Large deviations of the length of the longest increasing subsequence of random permutations and random walks

Jörn Börjes,[*] Hendrik Schawe,[†] and Alexander K. Hartmann[‡]
Institut für Physik, Universität Oldenburg, 26111 Oldenburg, Germany

(Received 17 January 2019; published 2 April 2019)

We study numerically the length distribution of the longest increasing subsequence (LIS) for random permutations and one-dimensional random walks. Using sophisticated large-deviation algorithms, we are able to obtain very large parts of the distribution, especially also covering probabilities smaller than 10^{-1000}. This enables us to verify for the length of the LIS of random permutations the analytically known asymptotics of the rate function and even the whole Tracy-Widom distribution. We observe a rather fast convergence in the larger than typical part to this limiting distribution. For the length L of LIS of random walks no analytical results are known to us. We test a proposed scaling law and observe convergence of the tails into a collapse for increasing system size. Further, we obtain estimates for the leading-order behavior of the rate functions in both tails.

DOI: 10.1103/PhysRevE.99.042104

I. INTRODUCTION

We study the length distribution of the *longest increasing subsequence* (LIS) [1] of different ensembles of random sequences. A subsequence of a sequence S consists of elements of S in the same order as in S. But neighbors in the subsequence are not necessarily neighbors in S. For a LIS it is required that the elements of the subsequence are increasing from left to right, and the number of elements in the subsequence is maximal.

The first mention of this problem involving random permutations (RPs) is from Stanisław Ulam [2] and is also known as "Ulam's problem." In his study the mean length L of LIS on RP of n integers were examined by means of Monte Carlo simulations. It was conjectured that, in the limit of large n, the length converges to $L = c\sqrt{n}$, with a constant c, which was later proven to be $c = 2$ [3]. In the following years much work was published scrutinizing the large-deviation behavior of this problem, and explicit expressions for both the left (lower) and right (upper) tail were derived rigorously [4–6]. Interestingly, for the LIS of RPs it was shown that the length distribution $P(L)$ is a Tracy-Widom distribution [7].

The Tracy-Widom distribution was at that time only known from random matrix theory, where it describes the fluctuations of the largest eigenvalues of the *Gaussian unitary ensemble* (GUE), an ensemble of Hermitian random matrices. In physics it came into focus after an explicit mapping of an $1+1$-dimensional polynuclear growth model [8]. Subsequently other mappings of $1+1$-dimensional growth models belonging to the Kardar-Parisi-Zhang universality like an anisotropic ballistic deposition were found [9]. Other models, in which the Tracy-Widom distribution appears, include the totally asymmetric exclusion process [10] and directed polymers [11]. For a pedagogical overview of the relations of different models exhibiting a Tracy-Widom distribution there are some review articles, e.g., Refs. [12–14]. Fluctuations in growth processes following the Tracy-Widom distribution could also be observed in experiments, e.g., from growing liquid crystals where the Tracy-Widom distribution of the GUE appears for circular growth [15] and of the *Gaussian orthogonal ensemble* (GOE) for growth from a flat surface [16].

The Tracy-Widom distribution seems to occur always together with a *third-order phase transition* between a *strongly interacting* phase in the left tail and a *weakly interacting* phase in the right tail [17]. For these third-order phase transitions, the probability density function behaves in the left tail as $P(x) \approx e^{-n\Phi_-}$ with the role of the free energy played by the *rate function* $\Phi_-(x) \sim (a-x)^3$ for $x \to a$ from the left, where the scaled mean value a is the critical point of the transition. Here n is some large parameter, e.g, the system size. The $O(x^3)$ leading-order behavior of Φ_- generally leads to a discontinuity in the third derivative of the free energy and therefore to a third-order phase transition. This seems to be a characteristic sign predicting the main region of the distribution to follow a Tracy-Widom distribution. Therefore the behavior of the far tails of these problems is of great interest to understand this connection better. Consequently the *large deviations* of some of these models were studied thoroughly [17,18].

For the length distribution in the RP case the large deviations, the behavior for large values of n including the far tails, are known analytically [4–7]. These show the characteristic behavior of the above mentioned left-tail rate function. For the case of random walks (RWs), bounds for the behavior of the mean are known [19], and there is also numerical work which is concerned with the distribution in the typical region [20], i.e., those LISs which occur with a high enough probability of about $\geqslant 10^{-6}$. We also deem it worthwhile to look closer at the tails of the distribution for finite systems.

For the purpose of studying the large deviations of this problem numerically, we utilize sophisticated large-deviation

[*]joern.boerjes@uni-oldenburg.de
[†]hendrik.schawe@uni-oldenburg.de
[‡]a.hartmann@uni-oldenburg.de

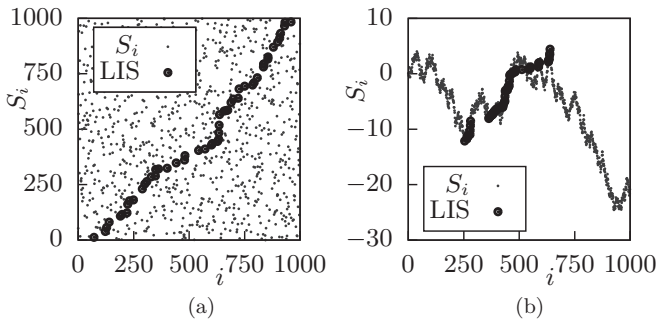

FIG. 1. Visualization of random sequences of length $n = 1000$ where the value is plotted over the corresponding index. Marked with circles are the entries of one possible LIS. (a) Random permutation (RP), (b) random walk (RW).

sampling methods to observe the distribution $P(L)$. In this way we can observe directly the far tails of the Tracy-Widom distribution for the RP case [7] and can confirm the known large n asymptotics [6]. The second ensemble are one-dimensional RWs with increments from a uniform distribution. While we can observe the scaling proposed in Ref. [20] for the main region, the tails are subject to considerable finite-size effects. Nevertheless the distributions collapse over larger regions for larger sizes n. Also, we give estimates for the leading-order behavior of the rate functions governing the left and right tails of the distribution $P(L)$.

This study first introduces the different ensembles of interest and the algorithms used to obtain the distribution of the length in Sec. II. Section III shows the results we gathered and interprets them. We conclude this study in Sec. IV.

II. MODELS AND METHODS

To define the LIS, we have to define a subsequence first. Given some sequence $S = (S_1, S_2, \ldots, S_n)$ a *subsequence* of length L is a sequence $s = (S_{i_1}, S_{i_2}, \ldots, S_{i_L})$ $(1 \leqslant i_j \leqslant n, i_j < i_{j+1}$ for all $j = 1, \ldots, L)$ containing only elements present in S in the same order as in S, though possibly with gaps. An *increasing subsequence* has elements such that every element in s is smaller than its predecessor, i.e., $S_{i_j} < S_{i_{j+1}}$ for $j = 1, \ldots, L-1$. The LIS is consequently the longest, i.e., the one with the highest number L of elements, of all possible increasing subsequences. Note that the LIS is not necessarily unique, but by definition its length is unique. As an example two different LISs are marked by overlines and underlines in the following sequence: $S = (\underline{3}, 9, \underline{4}, \overline{1}, \overline{2}, \underline{7}, \overline{6}, \overline{\underline{8}}, 0, 5)$.

In this study the sequence S is drawn either from the ensemble of *random permutations* of n consecutive integers or from the ensemble of *random walks* with increments δ_j $(j = 1, \ldots, n)$ from a uniform distribution $\delta_j \sim U(-1, 1)$, such that

$$S_i = \sum_{j=1}^{i} \delta_j. \tag{1}$$

An example of each sequence with the corresponding LIS marked is shown in Fig. 1.

To find L of any given sequence, we use the *patience sort algorithm*. We introduce only the very simple version to obtain the length, but a comprehensive review of the connection of patience sort with the LIS can be found in Ref. [3]. In short, the patience sort algorithm works as follows: We iterate over the n entries S_i and place each into an initially empty stack (or pile) a_j on the smallest j such that for the top entry $\mathrm{top}(a_j) > S_i$ holds. Note that this always ensures that the top entries of a are ascendingly sorted, such that we can determine j by a binary search in $O(\ln n)$. Finally, the number of nonempty stacks a_j is equal to the length L of the LIS.

A. Large-deviation sampling

To be able to gather statistics of the large-deviation regime numerically [21], we need to apply a sophisticated sampling scheme. Therefore we use a well-tested [22–24] Markov chain Monte Carlo sampling which treats the system as a canonical system at an artificial *temperature* with the observable of interest as its *energy*. Since the algorithm has been presented comprehensively in the literature, we here mainly state the details specific to the current application. In our case, we identify the state of the system with the sequence, the length L with the energy and sample the equilibrium state at temperature Θ using the Metropolis algorithm [25,26]. Controlling the temperature allows us to direct the sampling to different regimes of the distributions, to eventually cover the distribution over a large part of the support. To evolve our Markov chain of sequences, we have to introduce change moves, which modify a sequence and consequently the energy L. For the RP we swap two random entries, and for the RW we replace one of the increments δ_j [cf. Eq. (1)] by a new random number drawn from the same uniform distribution. These changes are accepted according to the Metropolis acceptance ratio

$$P_{\mathrm{acc}} = \min(1, e^{-\Delta L/\Theta}), \tag{2}$$

where ΔL is the change in energy due to the change move. This Markov chain of sequence realizations converges to an equilibrium state. As usual with Markov chain Monte Carlo simulations, we need to ensure equilibration and that the samples are decorrelated [26].

In equilibrium the realizations generally have a lower than typical energy for low temperatures and typical energies for high temperatures. We also introduce negative temperatures for larger than typical energies. This way the temperature can be tuned to guide the simulation towards realizations within a specific range of energies L. We know the equilibrium distribution $Q_\Theta(S)$ at temperature Θ of realizations, i.e., sequences S, to be

$$Q_\Theta(S) = \frac{1}{Z_\Theta} e^{-L(S)/\Theta} Q(S), \tag{3}$$

with the natural distribution $Q(S)$, i.e., the distribution of realizations arising by simply generating subsequences uniformly. This can be exploited to correct for the bias introduced by the temperature and arrive at the unbiased distribution $P(L)$ with good statistics also in the regions unreachable by simple sampling. Therefore consider the sampled equilibrium distributions $P_\Theta(L)$. To connect them to the distribution of realizations $Q_\Theta(S)$, we can sum all realizations with the same

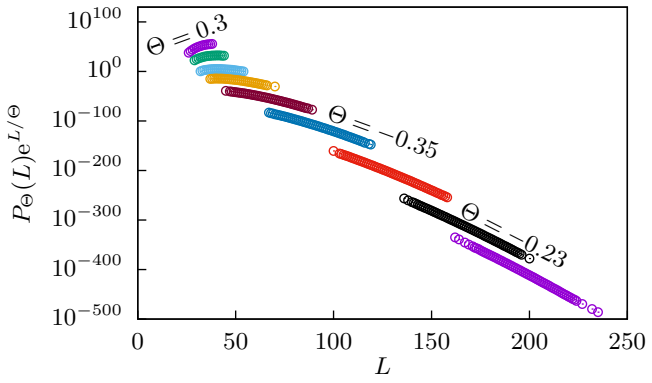

FIG. 2. Intermediate step after correction with Eq. (6) but before determination of the values Z_{Θ_i} (i.e., all $Z_{\Theta_i} = 1$). The data are gathered for RW sequences of length $n = 512$. Each shade of gray (color) is sampled at a different temperature Θ, and for three data sets the corresponding temperatures are annotated. (For clarity some evaluated temperatures are omitted.)

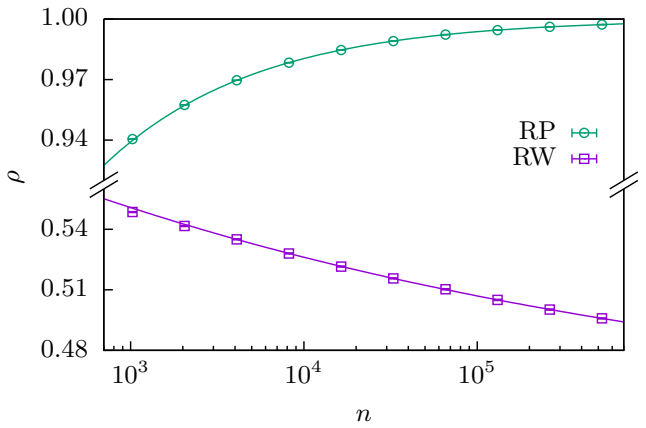

FIG. 3. Extrapolation of the span ρ. Measurements at different sizes n are used to extrapolate an asymptotic span according to a power law with offset $\rho = an^b + \rho_\infty$. Fits to this expression for $n \geqslant 4096$ are marked by a line. The two obtained asymptotic values are $\rho_\infty^{\mathrm{rp}} = 1.00005(2)$ for the RP and $\rho_\infty^{\mathrm{rp}} = 0.439(7)$ for the RW. Note the broken ρ axis. Error bars are smaller than the width of the line.

value of L, leading to

$$P_\Theta(L) = \sum_{\{S|L(S)=L\}} Q_\Theta(S) \tag{4}$$

$$= \sum_{\{S|L(S)=L\}} \frac{1}{Z_\Theta} e^{-L(S)/\Theta} Q(S) \tag{5}$$

$$= \frac{1}{Z_\Theta} e^{-L(S)/\Theta} P(L). \tag{6}$$

Solving this equation for $P(L)$ allows us to correct for the bias introduced by the temperature. An intermediate snapshot of this process is shown in Fig. 2.

The constants Z_Θ can be obtained by enforcing continuity of the distribution,

$$P_{\Theta_j}(L) e^{L/\Theta_j} Z_{\Theta_j} = P_{\Theta_i}(L) e^{L/\Theta_i} Z_{\Theta_i}, \tag{7}$$

for pairs of i, j for which the gathered data $P_{\Theta_i}(L)$ overlap with $P_{\Theta_j}(L)$. While this can be used to approximate the ratios of pairwise Z_{Θ_i}, the absolute value can then be obtained by normalization of the whole distribution. This procedure requires a clever choice of temperatures, since gaps in the sampled range of L would make it impossible to find a ratio of Z_{Θ_i} on the left and right sides of the gap. We use on the order of 100 distinct temperatures. In general, the larger the size n, the more temperatures are needed.

III. RESULTS

Before we look into the large-deviation tails, we in brief present some simple sampling results addressing the qualitative difference of RP and RW cases, which are visible in Fig. 1. The entries of the RW are strongly correlated such that the RW typically consists of runs with downward or upward trends. This means that the LIS is typically confined in an upward trend, and its entries therefore are close together. The RP, on the other hand, typically shows LISs with entries over the whole range.

To quantify this effect we measure the fraction of the sequence over which the LIS spans. For multiple system sizes $1024 \leqslant n \leqslant 524\,288$, 10^6 samples each, we measure the positions i, j of the first and last entries of a found LIS to calculate its relative *span* $\rho = (j - i)/n$. We extrapolate the mean span with an offset power law $\rho = an^b + \rho_\infty$ to extrapolate the asymptotic span ρ_∞, which is shown in Fig. 3.

For the RP we get a value of $\rho_\infty^{\mathrm{rp}} = 1.00005(2)$ and for the RW $\rho_\infty^{\mathrm{rp}} = 0.439(7)$. Note that these numbers are subject to two sources of systematical errors, which can explain, e.g., the impossible result of $\rho_\infty^{\mathrm{rp}} > 1$. First, the function we use to extrapolate is an ansatz, which considers only leading-order behavior of the actual scaling function. Second, we obtain only one LIS per sequence via the backpointer extension of patience sort [3], which might result in a biased selection of LISs. Both questions merit further research on their own but are beyond the scope of this article. This means that LISs of RPs typically span the whole sequence, while LISs of RWs typically span only less than half of the sequence, such that its entries are closer together.

To gather statistics of L, we apply the temperature-based sampling scheme for the two cases of RPs and of RWs with uniform increments. In both cases, we study five different system sizes n up to $n = 4096$ each.

A. Random permutations

First, we look at the LIS length distribution of RPs. For this case there are already many properties known in the limit of $n \to \infty$.

It is known that the distribution should converge to a suitably rescaled Tracy-Widom distribution χ of the GUE ensemble [7] for large values of n as

$$P_n[(L - 2\sqrt{n})n^{-1/6}] = \chi[(L - 2\sqrt{n})n^{-1/6}]. \tag{8}$$

Rescaled to compensate for this leading behavior, our results are shown in Fig. 4. By using the large-deviation approach, we are able to measure probabilities as small as 10^{-1000} and less, allowing us to go beyond the first numerical work [20] on the distribution of LISs. We can observe a very good collapse up

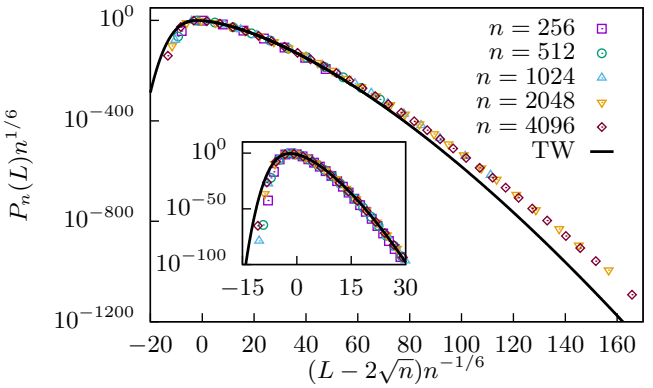

FIG. 4. Numerically obtained distributions for different system sizes n rescaled according to Eq. (8). The Tracy-Widom distribution is drawn as a black line [27] and is expected to be the curve all distributions collapse onto. The inset shows a zoom on the intermediate tails. On the left the tendency of our data towards the Tracy-Widom distribution with increasing system size n is visible. (For clarity some data points are discarded to show the same density of symbols for every system size.)

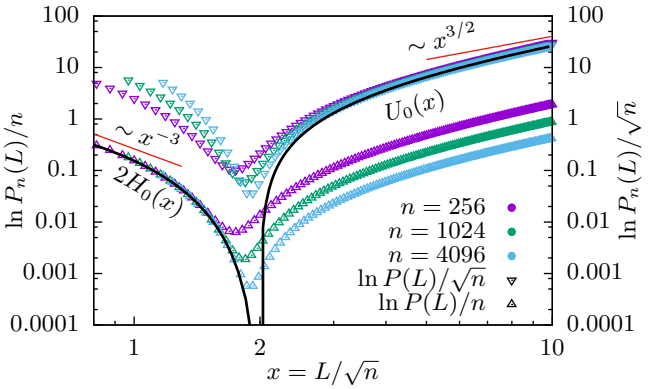

FIG. 5. Empirical rate functions for different system sizes n. On the top (triangles down) scaled as $\ln P_n(L)/\sqrt{n}$ to emphasize the right-tail behavior. On the bottom (triangles up) scaled as $\ln P_n(L)/n$ to emphasize the left tail behavior. The analytically known rate functions for both tails $2H_0$ and U_0 are shown in the correspondingly scaled region and a convergence of the data to these functions is well visible. The leading-order terms of the series expansion (cf. Ref. [4]) are also shown as straight lines next to the rate function.

to probabilities of 10^{-200} of our data onto the Tracy-Widom distribution given in the tables of Ref. [27].

Also note that the collapse works very well in the intermediate right tail but converges a bit slower in the left tail and far slower in the far-right tail. The inset zooms into the intermediate tail of the probability density function $P > 10^{-100}$, where the collapse fits very well to the expected Tracy-Widom distribution. In the far tails we observe considerable deviations, from the tabulated data, which are at least in part caused by finite-size effects due to the relatively small sizes n of our sequences. For a more extensive study of these finite-size effects, one could obtain the empirical distribution for more sizes, and extrapolate the finite-size effects to $n \to \infty$, as done in Ref. [28]. We do not attempt this analysis here, since the very small deviations between different values of n in the right tail suggest that data for much larger sizes would be needed for a meaningful extrapolation. This is at the moment not computationally feasible for us. Nevertheless, our numerically obtained tails fit very well to another expected form, which will be explained later, such that we assume a stronger influence of corrections to scaling in the far tails instead of systematic errors in our data.

Also note that while we can sample a very large part of the distribution $P(L)$ in the RP case—even including events with a probability less than 10^{-1000} for the largest systems—we cannot reach across the whole range of possible values. Possible approaches to extend this range are improvements to our sampling algorithm by, e.g., switching to a better change move or trying a different sampling algorithm like Wang-Landau's method [29].

The left-tail asymptotic, i.e., $L/\sqrt{n} = x < 2$, of the probability density function is given by the analytically known rate function [5,6]

$$\lim_{n\to\infty} \frac{1}{n} \ln P_n(L) = -2H_0(x) \tag{9}$$

with

$$H_0(x) = -\frac{1}{2} + \frac{x^2}{8} + \ln\frac{x}{2} - \left(1 + \frac{x^2}{4}\right) \ln\left(\frac{2x^2}{4 + x^2}\right); \tag{10}$$

the right-tail asymptotic, i.e., $L/\sqrt{n} = x > 2$, is given by [4,6]

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \ln P_n(L) = -U_0(x) \tag{11}$$

with

$$U_0(x) = 2x \cosh^{-1}(x/2) - 2\sqrt{x^2 - 4}. \tag{12}$$

Note that Eq. (11) behaves atypically for a rate function as the distribution behaves like $P_n \propto e^{-\sqrt{n}U_0}$, which according to the definition (e.g., Ref. [30]) does therefore not fulfill the large-deviation principle. Nevertheless, it describes the right-tail behavior of the distribution in leading order.

We use our sampled data to test these rate functions. If the data are suitably rescaled according to Eqs. (9) and (11), in the corresponding tails we can observe a very nice convergence of the data to the rate functions. This is plotted in Fig. 5. This excellent agreement of analytical and numerical results over hundreds of decades in probability gives us confidence that our approach works well and can be extended to cases where no analytical results are known. Also note that we can observe in our data the leading-order behavior of the left-tail rate function H_0, which goes with the exponent 3 characteristic for the third-order phase transition, confirming its connection with the Tracy-Widom distribution [17].

B. Random walks

The second class of sequences S we scrutinize are RWs. The distribution beyond the high-probability peak region seems to be unknown. Again, by applying the large-deviation approach, we sample basically the whole distribution and can even compare the right tail of our distribution with the

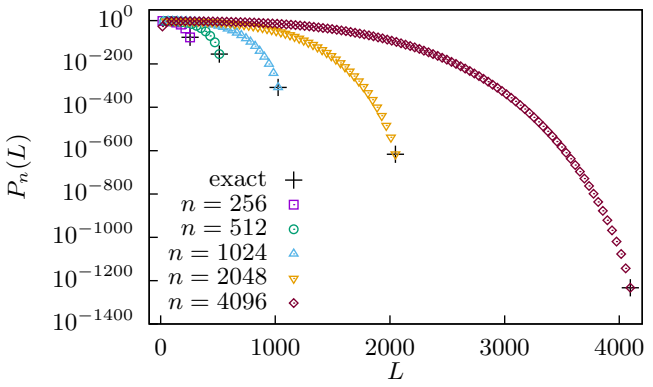

FIG. 6. Probability distributions $P_n(L)$ of the length of the LIS of RWs with exact extreme values for the $n = L$ case. (For clarity only every 40th bin is visualized, including the $n = L$ bin.)

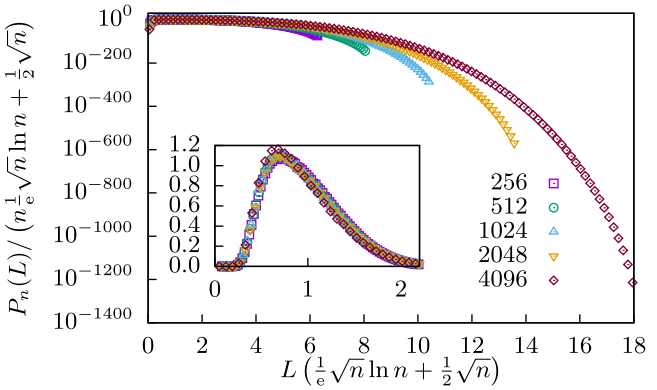

FIG. 7. Collapse of different system sizes on a common curve g from Eq. (13), with $\langle L\rangle$ given by Eq. (14). Apparently the far tail shows corrections to the proposed scaling for finite sizes, which are explained by finite-size effects, e.g., that there is a maximum length of n for finite systems. For increasing sizes n a convergence to a common curve is visible. The inset shows the same in linear scale around the maximum. (For clarity not all data points are drawn.)

corner case of $L = n$, which occurs only if all increments δ are positive and therefore with probability 2^{-n}. This case is marked in Fig. 6 to emphasize the quality of our data. For the left tail, we can not sample so far, as the very steep decline of the distribution is difficult to handle for our sampling scheme.

For RWs with increments from a symmetric uniform distribution, indeed for increments from any symmetric distribution with finite variance, the scaling of the mean as $\langle L\rangle \propto n^{\theta}$ and the variance as $\sigma^2 \propto n^{2\theta}$ was observed in Ref. [20] with $\theta = 0.5680(15)$ for finite system sizes. This observation lead to the assumption that the whole distribution follows the scaling form

$$P_n(L) = \langle L\rangle g(\langle L\rangle L), \tag{13}$$

with a not explicitly known function g. Even more, Ref. [20] suggests that their measurements can be explained, instead of the exponent θ, by a logarithmic correction to a square-root scaling:

$$\langle L\rangle \approx \frac{1}{e}\sqrt{n}\ln n + \frac{1}{2}\sqrt{n}. \tag{14}$$

Using our data for the tails of the distribution, we can test whether this scaling holds over the whole distribution or only in the main region. If we rescale the axis of the plot suitably, the distributions for different sizes n should collapse on the scaling function g, in the case that Eq. (13) holds. Note that the related problem of the longest weakly increasing subsequence for RW increments of ± 1 is known to scale also with $\sqrt{n}$ but does not exhibit the logarithmic correction [19]. Our collapse in Fig. 7 following Eq. (13) supports the validity of Eq. (14). The collapse does work except for the very far tails, which is an effect—at least partially—caused by finite-size effects, since the length of the LIS can for finite n never be longer than n. This pattern occurs often when looking at the far tails of discrete systems, e.g., for the convex hull of RWs on lattices in Refs. [31–34] or in a toy model for noninteracting Fermions in a landscape with n random energy levels [28].

Since for the rate functions characterizing the LIS length distribution of RWs there is no known result, we use our numerical data to give a rough estimate of the rate function. Therefore we look into the empirical rate function $\Phi_n(L) = \frac{1}{n}\ln P_n(L)$, which is plotted in Fig. 8 for the data already shown in Fig. 6.

Using the empirical rate function we can obtain the asymptotics of the rate function from our data. Note that to estimate the right-tail rate function we use the intermediate tail and not the far tail, which is bending up due to finite-size effects, as the very long LISs are suppressed by the hard limit of $L \leqslant n$.. Since we are interested only in the leading-order exponent of the rate function, i.e., assuming $\Phi(L) \propto L^{\kappa}$ for very small and very large values of L, we can rescale the axes arbitrarily due to the scale invariance of power laws. For convenience we look at $x = L/L_{\max}$ to limit the range to the interval [0,1]. For the left tail we observe a leading-order behavior of the rate function of approximately $\Phi(L) \sim L^{-1.6}$ and for the right tail $\Phi(L) \sim L^{2.9}$. Note that the exponent of the left tail is clearly distinct from 3, such that it does not show signs of a third-order phase transition. Also it does not show a Tracy-Widom distribution in the main region (also see Ref. [20]), which is

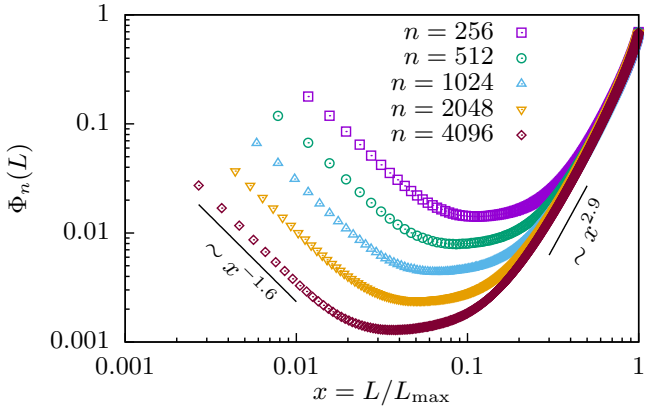

FIG. 8. Empirical rate function $\Phi_n(L)$ for the length of the LIS of RWs. The two straight lines are obtained by power-law fits and show the leading-order behavior of the rate function for each tail. (For clarity not every data point is shown.)

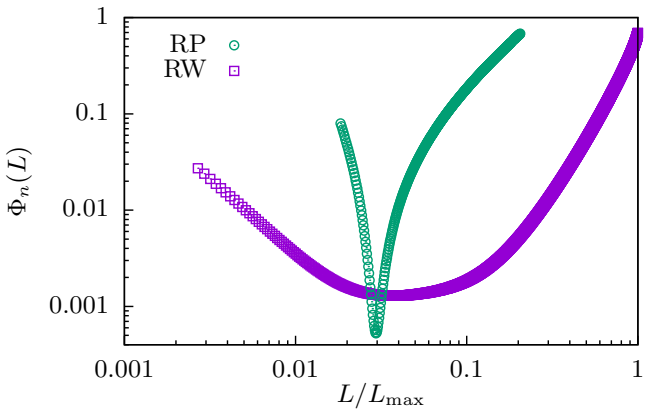

FIG. 9. Direct comparison of the empirical rate function $\Phi_{4096}(L)$ of the RP and the RW.

consistent with the expectation that these two properties do occur together [17].

A comparison of this leading-order behavior to the behavior of the RP case, as visualized in Fig. 5, shows that the tails decay differently. For a direct comparison of our results consider Fig. 9. While the right-tail exponent is larger in the RW case, the probability density decays slower (i.e., the empirical rate function increases slower). This apparent contradiction is understandable when considering that the rate function of the RP case grows much faster near the minimum at $\langle L \rangle$ before it settles into the asymptotic behavior. The RW case behaves exactly opposite, such that the branches left and right of the minimum show opposite curvature in the two cases. Generally, this leads to a distribution $P(L)$ which is much broader in the RW case, especially towards quite large values of L.

IV. CONCLUSIONS

We obtain numerical data for the distribution of the length of the longest increasing subsequence for two cases of sequences of random numbers, namely, for RPs and for one-dimensional RWs. By applying sophisticated large-deviation algorithms, we are able to sample the distributions over literally hundreds of decades in probability. The case of RPs is already well studied analytically in the literature, and we are able to confirm, to the best of our knowledge, for the first time these analytical results. Since our data are gathered for finite system sizes, we can observe a rather fast convergence to the analytical results valid in the $n \to \infty$ limit. These results also show the validity and convergence of our simulations. For the case of RWs we can observe the leading-order behavior of the rate function far into the tails. Also our data support the scaling assumption Eq. (13) [20] for the whole distribution including the logarithmic term, which is not present for weakly increasing subsequences in RWs with ± 1 steps [19]. This result could be used to guide analytical work on this topic and to test future analytical results. A direct comparison of the empirical rate functions in the tails shows qualitatively very different behavior. While the rate function of the RW seems to be a convex function, the RP case consists, in principle, of two concave parts.

A possible future direction extending this work would be an interpolation between the RP and RW case, where one could observe the change of the exponents governing the rate function. Since a set of distinct random numbers δ_j drawn uniformly from $[-1, 1]$ should show the same statistics for the longest increasing subsequence of a RP, we could introduce a parameter c governing the correlation length. The sequence would be constructed as $S_i = \sum_{j=\max(0,i-c)}^{i} \delta_j$. For $c = 0$ this would correspond to a RP and for $c = n$ to a RW. In addition to this simple type of correlation, one could study power-law correlated random numbers or increments, leading possibly to even more complicated behavior.

Furthermore, it is of interest to analyze the actual LIS in particular with taking the degeneracy into account. For this purpose one must use a dynamic programming approach, which exhibits a running time of $O(n^2)$ instead of the $O(n \ln n)$ complexity of the algorithm which obtains just the length of the LIS. Nevertheless, the dynamic programming approach would allow one to compare different LISs for every realization of the sequence, whether they are rather similar or possibly very different, depending on the type of sequence. Also one could study the distribution of the LIS entropy with similar large-deviation techniques as applied here. Furthermore, this would allow to measure a correlation between LIS length and span in a statistical unbiased way, going beyond the results shown in Fig. 3.

ACKNOWLEDGMENTS

We are indebted to Satya N. Majumdar, who brought this problem to our attention, and for his feedback on a draft of this manuscript. We also want to thank J. Ricardo G. Mendonça for useful discussions and his feedback on a draft of the manuscript. Further we thank Christoph Norrenbrock for helpful comments and advice during the preparation of the manuscript. We acknowledge usage of the HPC facilities of the GWDG Göttingen and the CARL cluster in Oldenburg funded by the DFG (INST 184/157-1 FUGG) and the Ministry of Science and Culture (MWK) of the Lower Saxony State. H.S. acknowledges support by DFG grant HA 3169/8-1.

[1] D. Romik, *The Surprising Mathematics of Longest Increasing Subsequences* (Cambridge University Press, New York, 2015).
[2] S. M. Ulam, in *Modern Mathematics for the Engineer: Second Series*, edited by E. Beckenbach and M. Hestenes, Dover Books on Engineering Series (Dover Publications, New York, 2013), Chap. 11, pp. 261–281.
[3] D. Aldous and P. Diaconis, Bull. Am. Math. Soc. **36**, 413 (1999).
[4] T. Seppäläinen, Probab. Theory Relat. Fields **112**, 221 (1998).
[5] B. F. Logan and L. A. Shepp, Adv. Math. **26**, 206 (1977).
[6] J.-D. Deuschel and O. Zeitouni, Comb. Probab. Comput. **8**, 247 (1999).
[7] J. Baik, P. Deift, and K. Johansson, J. Am. Math. Soc. **12**, 1119 (1999).
[8] M. Prähofer and H. Spohn, Phys. Rev. Lett. **84**, 4882 (2000).

[9] S. N. Majumdar and S. Nechaev, Phys. Rev. E **69**, 011103 (2004).
[10] K. Johansson, Commun. Math. Phys. **209**, 437 (2000).
[11] J. Baik and E. M. Rains, J. Stat. Phys. **100**, 523 (2000).
[12] T. Kriecherbauer and J. Krug, J. Phys. A: Math. Theor. **43**, 403001 (2010).
[13] S. N. Majumdar, in *Complex Systems: Lecture Notes of the Les Houches Summer School 2006*, edited by J. Bouchaud, M. Mézard, and J. Dalibard (Elsevier Science, Les Houches, 2006), Chap. 4, pp. 179–216.
[14] I. Corvin, Random Matrices: Theory Appl. **01**, 1130001 (2012).
[15] K. A. Takeuchi and M. Sano, Phys. Rev. Lett. **104**, 230601 (2010).
[16] K. A. Takeuchi, M. Sano, T. Sasamoto, and H. Spohn, Sci. Rep. **1**, 34 (2011).
[17] S. N. Majumdar and G. Schehr, J. Stat. Mech.: Theory Exp. (2014) P01012.
[18] P. L. Doussal, S. N. Majumdar, and G. Schehr, EPL (Europhys. Lett.) **113**, 60004 (2016).
[19] O. Angel, R. Balka, and Y. Peres, Math. Proc. Cambridge Philos. Soc. **163**, 173 (2017).
[20] J. R. G. Mendonça, J. Phys. A: Math. Theor. **50**, 08LT02 (2017).
[21] A. K. Hartmann, *Big Practical Guide to Computer Simulations* (World Scientific, Singapore, 2015).
[22] A. K. Hartmann, Phys. Rev. E **65**, 056102 (2002).
[23] A. K. Hartmann, Eur. Phys. J. B **84**, 627 (2011).
[24] A. K. Hartmann, Phys. Rev. E **89**, 052103 (2014).
[25] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[26] M. Newman and G. Barkema, *Monte Carlo Methods in Statistical Physics* (Oxford University Press, New York, 1999), pp. 3–86.
[27] M. Prähofer and H. Spohn, J. Stat. Phys. **115**, 255 (2004).
[28] H. Schawe, A. K. Hartmann, S. N. Majumdar, and G. Schehr, EPL (Europhys. Lett.) **124**, 40005 (2018).
[29] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).
[30] H. Touchette, Phys. Rep. **478**, 1 (2009).
[31] G. Claussen, A. K. Hartmann, and S. N. Majumdar, Phys. Rev. E **91**, 052104 (2015).
[32] H. Schawe, A. K. Hartmann, and S. N. Majumdar, Phys. Rev. E **96**, 062101 (2017).
[33] H. Schawe, A. K. Hartmann, and S. N. Majumdar, Phys. Rev. E **97**, 062159 (2018).
[34] H. Schawe and A. K. Hartmann, J. Phys.: Conf. Ser. (to be published).