+2)}\right) \\ &= O\left(m^{\frac{7}{4}} n^{-1} e^{-\sqrt{m}(\log n - \frac{1}{2}\log m - 1)}\right) = O\big(n^{-2}\big). \end{aligned}$$

Collecting these estimates, we get

$$\begin{aligned} &\sum_{1\leqslant \ell\leqslant m} \lambda_{n,m,\ell} e^{-\left(\phi(\frac{m}{n})-\phi(\frac{m-\ell}{n})\right)s} \prod_{m-\ell+1\leqslant r\leqslant m} \left(1-\frac{s}{r}\right) \\ &\quad = \Lambda_{n,m} - \frac{s}{m}\left(1+\alpha\phi'(\alpha)\right)\Lambda_{n,m}^{(1)} + O\big(n^{-2}\big) \\ &\quad = \Lambda_{n,m}\left(1-\frac{s}{m}\left(1+\alpha\phi'(\alpha)\right)\frac{\Lambda_{n,m}^{(1)}}{\Lambda_{n,m}} + O\left((mn)^{-1}\right)\right) \\ &\quad = \Lambda_{n,m}\left(1-\frac{s}{m}\left(1+\alpha\phi'(\alpha)\right)\frac{S_1(\alpha)}{S(\alpha)} + O\left((mn)^{-1}\right)\right), \end{aligned} \tag{65}$$

by applying (34).

By (64), (65) and the simple estimate

$$e^{\frac{s}{en}} = \left(1+\tfrac{s}{en}\right)\left(1+O\big(n^{-2}\big)\right),$$

we conclude (63). ▮

Corollary 7. *Let* $\phi(x) = \phi_1(x) = \int_0^x \left(\frac{1}{S_1(t)} - \frac{1}{t}\right) \mathrm{d}t$. *Then*

$$G_{n,m}(s) = 1 + O\big((mn)^{-1}\big),$$

where the O*-term holds uniformly for* $1 \leqslant m \leqslant n$*,* n *large enough and* $|s| \leqslant \kappa$.

Proof. To obtain the error term $O((mn)^{-1})$, we choose ϕ in a way that the two middle terms in the fraction of (63) are identical, which means

$$\frac{x}{S(x)} = (1+x\phi'(x))\frac{S_1(x)}{S(x)}.$$

Observe that $S(x) > 0$ for $x > 0$. This, together with $\phi_1(0) = 0$, implies $\phi = \phi_1$, which is not only a C^2-function but also analytic in the unit circle. ▮

Proof of Proposition 1. We now prove Proposition 1 by induction.

Lemma 11. *Let* $\phi = \phi_1$. *Then there exists an* $n_0 > 0$ *such that*

$$F_{n,m}(s) = 1 + O\left(n^{-1}H_m\right),$$

uniformly for $0 \leqslant m \leqslant n$*,* $n \geqslant n_0$ *and* $|s| \leqslant \kappa$*,* $\kappa \in (0,1)$.

Proof. We use induction on m and show that there exists a constant $C > 0$, such that

$$|F_{n,m}(s) - 1| \leqslant C n^{-1} H_m,$$

for all $1 \leqslant m \leqslant n$, $n \geqslant n_0$ and $|s| \leqslant \kappa$. Here C is independent of n_0.

When $m = 0$, the lemma holds, since $F_{n,0}(s) \equiv 1$.

Assume that the lemma holds for all functions $F_{n,k}(s)$ with $0 \leqslant k \leqslant m$ and $n \geqslant n_0$. By Corollary 7, there exists a constant $C_1 > 0$ such that for all $1 \leqslant m \leqslant n$, $n \geqslant n_1$ and $|s| \leqslant \kappa_1$, $0 < \kappa_1 < 1$,

$$|G_{n,m}(s) - 1| \leqslant C_1 (mn)^{-1}.$$

Now

$$\begin{aligned} |F_{n,m}(s) - 1| &= |F_{n,m}(s) - G_{n,m}(s) + G_{n,m}(s) - 1| \\ &\leqslant |F_{n,m}(s) - G_{n,m}(s)| + C_1 (mn)^{-1}. \end{aligned}$$

The first term on the right-hand side can be re-written as

$$\begin{aligned} &|F_{n,m}(s) - G_{n,m}(s)| \\ &= \frac{\left| e^{\frac{s}{en}} \sum_{1 \leqslant \ell \leqslant m} \lambda_{n,m,\ell} (F_{n,m-\ell}(s) - 1) e^{-\left(\phi\left(\frac{m}{n}\right) - \phi\left(\frac{m-\ell}{n}\right)\right)s} \prod_{m-\ell+1 \leqslant r \leqslant m} \left(1 - \frac{s}{r}\right) \right|}{\left| 1 - \left(1 - \Lambda_{n,m}\right) e^{\frac{s}{en}} \right|}. \end{aligned}$$

Since we assume $|s| \leqslant 1$, the product involved in the sum on the right-hand side is nonnegative and we have, by the induction hypothesis,

$$\begin{aligned} |F_{n,m}(s) - G_{n,m}(s)| &\leqslant \frac{e^{\frac{s}{en}} \sum_{1 \leqslant \ell \leqslant m} \lambda_{n,m,\ell} e^{-\left(\phi\left(\frac{m}{n}\right) - \phi\left(\frac{m-\ell}{n}\right)\right)s} \frac{C H_{m-\ell}}{n} \prod_{m-\ell+1 \leqslant r \leqslant m} \left(1 - \frac{s}{r}\right)}{1 - \left(1 - \Lambda_{n,m}\right) e^{\frac{s}{en}}} \\ &\leqslant \frac{C H_{m-1}}{n} G_{n,m}(s) \leqslant \frac{C H_{m-1}}{n} + \frac{C H_{m-1}}{n} \cdot |G_{n,m}(s) - 1| \\ &\leqslant \frac{C H_{m-1}}{n} + \frac{C H_{m-1}}{n} \cdot \frac{C_1}{mn}. \end{aligned}$$

It follows that

$$|F_{n,m}(s) - 1| \leqslant \frac{C H_m}{n} + \frac{1}{mn} \left(C_1 - C + C_1 C \frac{H_{m-1}}{n} \right).$$

Choose first $n_2 \geqslant n_1$ such that $\frac{H_{n_2 - 1}}{n_2} \leqslant \frac{1}{2C_1}$, which implies that $C_1 C \frac{H_{m-1}}{n} \leqslant \frac{C}{2}$ for $1 \leqslant m \leqslant n$ and $n \geqslant n_2$. Then choose $C = 2C_1$. We then have

$$C_1 - C + \frac{C_1 C H_{m-1}}{n} \leqslant C_1 - \frac{C}{2} \leqslant 0,$$

and thus

$$|F_{n,m}(s) - 1| \leqslant \frac{C H_m}{n}.$$

Note that apart from requiring $|s| \leqslant 1$ the only restriction on s comes from $G_{n,m}(s)$, thus we may choose $\kappa = \min(1, \kappa_1)$. This completes the proof. ∎

The Gumbel limit laws for $X_{n,m}$ ($m \to \infty$). We prove Theorem 6 by Proposition 1. Since $m \to \infty$, we have

$$\begin{aligned}
\mathbb{E}\big(e^{\frac{X_{n,m}}{en}s-(\log m+\phi_1(\alpha))s}\big) &= P_{n,m}\big(e^{\frac{s}{en}}\big)e^{-(H_m-\gamma+\phi_1(\frac{m}{n}))s}\left(1+O\left(m^{-1}\right)\right) \\
&= e^{\gamma s}F_{n,m}(s)\prod_{1\leqslant r\leqslant m}\frac{e^{-\frac{s}{r}}}{1-\frac{s}{r}}\left(1+O\left(m^{-1}\right)\right) \\
&= \Gamma(1-s)\left(1+O\left(\frac{\log m}{n}+\frac{1}{m}\right)\right). \qquad (66)
\end{aligned}$$

Thus Theorem 6 follows from an application of Curtiss' theorem (Flajolet and Sedgewick, 2009, IX 4.2), which is similar to Lévy's continuity theorem but with characteristic function replaced by moment generating function.

The Gumbel limit law for X_n. We now prove Theorem 7, starting from the moment generating function

$$\mathbb{E}\big(e^{X_n s}\big) = \sum_{0\leqslant m\leqslant n}\pi_{n,m}P_{n,m}\left(e^s\right),$$

where $\pi_{n,m} = \binom{n}{m}\rho^m\hat{\rho}^{n-m}$ ($\hat{\rho} := 1-\rho$). Then

$$\mathbb{E}\big(e^{\frac{X_n}{en}s-(\log\rho n+\phi_1(\rho))s}\big) = \sum_{0\leqslant m\leqslant n}\pi_{n,m}\underbrace{P_{n,m}\big(e^{\frac{s}{en}}\big)e^{-(H_m-\gamma+\phi_1(\alpha))s}}_{(66)}\times e^{\delta_{n,m}s},$$

where

$$\delta_{n,m} := H_m - \log\rho n - \gamma + \phi_1(\alpha) - \phi_1(\rho).$$

Since the binomial distribution is highly concentrated around the range $m = \rho n + x\sqrt{\rho\hat{\rho}n}$ where $x = o(n^{\frac{1}{6}})$, we see that

$$\delta_{n,m} = \frac{\sqrt{\rho\hat{\rho}}}{S_1(\rho)\sqrt{n}}x + \frac{1}{2\rho n}\left(1-\frac{\rho^2\hat{\rho}S_1'(\rho)}{S_1(\rho)^2}x^2\right) + O\left(\frac{|x|+|x|^3}{n^{\frac{3}{2}}}\right),$$

for m in this range.

By a standard argument (Gaussian approximation of the binomial and exponential tail estimates) using the expansion (66), we then deduce that

$$\begin{aligned}
\mathbb{E}\big(e^{\frac{X_n}{en}s-(\log\rho n+\phi_1(\rho))s}\big) &= \Gamma(1-s)\sum_{|m-\rho n|\leqslant n^{\frac{7}{12}}}\pi_{n,m}\left(1+O\left(n^{-1}\log n\right)\right) \\
&\quad + O\left(\sum_{|m-\rho n|>n^{\frac{7}{12}}}\pi_{n,m}\right) \\
&= \Gamma(1-s)\big(1+O\big(n^{-1}\log n\big)\big).
\end{aligned}$$

This proves Theorem 7.

7 Analysis of the $(1+1)$-EA for LEADINGONES

We consider the optimization time Y_n of the $(1+1)$-EA when the underlying fitness function is the number of leading 1s. This problem has been examined repeatedly in the literature due to the simple structural properties it exhibits; see Böttcher et al. (2010); Droste et al. (2002); Ladret (2005) and the references therein. The strongest results obtained were those in Ladret (2005) (but almost unknown in the EA literature) where she proved that the optimization time under LEADINGONES is asymptotically normally distributed with mean asymptotic to $\frac{e^c-1}{2c^2}n^2$ and variance to $\frac{3(e^{2c}-1)}{8c^3}n^3$, where $p=\frac{c}{n}$, $c>0$.

We revisit this problem and obtain similar type of results by an analytic-combinatorial approach (see Flajolet and Sedgewick (2009)) based on generating function and recurrence relations. This approach is not only very different from the purely probabilistic techniques used in Ladret (2005) but also differs significantly from the one we used for ONEMAX mainly because the generating functions here have simpler forms. It can also be readily amended for obtaining the convergence rate to the normal law.

Denote by Y_n the number of steps taken by Algorithm $(1+1)$-EA to reach the optimum state for LEADINGONES when starting from a random input (each bit being 1 with probability $\frac{1}{2}$). Then, similar to X_n for ONEMAX, its moment generating function satisfies

$$\mathbb{E}\left(e^{Y_n s}\right) := 2^{-n} + \sum_{1\leqslant m\leqslant n} 2^{m-n-1} Q_{n,m}(s), \tag{67}$$

where $Q_{n,m}(s)$ represents the moment generating function of $Y_{n,m}$, the conditional optimization time when beginning with a random input (each bit being 1 with probability $\frac{1}{2}$) having $n-m$ leading 1s.

Throughout this section, the probability p still carries the same meaning from Algorithm $(1+1)$-EA and $q=1-p$, namely, each bit has flipping probability p at each stage, and one independently of all the others.

Theorem 8. *Assume $p=\frac{c}{n}$, $c>0$. The random variables Y_n are asymptotically normally distributed*

$$\mathbb{P}\left(\frac{Y_n-\nu_n}{\varsigma_n}\leqslant x\right)\to\Phi(x) \qquad (x\in\mathbb{R}),$$

with the mean $\nu_n=\mathbb{E}(Y_n)$ and the variance $\varsigma_n=\mathbb{V}(Y_n)$ asymptotic to

$$\begin{aligned}\nu_n &= \frac{e^c-1}{2c^2}n^2+\frac{(c-2)e^c+2}{4c}n+O(1)\\ \varsigma_n &= \frac{e^{2c}-1}{8c^3}n^3+\frac{3e^{2c}(2c-3)-8e^c+17}{16c^2}n^2+O(n),\end{aligned} \tag{68}$$

respectively. Here $\Phi(x) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\,\mathrm{d}t$ denotes the standard normal distribution function.

In particular, we also have, by replacing the exact mean and variance by the corresponding asymptotic approximations

$$\mathbb{P}\left(\frac{Y_n-\frac{e^c-1}{2c^2}n^2}{\sqrt{\frac{e^{2c}-1}{8c^3}n^3}}\leqslant x\right)\to\Phi(x) \qquad (x\in\mathbb{R});$$

see Figure 9 for some graphical renderings of the distributions of Y_n.

Our method of proof consists in deriving first a more manageable expression for $Q_{n,m}$, and then work on the corresponding characteristic functions.

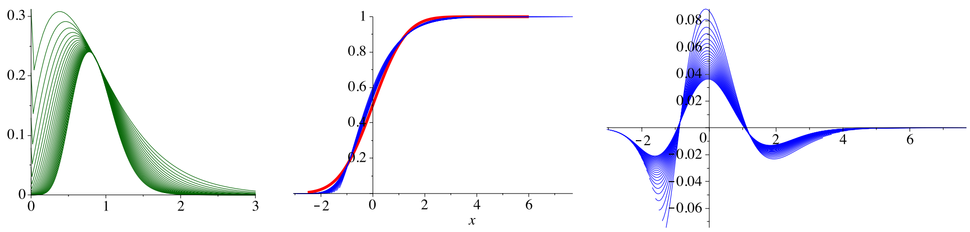


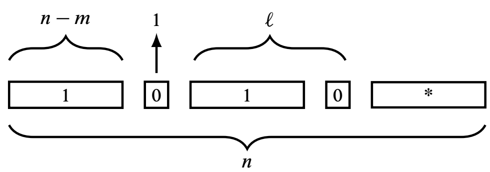
Figure 9: *Distributions of Y_n for $n = 5, \ldots, 20$ when the mutation rate is $p = \frac{1}{n}$: the density functions scaled by n^2 (left), distribution functions (middle), and the differences between the distribution function of $\frac{Y_n - \nu_n}{\varsigma_n}$ and $\Phi(x)$ (right).*

Lemma 12. *The moment generating function $Q_{n,m}(s)$ of $Y_{n,m}$ satisfies the recurrence relation*

$$Q_{n,m}(s) = \frac{pq^{n-m}e^s}{1-(1-pq^{n-m})e^s}\left(2^{1-m} + \sum_{1\leqslant \ell<m} \frac{Q_{n,\ell}(s)}{2^{m-\ell}}\right), \tag{69}$$

for $1 \leqslant m \leqslant n$, where $q = 1-p$.

Proof. The probability of jumping from a state with $n-m$ leading 1s to another state with $n-m+\ell$ leading 1s is given by



$$(1-p)^{n-m} \cdot p \cdot 2^{-\ell} \qquad (1 \leqslant \ell < m),$$

which corresponds to the situation when the first $n-m$ bits do not toggle their values, the $(n-m+1)$st bit toggles (from 0 to 1), together with the following $\ell-1$ bits also being 1. When $\ell = m$, the probability becomes

$$(1-p)^{n-m} \cdot p \cdot 2^{-m+1}.$$

We thus obtain the recurrence relation

$$Q_{n,m}(s) = pq^{n-m}e^s\left(2^{1-m} + \sum_{1\leqslant \ell<m} \frac{Q_{n,\ell}(s)}{2^{m-\ell}}\right) + (1-pq^{n-m})e^s Q_{n,m}(s), \tag{70}$$

which implies (69). ▮

A finite-product representation for $Q_{n,m}(s)$. Unlike the case of OneMax (see (4)), the recurrence relation (69) can be solved explicitly as follows.

Proposition 2. *The moment generating function $Q_{n,m}(s)$ of $Y_{n,m}$ has the closed-form*

$$Q_{n,m}(s) = \frac{1}{1-\frac{1-e^{-s}}{pq^{n-m}}} \prod_{1\leqslant j<m} \frac{1-\frac{1-e^{-s}}{2pq^{n-j}}}{1-\frac{1-e^{-s}}{pq^{n-j}}}, \tag{71}$$

for $m \geqslant 1$.

Let

$$G_m(t) := \frac{pq^{n-m}t}{1-(1-pq^{n-m})t}$$

denote the probability generating function of a geometric distribution $\mathrm{Geo}(pq^{n-m})$ with parameter pq^{n-m} and support $\{1, 2, \ldots\}$.

Corollary 8. *The random variables $Y_{n,m}$ can be decomposed as the sum of m independent random variables*

$$Y_{n,m} \stackrel{d}{=} Z_{n,m}^{[0]} + \cdots + Z_{n,m}^{[m-1]}, \tag{72}$$

where $Z_{n,m}^{[0]} \sim Geo(pq^{n-m})$ and the $Z_{n,m}^{[j]}$ are mixture of $Geo(pq^{n-j})$

$$\mathbb{E}\left(t^{Z_{n,m}^{[j]}}\right) = \frac{1}{2} \cdot \frac{1-(1-2pq^{n-j})t}{1-(1-pq^{n-j})t} \qquad (j = 1, \ldots, m-1).$$

For the proof of Proposition 2, we need a lemma.

Lemma 13. *The solution to the recurrence relation*

$$a_m = b_m + \sum_{1 \leqslant \ell < m} \frac{a_\ell}{2^{m-\ell}} \qquad (m \geqslant 1),$$

with $a_0 = b_0$, is given by the closed-form expression

$$a_m = b_m + \frac{1}{2} \sum_{1 \leqslant j < m} b_j \qquad (m \geqslant 0). \tag{73}$$

Proof. The corresponding generating functions $f(z) := \sum_{m\geqslant 1} a_m z^m$ and $g(z) := \sum_{m\geqslant 1} b_m z^m$ satisfy the equation

$$f(z) = g(z) + \frac{z}{2-z} f(z),$$

or

$$f(z) = \frac{1-\frac{z}{2}}{1-z} g(z).$$

This proves (73). ∎

Proof of Proposition 2. Let $\omega := \frac{1}{pq^n}(1-e^{-s})$. We start with the recurrence (from (70))

$$Q_{n,m}(s) = \omega q^m Q_{n,m}(s) + 2^{1-m} + \sum_{1 \leqslant \ell < m} \frac{Q_{n,\ell}(s)}{2^{m-\ell}},$$

which, by (73) with $b_m = \omega q^m Q_{n,m}(s) + 2^{1-m}$, has the alternative form

$$Q_{n,m}(s) = 1 + \omega q^m Q_{n,m}(s) + \frac{\omega}{2} \sum_{1 \leqslant h < m} q^h Q_{n,h}(s), \tag{74}$$

since $2^{1-m} + \frac{1}{2}\sum_{1\leqslant j<m} 2^{1-j} = 1$. From (74), we see that the bivariate generating function $Q_n(z, s) := \sum_{m\geqslant 1} Q_{n,m}(s) z^m$ of $Q_{n,m}(s)$ satisfies

$$Q_n(z, s) = \frac{z}{1-z} + \omega Q_n(qz, s) + \frac{\omega}{2} \cdot \frac{z}{1-z} Q_n(qz, s),$$

which implies the simpler functional equation

$$Q_n(z, s) = \frac{z}{1-z} + \omega \frac{1-\frac{z}{2}}{1-z} Q_n(qz, s).$$

Multiplying both sides by $1-z$ gives

$$(1-z)Q_n(z,s) = z + \omega\left(1-\frac{z}{2}\right)Q_n(qz,s),$$

implying the relation

$$\frac{Q_{n,m}(s)}{Q_{n,m-1}(s)} = \frac{1-\frac{1}{2}\omega q^{m-1}}{1-\omega q^m} \qquad (m \geqslant 2).$$

Accordingly, we obtain the closed-form expression (71). ∎

By (72), the mean $\nu_{n,m}$ of $Y_{n,m}$ is given by

$$\begin{aligned}\nu_{n,m} &= \sum_{0\leqslant j<m} \mathbb{E}\left(Z_{n,m}^{[j]}\right) = \frac{1}{pq^{n-m}} + \frac{1}{2}\sum_{1\leqslant j<m}\frac{1}{pq^{n-j}} \\ &= \frac{1}{pq^{n-1}}\left(\frac{1-q^{m-1}}{2p} + q^{m-1}\right).\end{aligned} \tag{75}$$

Similarly, the variance $\varsigma_{n,m}^2$ of $Y_{n,m}$ satisfies

$$\begin{aligned}\varsigma_{n,m}^2 &= \sum_{0\leqslant j<m} \mathbb{V}\left(Z_{n,m}^{[j]}\right) = \frac{3-2pq^{n-m}}{8(pq^{n-m})^2} + \frac{1}{2}\sum_{1\leqslant j<m}\frac{1-pq^{n-j}}{(pq^{n-j})^2} \\ &= -\nu_{n,m} + \frac{3q^2-(4q^2-1)q^{2m}}{4p^3(1+q)q^{2n}}.\end{aligned} \tag{76}$$

These expressions are valid for any p but the most interesting cases are when $p=\frac{c}{n}$ because the time complexity grows with c in an exponential rate.

Corollary 9. *Assume that $p=\frac{c}{n}$, where $c=o(\sqrt{n})$. Then, uniformly for $0\leqslant \alpha := \frac{m}{n} \leqslant 1$,*

$$\nu_{n,m} = \frac{e^c}{2c^2}\left(1-e^{-c\alpha}\right)n^2 + \frac{e^c}{4c}\left(c-2+e^{-c\alpha}\left(4-c+c\alpha\right)\right)n + O\left(c(c+1)e^c\right), \tag{77}$$

and

$$\varsigma_{n,m}^2 = \frac{3e^{2c}}{8c^3}\left(1-e^{-2c\alpha}\right)n^3 + O\left(c^{-2}e^{2c}(1+c)n^2\right). \tag{78}$$

These two approximations are straightforward from (75) and (76).

Proof. (of Theorem 8) By (67) and (75), we have

$$\nu_n = \sum_{1\leqslant m\leqslant n} 2^{-n+m-1}\nu_{n,m} = \frac{q}{2p^2}\left(q^{-n}-1\right),$$

and then the first estimate in (68) follows. Similarly, by (76),

$$\varsigma_n^2 = \sum_{1\leqslant m\leqslant n} 2^{-n+m-1}\mathbb{E}(Y_{n,m}^2) - \nu_n^2 = \frac{3q^2}{4p^3(1+q)}\left(q^{-2n}-1\right) - \nu_n,$$

and the second estimate in (68) also follows.

For the asymptotic normality, we consider the characteristic function

$$\mathbb{E}\big(e^{\frac{Y_n-\nu_n}{\varsigma_n}i\theta}\big) = 2^{-n} + \sum_{1\leqslant m\leqslant n} 2^{-n+m-1} Q_{n,m}\big(\tfrac{i\theta}{\varsigma_n}\big)e^{-\frac{\nu_n}{\varsigma_n}i\theta}.$$

We split the sum into two parts: $0 \leqslant n-m \leqslant 2\log_2 n$ and $1 \leqslant m < n-2\log_2 n$. Observe that when $n-m \leqslant 2\log_2 n$, we have the uniform estimate

$$\nu_n - \nu_{n,m} = O(n|n-m+1|) \quad \text{and} \quad \varsigma_n^2 - \varsigma_{n,m}^2 = O\left(n^2|n-m+1|\right),$$

by (77) and (78). We then have the local expansion

$$Q_{n,m}\left(\frac{i\theta}{\varsigma_n}\right)e^{-\frac{\nu_n}{\varsigma_n}i\theta} = \exp\left(\frac{\nu_{n,m}-\nu_n}{\varsigma_n}\,i\theta - \frac{\varsigma_{n,m}^2}{2\varsigma_n^2}\,\theta^2 + O\left(\frac{|\theta|^3}{\sqrt{n}}\right)\right),$$

uniformly for $|\theta| = o(n^{\frac{1}{6}})$. Thus

$$\begin{aligned}
Q_{n,m}\left(\frac{i\theta}{\varsigma_n}\right)e^{-\frac{\nu_n}{\varsigma_n}i\theta} &= \exp\left(-\frac{\theta^2}{2} + O\left(\frac{|n-m+1|}{\sqrt{n}}\,|\theta| + \frac{|n-m+1|}{n}\,\theta^2\right)\right)\\
&= \exp\left(-\frac{\theta^2}{2} + O\left(n^{-\frac{1}{2}}(\log n)|\theta| + n^{-1}(\log n)|\theta|^2\right)\right)\\
&= e^{-\frac{\theta^2}{2}}(1+o(1)),
\end{aligned}$$

uniformly in m. Consequently,

$$\sum_{n-2\log_2 n\leqslant m\leqslant n} 2^{-n+m-1}Q_{n,m}\left(\frac{i\theta}{\varsigma_n}\right)e^{-\frac{\nu_n}{\varsigma_n}i\theta} = e^{-\frac{\theta^2}{2}}(1+o(1)).$$

The remaining part is negligible since $|Q_{n,m}(e^{\frac{i\theta}{\sigma}})| \leqslant 1$ and

$$\sum_{1\leqslant m\leqslant n-2\log_2 n} 2^{-n+m-1}Q_{n,m}\left(\frac{i\theta}{\varsigma_n}\right)e^{-\frac{\nu_n}{\varsigma_n}i\theta} = O\left(\sum_{m>2\log_2 n} 2^{-m}\right) = O\left(2^{-2\log_2 n}\right) = O\big(n^{-2}\big).$$

We conclude that

$$\mathbb{E}\big(e^{\frac{Y_n-\nu_n}{\varsigma_n}i\theta}\big) \to e^{-\frac{\theta^2}{2}},$$

which implies the convergence in distribution of $\frac{Y_n-\nu_n}{\varsigma_n}$ to the standard normal distribution by Lévy's continuity theorem. ∎

8 Conclusion

Recurrence relations find their ubiquitous appearance in computer algorithms, and numerous techniques have been developed in the literature for a better understanding of their asymptotic behaviors. As a rich source of recurrences of diverse nature, evolutionary algorithms have mostly been analyzed via tools from probability theory. We showed in this paper the possibility of solving directly the recurrences via tools from asymptotic and complex analysis. It is common that such tools are often applicable under stronger settings, yet they achieve higher precision than the others whenever they apply, as Odlyzko commented in (Odlyzko, 1995, p. 1152)

"*Analytic methods are extremely powerful, and when they apply, they often yield estimates of unparalleled precision.*"

We analyzed in detail the performance of the $(1+1)$-EA under two simple fitness functions OneMax and LeadingOnes, and fully characterized the first two moments and the limit laws. Similar techniques based on recurrence relations may also be applied to other problems of a recursive nature; one of the simplest (even simpler than the $(1+1)$-EA) such algorithms is the *Randomized Local Search* or the *One-Bit-Flip* heuristic in which only one bit mutates at each stage; see Garnier et al. (1999); Ladret (2005); Doerr and Lengler (2016) for details. Such a simple heuristic yields the same type of stochastic behaviors for the optimization time as the $(1+1)$-EA (see Table 1) but with different (smaller) coefficients for the means and the variances. For completeness and ease of comparison, we list the corresponding results in Table 2; see He and Yao (2002, 2003) for other potential algorithms.

Properties	OneMax (X_n)	LeadingOnes (Y_n)
Reference	Garnier et al. (1999)	Ladret (2005)
Probability generating function	$\sum_{0\leqslant m\leqslant n} \frac{\binom{n}{m}}{2^n} \prod_{1\leqslant k\leqslant m} \frac{\frac{k}{n}t}{1-(1-\frac{k}{n})t}$	$\left(\frac{1}{2}+\frac{\frac{1}{n}t}{2(1-(1-\frac{1}{n})t)}\right)^n$
Mean $\sim$	$(\log\frac{n}{2}+\gamma)n+\frac{1}{2}$	$\frac{1}{2}n^2$
Variance $\sim$	$\frac{\pi^2}{6}n^2-(\log\frac{n}{2}+\gamma+1)n$	$\frac{3}{4}n^3$
Limit law	Gumbel distribution $\mathbb{P}\left(\frac{X_n}{n}-\log\frac{n}{2}\leqslant x\right)$ $\to e^{-e^{-x}}$	Gaussian distribution $\mathbb{P}\left(\frac{Y_n-\frac{1}{2}n^2}{\sqrt{\frac{3}{4}n^3}}\leqslant x\right)$ $\to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,\mathrm{d}t$
Approach	method of moments	characteristic function

Table 2: *A summary of known results for the optimization time of One-Bit-Flip under the* OneMax *(X_n) and* LeadingOnes *(Y_n) fitness function, respectively, when starting from a random initial state and when the mutation probability is* $\frac{1}{n}$.

Acknowledgements

We thank the editors and the seven referees for their constructive comments and helpful suggestions in improving the presentation of the paper. In particular, one referee's critical examination of the notion of uniformity led to further clarification of the proofs used in establishing the limit laws of $X_{n,m}$ when m is bounded.

References

Auger, A. and Doerr, B., editors (2011). *Theory of Randomized Search Heuristics: Foundations and Recent Developments*. World Scientific, Singapore.

Bäck, T. (1992). The interaction of mutation rate, selection, and self-adaptation within a genetic algorithm. In *Parallel Problem Solving from Nature (PPSN-II)*, pages 85–94, Brussels, Belgium.

Beyer, H.-G., Schwefel, H.-P., and Wegener, I. (2002). How to analyse evolutionary algorithms. *Theoretical Computer Science*, 287:101–130.

Böttcher, S., Doerr, B., and Neumann, F. (2010). Optimal fixed and adaptive mutation rates for the leadingones problem. In *Parallel Problem Solving from Nature (PPSN XI, Part I)*, volume 6238 of *Lecture Notes in Computer Science*, pages 1–10.

Coello Coello, C. A. (2006). Evolutionary multiobjective optimization: a historical view of the field. *IEEE Computational Intelligence Magazine*, 1:28–36.

Comtet, L. (1974). *Advanced Combinatorics, The Art of Finite and Infinite Expansions*. D. Reidel Publishing Co., Dordrecht.

de Bruijn, N. G. (1981). *Asymptotic methods in analysis*. Dover Publications, Inc., New York, third edition.

Deb, K. (2001). *Multi-Objective Optimization Using Evolutionary Algorithms*. John Wiley & Sons.

Doerr, B. and Doerr, C. (2014). The impact of random initialization on the runtime of randomized search heuristics. In *Genetic and Evolutionary Computation (GECCO 2014)*, pages 1375–1382. ACM.

Doerr, B., Doerr, C., and Ebel, F. (2013). Lessons from the black-box: fast crossover-based genetic algorithms. In *Genetic and Evolutionary Computation (GECCO 2013)*, pages 781–788. ACM.

Doerr, B., Fouz, M., and Witt, C. (2010a). Quasirandom evolutionary algorithms. In *Genetic and Evolutionary Computation (GECCO 2010)*, pages 1457–1464. ACM Press.

Doerr, B., Fouz, M., and Witt, C. (2011). Sharp bounds by probability-generating functions and variable drift. In *Genetic and Evolutionary Computation (GECCO 2011)*, pages 2083–2090. ACM Press.

Doerr, B., Johannsen, D., and Winzen, C. (2010b). Drift analysis and linear functions revisited. In *IEEE Congress on Evolutionary Computation*, pages 1–8.

Doerr, B., Johannsen, D., and Winzen, C. (2012). Multiplicative drift analysis. *Algorithmica*, 64:673–697.

Doerr, C. and Lengler, J. (2016). Onemax in black-box models with several restrictions. *Algorithmica*, pages 1–31.

Droste, S., Jansen, T., and Wegener, I. (1998). A rigorous complexity analysis of the $(1 + 1)$ evolutionary algorithm for linear functions with boolean inputs. In *IEEE International Conference on Evolutionary Computation ICEC'98*, pages 499–504, Piscataway, NJ. IEEE Press.

Droste, S., Jansen, T., and Wegener, I. (2002). On the analysis of the $(1 + 1)$ evolutionary algorithm. *Theoretical Computer Science*, 276:51–81.

Flajolet, P. and Sedgewick, R. (2009). *Analytic Combinatorics*. Cambridge University Press, Cambridge.

Garnier, J., Kallel, L., and Schoenauer, M. (1999). Rigorous hitting times for binary mutations. *Evolutionary Computation*, 7:167–203.

He, J. and Yao, X. (2001). Drift analysis and average time complexity of evolutionary algorithms. *Artificial Intelligence*, 127:57–85.

He, J. and Yao, X. (2002). From an individual to a population: An analysis of the first hitting time of population-based evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 6:495–511.

He, J. and Yao, X. (2003). Towards an analytic framework for analysing the computation time of evolutionary algorithms. *Artificial Intelligence*, 145(1):59–97.

Horn, J. (1997). Multicriterion decision making. In *Handbook of Evolutionary Computation*, pages F1.9:1–F1.9:15. IOP Publishing Ltd. and Oxford University Press.

Hwang, H.-K. (1998). On convergence rates of central limit theorems for combinatorial structures. *European Journal of Combinatorics*, 19:329–343.

Ladret, V. (2005). Asymptotic hitting time for a simple evolutionary model of protein folding. *Journal of Applied Probability*, 42:39–51.

Lehre, P. K. and Witt, C. (2014). Concentrated hitting times of randomized search heuristics with variable drift. In *Algorithms and Computation*, pages 686–697. Springer.

Mühlenbein, H. (1992). How genetic algorithms really work: mutation and hill-climbing. In *Parallel Problem Solving from Nature (PPSN II)*, pages 15–26.

Neumann, F. and Witt, C. (2009). Runtime analysis of a simple ant colony optimization algorithm. *Algorithmica*, 54:243–255.

Neumann, F. and Witt, C. (2010). *Bioinspired Computation in Combinatorial Optimization – Algorithms and Their Computational Complexity*. Springer.

Odlyzko, A. M. (1995). Asymptotic enumeration methods. In *Handbook of Combinatorics, Vol. 1, 2*, pages 1063–1229. Elsevier, Amsterdam.

Sloane, N. J. A. (since 1995). *Online Encyclopedia of Integer Sequences (OEIS)*. https://oeis.org.

Sudholt, D. (2010). General lower bounds for the running time of evolutionary algorithms. In *Parallel Problem Solving from Nature (PPSN X)*, volume 6238 of *Lecture Notes in Computer Science*, pages 124–133.

Sudholt, D. (2013). A new method for lower bounds on the running time of evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 17(3):418–435.

Sudholt, D. and Witt, C. (2010). Runtime analysis of a binary particle swarm optimizer. *Theoretical Computer Science*, 411:2084–2100.

van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press, Cambridge.

Witt, C. (2013). Tight bounds on the optimization time of a randomized search heuristic on linear functions. *Combin. Probab. Comput.*, 22(2):294–318.

Witt, C. (2014). Fitness levels with tail bounds for the analysis of randomized search heuristics. *Information Processing Letters*, 114(1-2):38–41.

Wong, R. (2014). Asymptotics of linear recurrences. *Analysis and Applications (Singapore)*, 12(4):463–484.

Zhou, D., Luo, D., Lu, R., and Han, Z. (2012). The use of tail inequalities on the probable computational time of randomized search heuristics. *Theoretical Computer Science*, 436:106–117.

Appendices

A Some properties of $S_r(z)$

We collected here some interesting expressions for $S_r(z)$.

We begin with proving that all S_r can be expressed in terms of S_0 and the two modified Bessel functions

$$\bar{I}_0(\alpha) := I_0\left(2\sqrt{\alpha(1-\alpha)}\right) = \sum_{\ell\geqslant 0} \frac{\alpha^\ell(1-\alpha)^\ell}{\ell!\ell!},$$
$$\bar{I}_1(\alpha) := \sqrt{\frac{\alpha}{1-\alpha}} I_1\left(2\sqrt{\alpha(1-\alpha)}\right) = \sum_{\ell\geqslant 1} \frac{\alpha^\ell(1-\alpha)^{\ell-1}}{\ell!(\ell-1)!}.$$

The starting point is the obvious relation ($E_r(z) := \sum_{\ell\geqslant 1} \ell^r z^{\ell-1}$)

$$E_r(z) = zE'_{r-1}(z) + E_{r-1}(z) \qquad (r \geqslant 1).$$

Applying the integral representation (32) and integration by parts, we have

$$S_r(\alpha) = \frac{1}{2\pi i}\oint_{|z|=c} \left(\frac{\alpha}{z} - (1-\alpha)z\right) E_{r-1}(z) e^{\frac{\alpha}{z}+(1-\alpha)z}\, \mathrm{d}z.$$

By the same argument used for Corollary 3, we deduce the recurrence

$$S_r(\alpha) = \alpha\bar{I}_0(\alpha) + \sum_{0\leqslant j<r}\binom{r-1}{j}\left(\alpha + (-1)^{r-j}(1-\alpha)\right) S_j(\alpha), \tag{79}$$

for $r \geqslant 2$ with

$$S_1(\alpha) = (2\alpha-1)S_0(\alpha) + \alpha\bar{I}_0(\alpha) + (1-\alpha)\bar{I}_1(\alpha).$$

A closed-form expression can be obtained for the recurrence (79) but it is very messy. More precisely, let $f(z) := \sum_{r\geqslant 0} S_r(\alpha)\frac{z^r}{r!}$. Then f satisfies the first-order differential equation

$$f'(z) = (\alpha e^z - (1-\alpha)e^{-z})\, f(z) + \alpha\bar{I}_0(\alpha) + (1-\alpha)\bar{I}_1(\alpha).$$

The solution to the differential equation with the initial condition $f(0) = S_0(\alpha)$ is given by

$$f(z) = S_0(\alpha)e^{\alpha(e^z-e^{-z})+e^{-1}-1} + e^{\alpha e^z+(1-\alpha)e^{-z}}\int_0^z \left(\alpha\bar{I}_0(\alpha)e^u + (1-\alpha)\bar{I}_1(\alpha)\right) e^{-\alpha e^u-(1-\alpha)e^{-u}}\, \mathrm{d}u.$$

This implies that $S_r(\alpha)$ has the general form

$$S_r(\alpha) = p_r^{[0]}(\alpha)\bar{I}_0(\alpha) + p_r^{[1]}(\alpha)\bar{I}_1(\alpha) + p_r^{[2]}(\alpha)S_0(\alpha) \qquad (r \geqslant 1),$$

where the $p_r^{[i]}$ are polynomials of α of degree r. Closed-form expressions can be derived but are less simple than the recurrence (79) for small values of r.

On the other hand, the same argument also leads to

$$S_r'(\alpha) = \bar{I}_0(\alpha) + \sum_{0\leqslant j<r} \binom{r}{j} \big(1-(-1)^{r-j}\big) S_j(\alpha)(\qquad (r \geqslant 1).$$

In particular, $S_1'(\alpha) = \bar{I}_0(\alpha) + 2S_0(\alpha)$. Note that

$$S_0'(\alpha) = \bar{I}_0(\alpha) + \bar{I}_1(\alpha),$$

implying that

$$S_0(\alpha) = \int_0^\alpha \big(\bar{I}_0(u) + \bar{I}_1(u)\big)\,\mathrm{d}u.$$

This in turn gives

$$S_1(\alpha) = \int_0^\alpha \big((1+2(\alpha-u))\bar{I}_0(u) + 2(\alpha-u)\bar{I}_1(u)\big)\,\mathrm{d}u.$$

This expression can be further simplified by taking second derivative with respect to α of the integral representation

$$S_1(\alpha) = \frac{1}{2\pi i}\oint_{|z|=c} \frac{e^{\frac{\alpha}{z}+(1-\alpha)z}}{(1-z)^2}\,\mathrm{d}z,$$

giving

$$S_1''(\alpha) = 2\bar{I}_0(\alpha) + \alpha^{-1}\bar{I}_1(\alpha),$$

which implies that (with $S_1(0) = 0$, $S_1'(0) = 1$)

$$S_1(\alpha) = \int_0^\alpha (\alpha-u)\left(2\bar{I}_0(u) + u^{-1}\bar{I}_1(u)\right)\mathrm{d}u$$

Similarly, since $S_2' = I_0 + 4S_1$, we have

$$S_2(\alpha) = \int_0^\alpha (1+4(\alpha-u)(\alpha-u+1))\bar{I}_0(u)\,\mathrm{d}u + 4\int_0^\alpha (\alpha-u)^2\bar{I}_1(u)\,\mathrm{d}u.$$

These expressions show not only the intimate connections of S_r to Bessel functions but also their rich algebraic aspects.

We now consider $S_r(1-\alpha)$. By the same integral representation and a change of variables, we see that, for $r \geqslant 1$,

$$(-1)^r S_r(\alpha) + S_r(1-\alpha) = [z^0]E_r(1-z)e^{\frac{\alpha}{1-z}+(1-\alpha)(1-z)}.$$

Now

$$E_r(1-z) = r![w^r]\frac{e^w}{1-(1-z)e^w} = \sum_{0\leqslant j\leqslant r} (-1)^{r+j} j!\,\mathrm{Stirling}_2(r,j) z^{-j-1}.$$

Thus we deduce the identity (for $r \geqslant 1$)

$$\begin{aligned}&(-1)^r S_r(\alpha) + S_r(1-\alpha)\\ &\quad = e\sum_{0\leqslant \ell\leqslant r} (-1)^{r+\ell}\ell!\,\mathrm{Stirling}_2(r,\ell) \sum_{\substack{0\leqslant h\leqslant \ell\\ 0\leqslant j<\frac{h}{2}}} \binom{h-j-1}{j-1}\frac{(2\alpha-1)^{\ell-h}\alpha^j}{(\ell-h)!\,j!}\end{aligned}$$

or

$$(-1)^r S_r(\alpha) + S_r(1-\alpha) \\ = e \sum_{0 \leqslant \ell \leqslant r} (-1)^{r+\ell} \ell! \operatorname{Stirling}_2(r, \ell) \left(\frac{(\alpha-1)^\ell}{\ell!} + \sum_{\substack{0 \leqslant h \leqslant \ell \\ 0 \leqslant j < h}} \binom{h-1}{j} \frac{\alpha^{h-j}(\alpha-1)^{\ell-h}}{(\ell-h)!(h-j)!} \right).$$

Note that for $r = 0$

$$S_0(\alpha) + S_0(1-\alpha) = e - \bar{I}_0(\alpha).$$

In particular, this gives $S(\frac{1}{2}) = \frac{1}{2}(e - I_0(1)) \approx 0.726107$. For $r \geqslant 1$

$$\begin{aligned} S_1(\alpha) - S_1(1-\alpha) &= e(2\alpha - 1) \\ S_2(\alpha) + S_2(1-\alpha) &= e(4\alpha^2 - 4\alpha + 2) \\ S_3(\alpha) - S_3(1-\alpha) &= e(8\alpha^3 - 12\alpha^2 + 14\alpha - 5) \\ S_4(\alpha) + S_4(1-\alpha) &= e(16\alpha^4 - 32\alpha^3 + 64\alpha^2 - 48\alpha + 15). \end{aligned}$$

On the other hand, we also have the limiting behaviors

$$\lim_{\alpha \to 0} \frac{S_r(\alpha)}{\alpha} = 1,$$

and

$$\lim_{\alpha \to 1} S_r(\alpha) = \sum_{\ell \geqslant 1} \frac{\ell^r}{\ell!} = \{e-1, e, 2e, 5e, 15e, \cdots\}.$$

Without the first term, the right-hand side is, up to e, the Bell numbers (all partitions of a set; Sequence A000110 in Sloane's Encyclopedia of Integer Sequences Sloane (1995)).

B The refined approximation (39) to the asymptotics of $\mu^*_{n,m}$

We consider now the difference

$$\Delta^*_{n,m} := \mu^*_{n,m} - (H_m + \phi_1(\alpha)) - \frac{1}{n}(b_1 H_m + \phi_2(\alpha)),$$

and will determine the constant b_1 and the function $\phi_2(z)$ such that

$$\Delta^*_{n,m} = O(n^{-2} H_m), \tag{80}$$

uniformly for $1 \leqslant m \leqslant n$, which then proves Theorem 2. By (16), $\Delta^*_{n,m}$ satisfies, for $1 \leqslant m \leqslant n$, the recurrence

$$\sum_{1 \leqslant \ell \leqslant m} \lambda^*_{n,m,\ell} \left(\Delta^*_{n,m} - \Delta^*_{n,m-\ell} \right) = E_2(n, m), \tag{81}$$

where

$$E_2(n,m) := \frac{1}{n} - \sum_{1 \leqslant \ell \leqslant m} \lambda^*_{n,m,\ell} (H_m - H_{m-\ell}) - \sum_{1 \leqslant \ell \leqslant m} \lambda^*_{n,m,\ell} \left(\phi_1\left(\frac{m}{n}\right) - \phi_1\left(\frac{m-\ell}{n}\right) \right) \\ - \frac{b_1}{n} \sum_{1 \leqslant \ell \leqslant m} \lambda^*_{n,m,\ell} (H_m - H_{m-\ell}) - \frac{1}{n} \sum_{1 \leqslant \ell \leqslant m} \lambda^*_{n,m,\ell} \left(\phi_2\left(\frac{m}{n}\right) - \phi_2\left(\frac{m-\ell}{n}\right) \right).$$

In particular, $\Delta^*_{n,0} = -\frac{\phi_2(0)}{n}$.

The hard part here is to derive an asymptotic expansion for $E_2(n,m)$ that holds uniformly for $1 \leqslant m \leqslant n$ as $n \to \infty$. To that purpose, we first extend Lemma 5 by using a Taylor expansion of third order for a $C^\infty[0,1]$-function $\phi(z)$, which then gives, uniformly for $1 \leqslant m \leqslant n$,

$$\begin{aligned}
&\sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell}\left(\phi\Big(\frac{m}{n}\Big) - \phi\Big(\frac{m-\ell}{n}\Big)\right)\\
&\quad= \frac{\phi'(\alpha)}{n}\sum_{1\leqslant \ell\leqslant m} \ell\lambda^*_{n,m,\ell} - \frac{\phi''(\alpha)}{2n^2}\sum_{1\leqslant \ell\leqslant m} \ell^2\lambda^*_{n,m,\ell} + O\big(n^{-3}\big)\\
&\quad= \frac{\phi'(\alpha)}{n}\, S_1(\alpha) + \frac{1}{2n^2}\big(2\phi'(\alpha)U_1(\alpha) - \phi''(\alpha)S_2(\alpha)\big) + O\big(n^{-3}\big), \qquad (82)
\end{aligned}$$

where we used Corollary 3.

On the other hand, by Corollary 6, we have

$$H_m - H_{m-\ell} = \frac{\ell}{m} + \frac{[\![m \geqslant 2]\!]\,\ell(\ell-1)}{2m(m-1)} + O\Big(\frac{\ell(\ell-1)(\ell-2)}{m^3}\Big),$$

for $0 \leqslant \ell \leqslant \frac{m}{2}$ and $m \geqslant 1$. By the same argument used above for (45), we get the expansion

$$\begin{aligned}
&\sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell}\,(H_m - H_{m-\ell})\\
&\quad= \sum_{1\leqslant \ell\leqslant m}\left(\frac{\ell}{m} + \frac{[\![m \geqslant 2]\!]\,\ell(\ell-1)}{2m(m-1)} + O\Big(\frac{\ell(\ell-1)(\ell-2)}{m^3}\Big)\right)\lambda^*_{n,m,\ell} + O\big(n^{-3}\big)\\
&\quad= \frac{1}{m}\sum_{1\leqslant \ell\leqslant m} \ell\lambda^*_{n,m,\ell} + \frac{[\![m \geqslant 2]\!]}{2m(m-1)}\sum_{1\leqslant \ell\leqslant m} \ell(\ell-1)\lambda^*_{n,m,\ell} + O\big(n^{-3}\big)\\
&\quad= \frac{S_1(\alpha)}{\alpha n} + \frac{1}{2n^2}\left(\frac{U_1(\alpha)}{\alpha} + \frac{S_2(\alpha) - S_1(\alpha)}{\alpha^2}\right) - \frac{[\![m = 1]\!]}{2n^2} + O\big(n^{-3}\big), \qquad (83)
\end{aligned}$$

which holds uniformly for $1 \leqslant m \leqslant n$. Note that for $m = 1$ a correction term is needed; more correction terms have to be introduced in more refined expansions (see Appendix C).

Combining the estimates (82) (with $\phi = \phi_1, \phi_2$) and (83), we see that

$$E_2(n,m) = \frac{J_1(\alpha)}{n} + \frac{J_2(\alpha)}{2n^2} + \frac{[\![m = 1]\!]}{2n^2} + O\big(n^{-3}\big),$$

uniformly for $1 \leqslant m \leqslant n$, where

$$\begin{aligned}
J_1(z) &= 1 - \frac{S_1(z)}{z} - \phi_1'(z)S_1(z),\\
J_2(z) &= -\frac{S_2(z) - S_1(z)}{z^2} - \frac{2b_1}{z}\,S_1(z) - \left(\frac{2}{z} + 2\phi_1'(z)\right)U_1(z)\\
&\qquad + \phi_1''(z)S_2(z) - 2\phi_2'(z)S_1(z).
\end{aligned}$$

Obviously, $J_1(z) = 0$ because $\phi_1'(z) = \frac{1}{S_1(z)} - \frac{1}{z}$. To determine b_1 and ϕ_2, we observe that

$$\lim_{z\to 0} J_2(z) = \lim_{z\to 0}\left(-\frac{S_2(z) - S_1(z)}{z^2} - \frac{2b_1}{z}\,S_1(z) - \frac{2}{z}\,U_1(z)\right) = 2b_1 - 2,$$

where we used the relation $U_1(\alpha) = -S_0(\alpha) - \frac{1}{2}S_1(\alpha)$ (see (31)). In order that $E_2 = o(n^{-2})$ uniformly for $1 \leqslant m \leqslant n$, we need $2b_1 - 2 = 0$, so that $b_1 = 1$.

Now the equation $J_2(z) = 0$ also implies, by (31), that

$$\phi_2'(z) = -\frac{S_1'(z)S_2(z)}{2S_1^3(z)} + \frac{S_0(z)}{S_1(z)^2} + \frac{1}{2S_1(z)} + \frac{1}{2z^2} - \frac{1}{z}. \tag{84}$$

With these choices of b_1 and $\phi_2(z)$, we have

$$E_2(n,m) = \frac{[\![m=1]\!]}{2n^2} + O\big(n^{-3}\big),$$

uniformly for $1 \leqslant m \leqslant n$.

The exact solution to the differential equation (84) requires the constant term $\phi_2(0)$, which we have not yet specified. To specify this value, we take $m = 1$ in (80) and then obtain, by the recurrence (81),

$$\Delta^*_{n,1} = \Delta^*_{n,0} + \frac{E_2(n,1)}{\Lambda^*_{n,1}} = -\frac{\phi_2(0)}{n} + nE_2(n,1) = -\frac{\phi_2(0)}{n} + \frac{1}{2n} + O\big(n^{-2}\big).$$

This entails the choice $\phi_2(0) = \frac{1}{2}$ in order that $\Delta^*_{n,1} = O\big(n^{-2}\big)$. Thus we obtain the integral solution (38) for $\phi_2(z)$. In particular, the first few terms of $\phi_2(z)$ in the Taylor expansion are given as follows.

$$\phi_2(z) = \tfrac{1}{2} - \tfrac{7}{4}z + \tfrac{23}{18}z^2 - \tfrac{19951}{17280}z^3 + \tfrac{64903}{57600}z^4 - \tfrac{13803863}{12096000}z^5 + \cdots.$$

As a function in the complex plane, the region where $\phi_2(z)$ is analytic is dictated by the first zeros of $S_1(z)$, which exceeds unity.

To complete the proof of (80), we require a variation of Lemma 4, since the assumption on $a_{n,0}$ given there is not satisfied here.

Lemma 14. *Consider the recurrence*

$$\sum_{1\leqslant \ell \leqslant m} \lambda^*_{n,m,\ell}\,(a_{n,m} - a_{n,m-\ell}) = b_{n,m} \qquad (m \geqslant 1),$$

where $b_{n,m}$ is defined for $1 \leqslant m \leqslant n$ and $n \geqslant 1$. Assume that $|a_{n,0}| \leqslant cn$ for $n \geqslant 1$, and $|a_{n,1}| \leqslant 2c$ for $n \geqslant 1$. If there exists a constant $c > 0$ such that $|b_{n,m}| \leqslant \frac{c}{n}$ holds uniformly for $2 \leqslant m \leqslant n$ and $n \geqslant 1$, then

$$|a_{n,m}| \leqslant 2cH_m \qquad (1 \leqslant m \leqslant n). \tag{85}$$

Proof. The inequality (85) holds when $m = 1$ by assumption. For $m \geqslant 2$, we write the recurrence as follows

$$a_{n,m} = \frac{1}{\Lambda^*_{n,m}} \sum_{1\leqslant \ell < m} \lambda^*_{n,m,\ell}\, a_{n,m-\ell} + \frac{\lambda^*_{n,m,m} a_{n,0}}{\Lambda^*_{n,m}} + \frac{b_{n,m}}{\Lambda^*_{n,m}}.$$

By induction hypothesis and the two inequalities (see Lemma 4)

$$\Lambda^*_{n,m} \geqslant \frac{m}{n}, \quad \text{and} \quad \lambda^*_{n,m,m} = n^{-m} \leqslant n^{-2},$$

we obtain

$$|a_{n,m}| \leqslant 2cH_{m-1} + \frac{n}{m}\cdot\frac{1}{n^2}\cdot cn + \frac{n}{m}\cdot\frac{c}{n} \leqslant 2cH_m,$$

and this proves the lemma. ∎

In view of the estimates $\Delta^*_{n,0} = O\big(n^{-1}\big)$, $\Delta^*_{n,1} = O\big(n^{-2}\big)$ and $E_2(n,m) = O\big(n^{-3}\big)$, for $2 \leqslant m \leqslant n$, there exists a constant $c > 0$ such that the quantity $\Delta_{n,m} := n^2\Delta^*_{n,m}$ satisfies the assumptions of Lemma 14, which implies the bound $\Delta_{n,m} = O(H_m)$, or, equivalently $\Delta^*_{n,m} = O(n^{-2}H_m)$, uniformly for $1 \leqslant m \leqslant n$. This completes the proof of Theorem 2. ∎

C Closeness of the approximation (39) for $\mu^*_{n,m}$: graphical representations

The successive improvements attained by adding more terms on the right-hand side of (39) can be viewed in Figures 10 and 11.

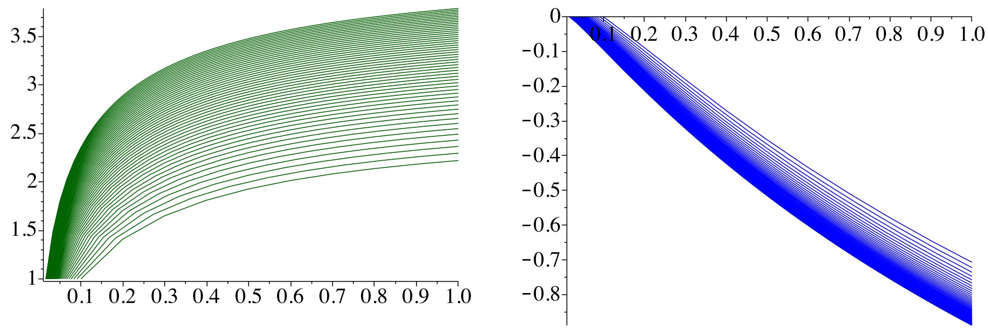


Figure 10: *Left: the sequence* $\mu^*_{n,m}$ *for* $1 \leqslant m \leqslant n$ *and* $n = 10, \dots, 60$*; Right: the difference* $\mu^*_{n,m} - H_m$ *for* n, m *in the same ranges.*

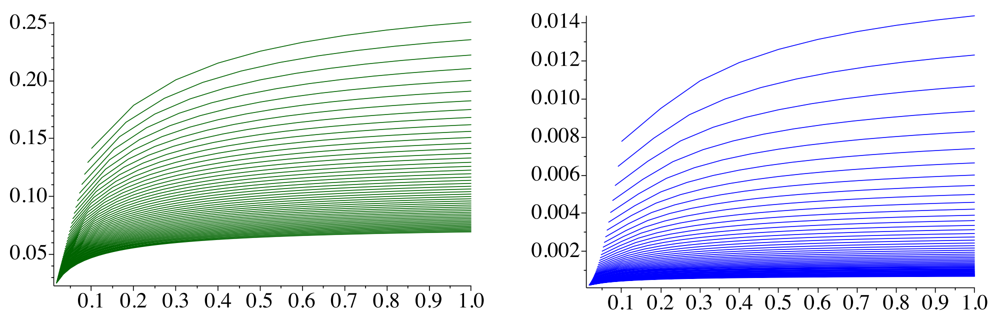


Figure 11: *The difference* $\mu^*_{n,m} - (H_m + \phi_1(\frac{m}{n}))$ *(left) and* $\mu^*_{n,m} - \big(H_m + \phi_1(\frac{m}{n}) + \frac{H_m + \phi_2(\frac{m}{n})}{n}\big)$ *(right) for* $1 \leqslant m \leqslant n$ *and* $n = 10, \dots, 60$.

D An asymptotic expansion for the mean

The above procedure can be extended to get more smaller-order terms, but the expressions for the coefficients soon become very involved. However, it follows from the discussions in § 2.4 that we expect the asymptotic expansion

$$\mu^*_{n,m} \sim \sum_{k \geqslant 0} \frac{b_k H_m + \phi_{k+1}(\alpha)}{n^k}, \tag{86}$$

in the sense that the truncated asymptotic expansion

$$\mu^*_{n,m} = \sum_{0\leqslant k\leqslant K} \frac{b_k H_m + \phi_{k+1}(\alpha)}{n^k} + O\left(n^{-K-1} H_m\right) \tag{87}$$

holds uniformly for $K \leqslant m \leqslant n$ and introduces an error of order $n^{-K-1}H_m$. This asymptotic approximation may not hold when $1 \leqslant m < K$ because additional correction terms are needed in that case. Technically, the correction terms stem from asymptotic expansions for sums of the form $\sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell}(H_m - H_{m-\ell})$; see (83) and the comments given there.

We propose here a codable procedure for the coefficients in the expansion, whose justification follows the same error analysis as above. We start with the formal expansion (86) and expand in all terms for large $m = \alpha n$ in decreasing powers of n, match the coefficients of n^{-K-1} on both sides for each $K \geqslant 0$, and then adjust the initial condition $\phi_{K+1}(0)$ by taking into account the extremal case when $m = K$ (for $m < K$ the expansion up to that order may not hold). With this algorithmic approach it is possible to determine the coefficients b_K and the functions $\phi_{K+1}(z)$ successively one after another.

Observe first that

$$H_m - H_{m-\ell} = \sum_{0\leqslant j<\ell} \frac{1}{m-j} = \sum_{0\leqslant j<\ell}\sum_{r\geqslant 1} j^{r-1} m^{-r} = \sum_{r\geqslant 1} m^{-r}\beta_r(\ell) = \sum_{r\geqslant 1} n^{-r}\alpha^{-r}\beta_r(\ell),$$

where ($0^0 = 1$)

$$\beta_r(\ell) := \sum_{0\leqslant j<\ell} j^{r-1} = \frac{1}{r}\sum_{0\leqslant j<r}\binom{r}{j} B_j \ell^{r-j},$$

the B_j representing the Bernoulli numbers. On the other hand,

$$\phi_{k+1}(\alpha) - \phi_{k+1}\left(\alpha - \frac{\ell}{n}\right) = -\sum_{r\geqslant 1}\frac{\phi^{(r)}_{k+1}(\alpha)}{r!}\left(-\frac{\ell}{n}\right)^r.$$

Thus, by substituting these expansions into (86),

$$\mu^*_{n,m} - \mu^*_{n,m-\ell} \sim \sum_{r\geqslant 1} n^{-r} \sum_{1\leqslant j\leqslant r}\left(\frac{\beta_j(\ell) b_{r-j}}{\alpha^j (r-j)!} - \frac{(-\ell)^j}{j!}\phi^{(j)}_{r-j+1}(\alpha)\right).$$

Then, by (25),

$$\sum_{1\leqslant\ell\leqslant m}\lambda^*_{n,m,\ell}\left(\mu^*_{n,m} - \mu^*_{n,m-\ell}\right) \sim \sum_{r\geqslant 1} n^{-r}[t^{-1}]\left(1+\frac{1}{nt}\right)^m\left(1+\frac{t}{n}\right)^{n+1-m} f_r(t),$$

where

$$f_r(t) := \sum_{\ell\geqslant 1} t^{\ell-1}\sum_{1\leqslant j\leqslant r}\left(\frac{\beta_j(\ell) b_{r-j}}{\alpha^j (r-j)!} - \frac{(-\ell)^j}{j!}\phi^{(j)}_{r-j+1}(\alpha)\right).$$

Now

$$\left(1+\frac{1}{nt}\right)^m\left(1+\frac{t}{n}\right)^{n+1-m} = \exp\left(\sum_{j\geqslant 1}\frac{(-1)^{j-1}}{j}\left(\frac{\alpha t^{-j} + (1-\alpha)t^j}{n^{j-1}} + \frac{t^j}{n^j}\right)\right).$$

A direct expansion using Bell polynomials $B_k^*(t_1,\dots,t_k)$ (see Comtet (1974)) then gives

$$\begin{aligned}\left(1+\frac{1}{nt}\right)^m\left(1+\frac{t}{n}\right)^{n+1-m} &= e^{\frac{\alpha}{t}+(1-\alpha)t}\sum_{k\geqslant 0}\frac{B_k^*(t_1,\dots,t_k)}{k!}n^{-k}\\ &= e^{\frac{\alpha}{t}+(1-\alpha)t}\sum_{k\geqslant 0}\frac{\tilde{B}_k(\mathbf{t})}{k!t^{2k}}n^{-k}\end{aligned}$$

where $\tilde{B}_0=1$,

$$t_j := \frac{(-1)^j j!}{j+1}\left(\frac{\alpha}{t^{j+1}}+(1-\alpha)t^{j+1}\right)+(-1)^{j-1}(j-1)!t^j \qquad (j=1,2,\dots),$$

and $\tilde{B}_k(\mathbf{t})$ is a polynomial of degree $4k$.

Collecting these expansions, we get

$$\sum_{1\leqslant \ell\leqslant m}\lambda^*_{n,m,\ell}\left(\mu^*_{n,m}-\mu^*_{n,m-\ell}\right)\sim\sum_{K\geqslant 0}n^{-(K+1)}\sum_{0\leqslant r\leqslant K}[t^{2r-1}]e^{\frac{\alpha}{t}+(1-\alpha)t}\frac{\tilde{B}_r(\mathbf{t})}{r!}f_{K+1-r}(t).$$

All terms now have the form

$$\begin{aligned}[t^{2r-1}]e^{\frac{\alpha}{t}+(1-\alpha)t}F(t) &= \sum_{\ell\geqslant 0}\frac{\alpha^\ell}{\ell!}[t^{\ell+2r-1}]e^{(1-\alpha)t}F(t)\\ &= \sum_{\ell\geqslant 0}\frac{\alpha^\ell}{\ell!}\sum_{0\leqslant j<2r+\ell}\frac{(1-\alpha)^j}{j!}\cdot\frac{F^{(\ell+2r-1-j)}(0)}{(\ell+2r-1-j)!}.\end{aligned}$$

Since we are solving the recurrence

$$\sum_{1\leqslant\ell\leqslant m}\lambda^*_{n,m,\ell}\big(\mu^*_{n,m}-\mu^*_{n,m-\ell}\big)=\frac{1}{n},$$

we have the relations

$$\begin{cases}\left(\dfrac{b_0}{\alpha}+\phi_1'(\alpha)\right)[t^{-1}]\dfrac{e^{\frac{\alpha}{t}+(1-\alpha)t}}{(1-t)^2}=1,\\ \displaystyle\sum_{0\leqslant r\leqslant K}[t^{2r-1}]e^{\frac{\alpha}{t}+(1-\alpha)t}\frac{\tilde{B}_r(\mathbf{t})}{r!}f_{K+1-r}(t)=0, \quad (K\geqslant 1).\end{cases}$$

By induction, each ϕ_{K+1} satisfies a differential equation of the form

$$\left(\frac{b_K}{\alpha}+\phi'_{K+1}(\alpha)\right)S_1(\alpha)=\mathscr{L}_{K+1}[\phi_1,\dots,\phi_K](\alpha),$$

for some functional $\mathscr{L}_{K+1}$. Since $S_1(\alpha)\sim\alpha$ as $\alpha\to 0$, we also have the relation

$$b_K=\mathscr{L}_{K+1}[\phi_1,\dots,\phi_K](0).$$

Once the value of b_K is determined, we can then write

$$\phi_{K+1}(\alpha)=\phi_{K+1}(0)+\int_0^\alpha\left(\frac{\mathscr{L}_{K+1}[\phi_1,\dots,\phi_K](x)}{S_1(x)}-\frac{b_K}{x}\right)\mathrm{d}x,$$

and it remains to determine the initial value $\phi_{K+1}(0)$, which is far from being obvious. The crucial property we need is that the truncated expansion (87) holds when $K \leqslant m \leqslant n$, and particularly when $m = K$. So we compute (87) with $m = K$ and drop all terms of order smaller than or equal to n^{-K-1}. Then we match the coefficient of n^{-K} with that in the expansion of $\mu^*_{n,K}$ obtained by a direct calculation from the recurrence (16).

We illustrate this procedure by computing the first two terms in (86). First, we have

$$\left(\frac{b_0}{\alpha} + \phi_1'(\alpha)\right) S_1(\alpha) = 1,$$

which implies $b_0 = 1$ and $\phi_1'(\alpha) = \frac{1}{S_1(\alpha)} - \frac{1}{\alpha}$. Moreover, substituting the initial value $m = K = 0$, we get

$$0 = \mu^*_{n,0} = H_0 + \phi_1(0) + O\big(n^{-1}\big) = \phi_1(0) + O\big(n^{-1}\big),$$

entailing $\phi_1(0) = 0$, which is consistent with what we obtained above.

The next-order term when $K = 1$ is (after substituting the relations $b_0 = 1$, $\phi_1'(\alpha) = \frac{1}{S_1(\alpha)} - \frac{1}{\alpha}$ and $\phi_1''(\alpha) = \frac{1}{\alpha^2} - \frac{S_1'(\alpha)}{S_1(\alpha)^2}$)

$$\begin{aligned}&\left(\frac{b_1}{\alpha} + \phi_2'(\alpha)\right)[t^{-1}]\frac{e^{\frac{\alpha}{t}+(1-\alpha)t}}{(1-t)^2}\\ &\quad = [t^{-1}]e^{\frac{\alpha}{t}+(1-\alpha)t}\left(\frac{1}{2\alpha^2(1-t)^2} - \frac{(1+t)S_1'(\alpha)}{2(1-t)^3 S_1(\alpha)} + \frac{\alpha - 2t^3 + (1-\alpha)t^4}{2t^2(1-t)^2 S_1(\alpha)}\right),\end{aligned}$$

implying that

$$\left(\frac{b_1}{\alpha} + \phi_2'(\alpha)\right) S_1(\alpha) = -\frac{S_1'(\alpha)S_2(\alpha)}{2S_1(\alpha)^2} + \frac{S_0(\alpha)}{S_1(\alpha)} + \frac{1}{2} + \frac{S_1(\alpha)}{2\alpha^2}. \tag{88}$$

As $\alpha \to 0$, the right-hand side of (88) has the local expansion $1 - \frac{1}{4}\alpha + \cdots$, forcing $b_1 = 1$, and, accordingly, we obtain the same differential equation (84). Substituting the value $m = K = 1$ in (87) yields

$$\begin{aligned}1 = \mu^*_{n,1} &= H_1 + \phi_1\Big(\frac{1}{n}\Big) + \frac{1}{n}\left(H_1 + \phi_2\Big(\frac{1}{n}\Big)\right) + O\big(n^{-2}\big)\\ &= 1 + \frac{\phi_1'(0)}{n} + \frac{1}{n} + \frac{\phi_2(0)}{n} + O\big(n^{-2}\big),\end{aligned}$$

implying, by using (43), $\phi_2(0) = -\phi_1'(0) - 1 = \frac{1}{2}$, which is consistent with Theorem 2.

Although the expressions become rather involved for higher-order terms, all calculations (symbolic or numerical) are easily codable. For example, we have

$$\phi_3(z) = \tfrac{1}{12} - \tfrac{575}{432}z + \tfrac{15101}{11520}z^2 - \tfrac{8827}{5400}z^3 + \tfrac{2229089}{1036800}z^4 - \tfrac{361022171}{127008000}z^5 + \cdots.$$

E Asymptotic expansions for $V^*_{n,m}$ for small m and the refined approximation (52) to $V^*_{n,m}$

Recall that

$$V^*_{n,m} = \frac{e_n^2}{n^2}\big(\mathbb{V}(X_{n+1,m}) + \mathbb{E}(X_{n+1,m})\big).$$

This sequence satisfies $V^*_{n,0} = 0$, and for $1 \leqslant m \leqslant n$,

$$\sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell} \left(V^*_{n,m} - V^*_{n,m-\ell}\right) = T^*_{n,m}, \tag{89}$$

where

$$T^*_{n,m} := \sum_{1\leqslant \ell\leqslant m} \lambda^*_{n,m,\ell} \left(\mu^*_{n,m} - \mu^*_{n,m-\ell}\right)^2.$$

From this recurrence, we obtain the expansions

$$\begin{aligned}
V^*_{n,1} &= 1,\\
V^*_{n,2} &= \tfrac{5}{4} - \tfrac{1}{2}n^{-1} + \tfrac{3}{4}n^{-2} - \tfrac{5}{4}n^{-3} + \tfrac{31}{16}n^{-4} - 3n^{-5} + O\big(n^{-6}\big),\\
V^*_{n,3} &= \tfrac{49}{36} - \tfrac{17}{18}n^{-1} + \tfrac{52}{27}n^{-2} - \tfrac{139}{36}n^{-3} + \tfrac{3157}{432}n^{-4} - \tfrac{361}{27}n^{-5} + O\big(n^{-6}\big),\\
V^*_{n,4} &= \tfrac{205}{144} - \tfrac{95}{72}n^{-1} + \tfrac{1489}{432}n^{-2} - \tfrac{1243}{144}n^{-3} + \tfrac{33091}{1728}n^{-4} - \tfrac{28979}{864}n^{-5} + O\big(n^{-6}\big).
\end{aligned}$$

Observe that the leading constant terms are exactly given by

$$\left\{H^{(2)}_m\right\} = \left\{1, \tfrac{5}{4}, \tfrac{49}{36}, \tfrac{205}{144}, \tfrac{5269}{3600}, \tfrac{5369}{3600}, \dots\right\}.$$

These expansions suggest the general form

$$V^*_{n,m} \approx H^{(2)}_m + \sum_{k\geqslant 1} \frac{\tilde{d}_k(m)}{n^k}.$$

With this form using the technique of matched asymptotics, we are then led to the following explicit expressions.

$$\begin{aligned}
\tilde{d}_1(m) &= -2H_m + 2H^{(2)}_m, \quad \text{for } m \geqslant 0,\\
\tilde{d}_2(m) &= -\tfrac{11}{2}H_m + \tfrac{7}{3}H^{(2)}_m + \tfrac{7}{12} + \tfrac{11}{4}m, \quad \text{for } m \geqslant 2,\\
\tilde{d}_3(m) &= -\tfrac{73}{9}H_m + \tfrac{7}{3}H^{(2)}_m + \tfrac{1}{6} + \tfrac{239}{36}m - \tfrac{49}{36}m^2, \quad \text{for } m \geqslant 2,\\
\tilde{d}_4(m) &= -\tfrac{1349}{144}H_m + 2H^{(2)}_m + \tfrac{197}{144} + \tfrac{14135}{1728}m - \tfrac{6283}{2880}m^2 + \tfrac{2473}{4320}m^3, \quad \text{for } m \geqslant 4.
\end{aligned}$$

The above expansions for small m suggest the more uniform asymptotic expansion for $V^*_{n,m}$ for $1 \leqslant m \leqslant n$

$$V^*_{n,m} \sim H^{(2)}_m + \sum_{k\geqslant 1} \frac{a_k H_m + \psi_k(\alpha) + c_k H^{(2)}_m}{n^k}, \tag{90}$$

in the sense that omitting all terms with indices $k > K$ introduces an error of order $n^{-(K+1)}H_m$; furthermore, the expansion holds uniformly for $K \leqslant m \leqslant n$. We elaborate this approach by carrying out the required calculations up to $k = 2$, which then characterizes particularly the constant a_3 and the function $\psi_2(z)$.

We start with the formal expansion (90) and expand in recurrence (89) all terms for large $m = \alpha n$ in decreasing powers of n; we then match the coefficients of $n^{-(K+1)}$ on both sides for each $K \geqslant 1$. To specify the initial condition $\psi_K(0)$ we incorporate the information from the asymptotic expansion for $V^*_{n,K}$ (obtained by exact solution). With this algorithmic approach it is possible to determine the coefficients a_k and c_k and the functions $\psi_k(z)$ successively one after another. We remark that a formalization of this procedure at the generating function level as carried out for the expectation in Appendix D could be given also, but here we do not pursue this any further.

We use the expansions

$$\begin{aligned}
\phi\Big(\frac{m}{n}\Big) - \phi\Big(\frac{m-\ell}{n}\Big) &= \phi'(\alpha)\frac{\ell}{n} - \phi''(\alpha)\frac{\ell^2}{2n^2} + \phi'''(\alpha)\frac{\ell^3}{6n^3} + \cdots,\\
H_m - H_{m-\ell} &= \frac{\ell}{\alpha\, n} + \frac{\ell(\ell-1)}{2\alpha^2 n^2} + \frac{\ell(\ell-1)(2\ell-1)}{6\alpha^3 n^3} + \cdots,\\
H_m^{(2)} - H_{m-\ell}^{(2)} &= \frac{\ell}{\alpha^2 n^2} + \frac{\ell(\ell-1)}{\alpha^3 n^3} + \frac{\ell(\ell-1)(2\ell-1)}{2\alpha^4 n^4} + \cdots
\end{aligned}$$

as well as those for $\mu^*_{n,m}$ and $\Lambda^{*(r)}_{n,m}$ in (86) and (29), respectively. The expansion of the right-hand side of (89) then starts as follows.

$$T^*_{n,m} = \frac{T_1(\alpha)}{n^2} + \frac{T_2(\alpha)}{n^3} + \cdots,$$

where

$$\begin{aligned}
T_1(z) &= \frac{S_2(z)}{S_1^2(z)},\\
T_2(z) &= -\frac{S_2^2(z)S_1'(z)}{S_1^4(z)} + \frac{S_3(z)S_1'(z)}{S_1^3(z)} + \frac{2S_0(z)S_2(z)}{S_1^3(z)} + \frac{S_0(z)}{S_1^2(z)} - \frac{S_2(z)}{2S_1^2(z)} - \frac{2}{S_1(z)}.
\end{aligned}$$

For the left-hand side of (89), the asymptotic form (90) leads to

$$\sum_{1\leqslant\ell\leqslant m} \lambda^*_{n,m,\ell}\left(V^*_{n,m} - V^*_{n,m-\ell}\right) = \frac{V_1(\alpha)}{n^2} + \frac{V_2(\alpha)}{n^3} + \cdots,$$

where

$$\begin{aligned}
V_1(z) &= \left(\frac{1}{z^2} + \frac{a_2}{z} + \psi_1'(z)\right) S_1(z),\\
V_2(z) &= \left(-\frac{1}{z^2} - \frac{a_2}{z} - \psi_1'(z)\right) S_0(z)\\
&\quad + \left(-\frac{1}{z^3} - \frac{1}{2z^2} + \frac{c_6}{z^2} - \frac{a_2}{2z^2} - \frac{a_2}{2z} + \frac{a_3}{z} - \frac{\psi_1'(z)}{2} + \psi_2'(z)\right) S_1(z)\\
&\quad + \left(\frac{1}{z^3} + \frac{a_2}{2z^2} - \frac{\psi_1''(z)}{2}\right) S_2(z).
\end{aligned}$$

Observe that all functions $V_k(z)$, $T_k(z)$ have a simple pole at $z = 0$.

We match the terms in the expansion and consider $V_1(z) = T_1(z)$. First we compare the first two terms of the Laurent expansions of both functions. Using (36), we get

$$\begin{aligned}
V_1(z) &= \frac{1}{z} + \left(\frac{3}{2} + a_2\right) + O(z),\\
T_1(z) &= \frac{1}{z} - \frac{1}{2} + O(z),
\end{aligned}$$

and by matching the two constant terms, we see that $a_2 = -2$. The equation $V_1(z) = T_1(z)$ characterizes then the function $\psi_1'(z)$ of the form

$$\psi_1'(z) = \frac{S_2(z)}{S_1^3(z)} - \frac{1}{z^2} + \frac{2}{z}.$$

Next we consider $V_2(z) = T_2(z)$ and obtain

$$\begin{aligned} V_2(z) &= \left(-\frac{1}{2} + c_6\right)\frac{1}{z} + \left(-\frac{5}{12} - a_2 + \frac{3c_6}{2} + a_3\right) + O(z), \\ T_2(z) &= \frac{3}{2z} - \frac{11}{12} + O(z), \end{aligned}$$

and thus, by matching the terms and using the values already computed in the first-order approximation for $V_{n,m}^*$, $c_6 = 2$ and $a_3 = -\frac{11}{2}$. Then the function $\psi_2'(z)$ can be characterized by equating $V_2(z) = T_2(z)$, which then gives

$$\begin{aligned} \psi_2'(z) = &-\frac{5S_2^2(z)S_1'(z)}{2S_1^5(z)} + \frac{S_3(z)S_1'(z)}{S_1^4(z)} + \frac{3S_2(z)S_0(z)}{S_1^4(z)} + \frac{S_2(z)S_2'(z)}{2S_1^4(z)} \\ &+ \frac{S_0(z)}{S_1^3(z)} - \frac{2}{S_1^2(z)} + \frac{1}{z^3} - \frac{3}{z^2} + \frac{11}{2z}. \end{aligned}$$

All constants and functions here match with those obtained earlier in previous paragraphs, and we can pursue the same calculations further and obtain finer approximations. For example, we have $c_7 = \frac{7}{3}$. But the calculations are long and laborious.

Finally, it remains to determine the constant terms in the Taylor expansion of the functions $\psi_k(z)$ by adjusting them to the expansion of $V_{n,m}^*$ for small m. This yields $\psi_1(0) = 0$, and

$$\psi_2(0) = \frac{7}{12}.$$

This characterizes the function $\psi_2(z)$ in Theorem 4 as follows.

$$\begin{aligned} \psi_2(z) = \frac{7}{12} + \int_0^z \Bigg(&-\frac{5S_2^2(t)S_1'(t)}{2S_1^5(t)} + \frac{S_3(t)S_1'(t)}{S_1^4(t)} + \frac{3S_2(t)S_0(t)}{S_1^4(t)} + \frac{S_2(t)S_2'(t)}{2S_1^4(t)} \\ &+ \frac{S_0(t)}{S_1^3(t)} - \frac{2}{S_1^2(t)} + \frac{1}{t^3} - \frac{3}{t^2} + \frac{11}{2t} \Bigg) \mathrm{d}t. \end{aligned}$$

In particular, the first few terms in the Taylor expansion of $\psi_2(z)$ are given by

$$\psi_2(z) = \tfrac{7}{12} + \tfrac{239}{36}z - \tfrac{6283}{2880}z^2 - \tfrac{4529}{3600}z^3 + \tfrac{9283591}{1814400}z^4 - \tfrac{137478949}{14112000}z^5 + \cdots .$$

F Closeness of the approximation (52) for $V_{n,m}^*$: graphical representations

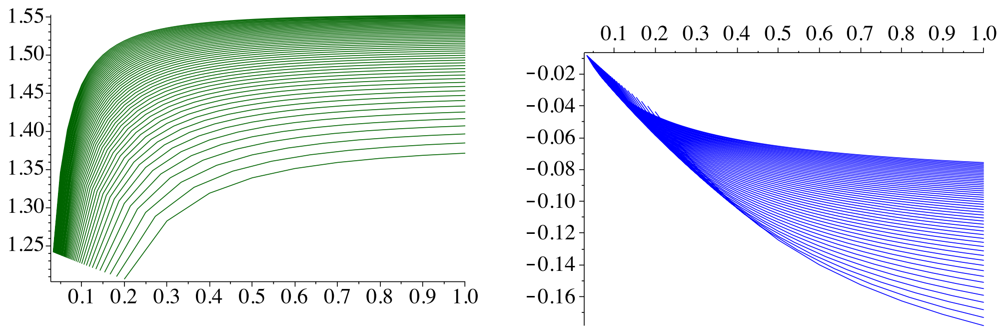


Figure 12: *Left: the sequence* $V^*_{n,m}$ *for* $2 \leqslant m \leqslant n$ *and* $n = 10, \ldots, 60$*; Right: the difference* $V^*_{n,m} - H_m^{(2)}$ *for* n, m *in the same ranges.*

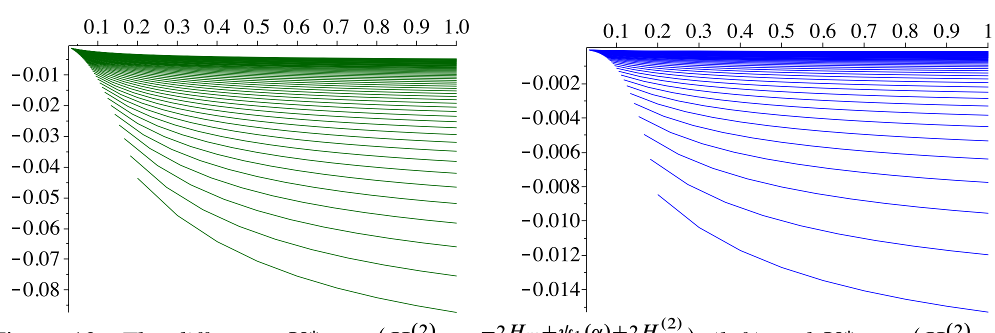


Figure 13: *The difference* $V^*_{n,m} - \left(H_m^{(2)} + \frac{-2H_m + \psi_1(\alpha) + 2H_m^{(2)}}{n}\right)$ *(left) and* $V^*_{n,m} - \left(H_m^{(2)} + \frac{-2H_m + \psi_1(\alpha) + 2H_m^{(2)}}{n} + \frac{-\frac{11}{2}H_m + \psi_2(\alpha) + \frac{7}{3}H_m^{(2)}}{n^2}\right)$ *(right) for* $2 \leqslant m \leqslant n$ *and* $n = 10, \ldots, 60$.