How fast does a random walk cover a torus?

Peter Grassberger

JSC, FZ Jülich, D-52425 Jülich, Germany

We present high statistics simulation data for the average time $\langle T_{\text{cover}}(L) \rangle$ that a random walk needs to cover completely a two-dimensional torus of size $L \times L$. They confirm the mathematical prediction that $\langle T_{\text{cover}}(L) \rangle \sim (L \ln L)^2$ for large L, but the prefactor *seems* to deviate significantly from the supposedly exact result $4/\pi$ derived by Dembo *et al.* [Ann. Math. **160**, 433 (2004)], if the most straightforward extrapolation is used. On the other hand, we find that this scaling does hold for the time $T_{N(t)=1}(L)$ at which the average number of yet unvisited sites is 1, as also predicted previously. This might suggest (wrongly) that $\langle T_{\text{cover}}(L) \rangle$ and $T_{N(t)=1}(L)$ scale differently, although the distribution of rescaled cover times becomes sharp in the limit $L \to \infty$. But our results can be reconciled with those of Dembo *et al.* by a very slow and *nonmonotonic* convergence of $\langle T_{\text{cover}}(L) \rangle / (L \ln L)^2$, as had been indeed proven by Belius *et al.* [Probab. Theory Relat. Fields **167**, 461 (2017)] for Brownian walks, and was conjectured by them to hold also for lattice walks.

The problem of how fast a random walk covers a two-dimensional torus was introduced in the mathematical literature by Wilf [1], who called it the "white screen problem." But it is also of considerable interest for other sciences, as it relates, e.g., to how fast a grazing animal can collect as much food as possible [2–4], or how fast information can be spread on or collected from a network (such as a mobile *ad hoc* network) whose topology is not known [5–7]. For that reason, it has also been discussed extensively in the statistical physics literature [8–13].

Let us denote by $\langle T_{\text{cover}}(L) \rangle$ the average time needed to cover a torus of $L \times L$ sites completely, and by $T_{N(t)=1}(L)$ the time at which the average number of yet uncovered sites is 1. Naively, one would expect that both diverge in the same way with L, at least if the distribution of cover times is not too broad.

Aldous [14,15] proved that

$$\langle T_{\text{cover}}(L) \rangle \lesssim \frac{4}{\pi} L^2 \ln^2 L, \tag{1}$$

and proved that the rescaled time $T_{\text{cover}}(L)/(L \ln L)^2$ is indeed δ distributed in the limit $L \to \infty$. He furthermore conjectured that Eq. (1) becomes sharp in this limit.

This conjecture was supported by heuristic arguments in Refs. [9,10], where the main quantity of interest was not $\langle T_{\text{cover}}(L) \rangle$ but $T_{N(t)=1}(L)$. These authors argued convincingly that

$$T_{N(t)=1}(L)/(L \ln L)^2 \to \frac{4}{\pi} \quad \text{for } L \to \infty, \tag{2}$$

and then conjectured that the same is true also for the cover times, because mean cover times and times at which the average number of uncovered sites is 1 should scale in the same way.

The story was seemingly closed when Dembo *et al.* [16] proved rigorously that

$$\lim_{L \to \infty} \frac{T_{\text{cover}}(L)}{(L \ln L)^2} = \frac{4}{\pi} \quad \text{in probability}, \tag{3}$$

i.e., Aldous' inequality Eq. (1) is saturated and the limit distribution is indeed sharp.

When I reconsidered this problem, I was primarily interested in the way how "true self-avoiding" walks (or "self-repelling walks") [17] cover the torus or any other finite lattice [7,18], and wanted just to document the dramatic difference between self-repelling and ordinary random walks. However, soon after I started to simulate ordinary random walks on the 2-torus, it became clear that the data agreed with Eqs. (1) and (2), but not easily with Eq. (3).

The results presented in the following come from simulations that altogether took about 1 year of CPU time on modern workstations. Lattice sizes ranged from $L = 16$ to $L = 65\,536$ in steps of powers of 2. The number of walks simulated varied between $\approx 4 \times 10^7$ for $L = 16$ and 1350 for $L = 65\,536$. For easier coding and faster codes, boundary conditions (BCs) were not strictly periodic but helical [19]. For large L the difference is negligible. In particular, also for helical BCs the lattice is a torus, and the difference with periodic BCs is just that one of the coordinate axes is slightly tilted. We verified that the results obtained with periodic BCs were identical within statistical errors for $L \geqslant 16$. We also tested two different random number generators (Ziff's four-tap generator [20] and the UNIX generator rand48), again with no significant differences.

Results for $\langle T_{\text{cover}}/(L \ln L)^2 \rangle$ against L are shown in Fig. 1. Whenever error bars are not visible on the data points, they are smaller than the line thickness. Also shown is the prediction of Dembo *et al.* [16] (horizontal line) and a fit for large L. This fit is a least-square fit (with all three constants fitted) to all data with $L \geqslant 128$, but the quoted error in the first term is much bigger that the purely statistical error, in order to include plausible further correction terms, where we assume that "plausible" correction terms do not ruin the monotonicity. Our first conclusion is thus that

$$\lim_{L \to \infty} \frac{\langle T_{\text{cover}}(L) \rangle}{(L \ln L)^2} = 1.2473 \pm 0.0012. \tag{4}$$

The right-hand side disagrees with the supposedly exact value $4/\pi = 1.2732\ldots$ by about 22 standard deviations (similar results have been obtained in Ref. [12], albeit with less statistics). This discrepancy can hardly be blamed on statistical fluctuations (the likelihood being about 10^{-100}). It cannot be
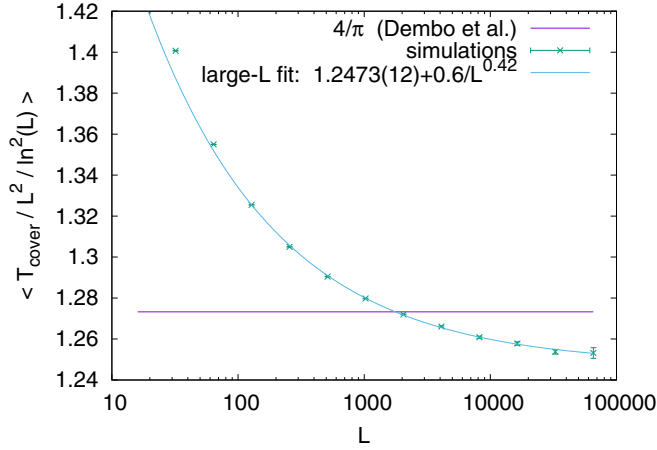
FIG. 1. Log-linear plot of average cover time for 2-tori of size $L \times L$, plotted against L. Whenever no error bars are visible on the data points, they are smaller than the line thickness. The fit is for all data with $L \geqslant 128$. The error of the leading term is supposed to take into account the possibility of further corrections to scaling that would, however, leave the cover times monotonically decreasing with L.

blamed on the used random number generators, both of which have been proven to be reliable even in problems involving much higher statistics. In view of the extreme simplicity of the code (about one page), also a programming error is very unlikely.

A next problem that could cause a wrong asymptotic estimate could be a very skewed and broad distribution of cover times. But the distribution of normalized cover times is expected [21] to be a (randomly shifted) Gumbel distribution in the limit $L \to \infty$. This gives a roughly exponential tail, which could not significantly bias any estimates of average cover times.

In any case, in Figs. 2 and 3 we show such distributions. They seem to be indeed exponentially cut off at large times, and definitely do not suggest that estimates of the averages could be influenced significantly by large T tails.
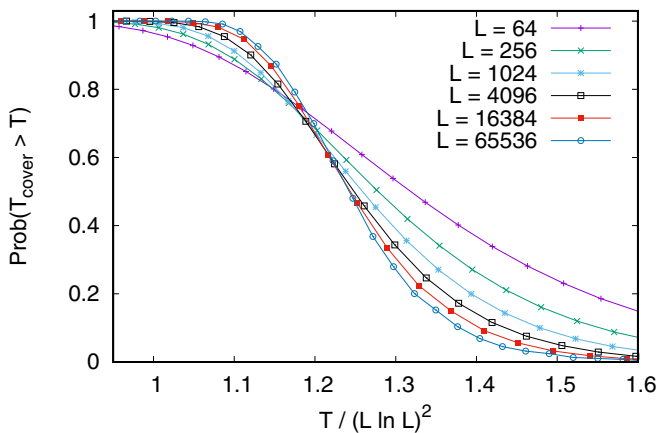

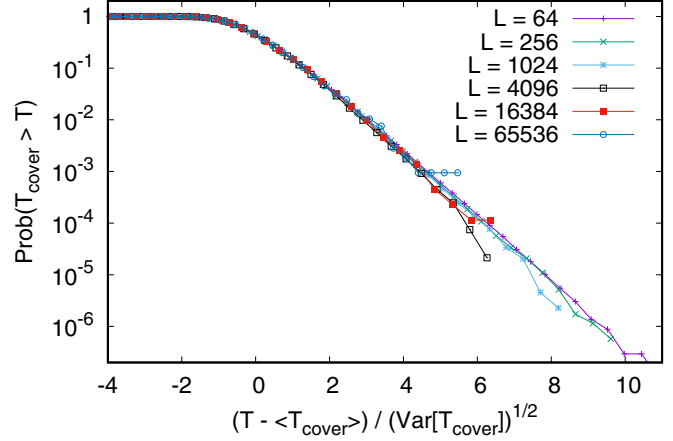FIG. 2. Cumulative distributions $\mathrm{Prob}(T_{\mathrm{cover}} > T)$ for six values of L.

FIG. 3. Same data as in Fig. 2, but (i) plotted on a logarithmic y scale, and (ii) plotted against $(T - \langle T_{\mathrm{cover}} \rangle)/\{\mathrm{Var}[T_{\mathrm{cover}}]\}^{1/2}$.

To add to the last point, we show in Fig. 4 our estimates of the relative fluctuations of T_{cover}, defined as $\{\mathrm{Var}[T_{\mathrm{cover}}]\}^{1/2}/\langle T_{\mathrm{cover}} \rangle$. We see that they decrease with L, as predicted by Aldous, although our data are not precise enough to distinguish between a power-law decay with a very small exponent (≈ 0.11) and a logarithmic behavior.

To shed more light on this problem, we considered next the average number $N(t)$ of uncovered sites at time t. For $t \ll L^2$, the number of *covered* sites is independent of L, and given asymptotically by [22]

$$s(t) \equiv L^2 - N(t) = \frac{\pi t}{\ln t}\left[1 + O\left(\frac{\ln \ln t}{\ln t}\right)\right]. \quad (5)$$

The finiteness of the lattice becomes relevant for $t \approx L^2$, and for $t \gg L^2$ the decay of $N(t)$ is purely exponential [9]. The crossover between these two regimes is shown in Fig. 5. There we show on the y axis not $N(t)/L^2$ itself, but we multiplied it with $\exp[t/\tau(L)]$, where the characteristic time $\tau(L)$ (the inverse decay rate) was estimated from fits in the regime $L^2 < t < (L \ln L)^2$. The quality of the exponential decay in this regime is illustrated in Fig. 6 for $L = 8192$ (but similarly nice exponentials were also found for all other lattice sizes). In Fig. 6 we plotted $N(t)$ itself, and we verified that the
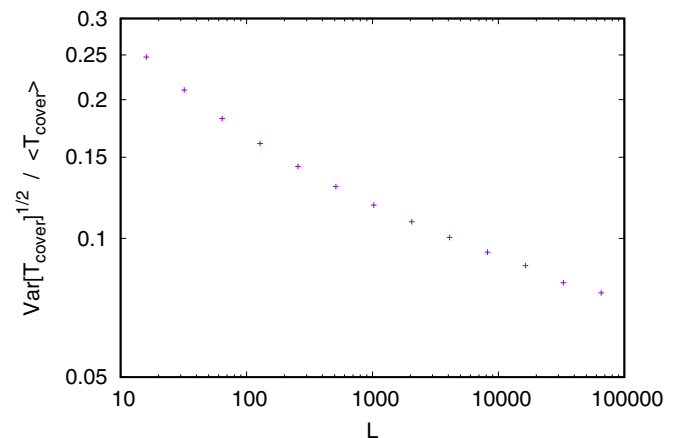

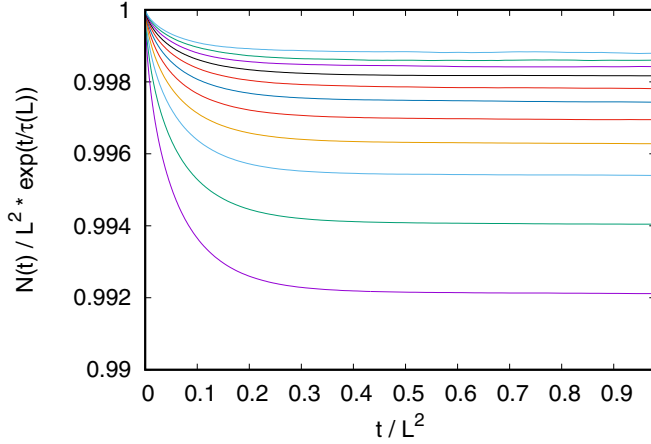FIG. 4. Relative fluctuations of cover times, plotted against L.

FIG. 5. The average number $N(t)$ of uncovered sites at time, plotted for different values of L in the regime $t < L^2$. For clarity, we show on the y axis not $N(t)$ itself but $N(t)/L^2 \exp[t/\tau(L)]$, where $\tau(L)$ is the numerically found inverse decay rate of $N(t)$ for $t \gg L^2$. The uppermost curve is for $L = 65\,536$, and the lowest is for $L = 64$.

exponential decay continued also for $t \gg (L \ln L)^2$, although statistical errors increase rapidly for large t.

This purely exponential decay can be used to determine $\tau(L)$ either by a fit in the regime $L^2 < t < (L \ln L)^2$ or by just finding the value of t where $N(t) = 1$. In the second method we of course have to take into account that the exponential decay holds only for $t > L^2$, but this correction becomes negligible for $L \to \infty$, i.e.,

$$T_{N(t)=1} = 2\tau(L) \ln L[1 + O(1/\ln^2 L)]. \tag{6}$$

Direct numerical estimates of $T_{N(t)=1}/(L \ln L)^2$ are shown in Fig. 7. We see a much slower (probably logarithmic) convergence than for average cover times, but the data are completely compatible with Eq. (2).

A last reason for a wrong asymptotic estimate would be a very slow (and nonmonotonic) convergence with L. We found no indication for this in our data, but it is conjectured in
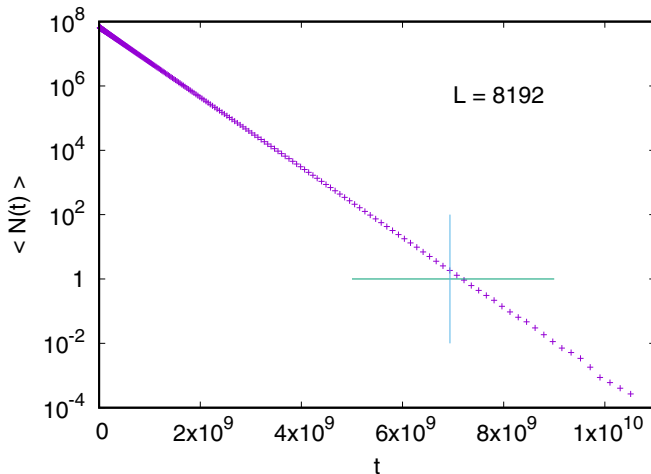

FIG. 6. The average number $N(t)$ of uncovered sites at time, plotted against t, for $L = 8192$. The horizontal and vertical straight lines indicate the values $N(t) = 1$ and $t = \langle T_{\text{cover}} \rangle$.
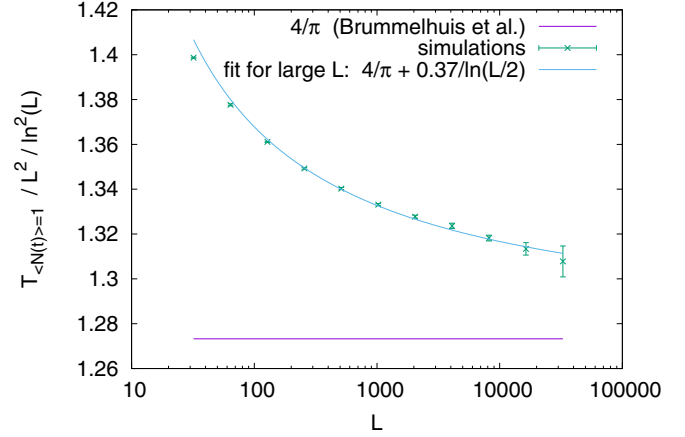

FIG. 7. Direct estimates of $T_{N(t)=1}/(L \ln L)^2$ plotted against L on a log-linear plot. The fit just demonstrates that the data are compatible with Eq. (2).

Ref. [23] that the behavior for walks on the square lattice is as for off-lattice Brownian walks, which would suggest [23]

$$\frac{\pi \langle T_{\text{cover}}(L) \rangle}{4(L \ln L)^2} = 1 - \frac{1}{2} \ln \ln L / \ln L + D/\ln L + o(1/\ln L), \tag{7}$$

with an unknown constant D (indeed, the conjecture in Ref. [23] for lattice walks was slightly weaker). In Fig. 8 we show the data shown already in Fig. 1, together with two analytic curves representing Eq. (7): one with $D = 0$, and the other with $D = 2$. We see that the latter gives a very good fit, from which we conclude that the mathematical predictions are presumably all correct, and $D = 2.02(2)$. We should warn, however, that we could also give decent fits with different coefficients of the $\ln \ln L / \ln L$ term (and, of course, different D).

Finally, we show in Fig. 9 the ratios $T_{N(t)=1}/\langle T_{\text{cover}} \rangle$. For very small L, they are < 1, because the large-T tails contribute more to $\langle T_{\text{cover}} \rangle$ than to $T_{N(t)=1}$. For larger L, this effect is outweighed by the fact that $N(\langle T_{\text{cover}} \rangle) > 1$ because walks that do not yet cover at $t = \langle T_{\text{cover}} \rangle$ might have $\gg 1$ uncovered
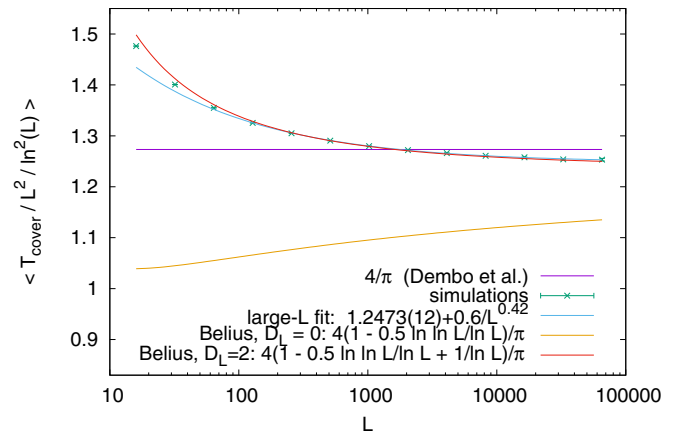

FIG. 8. Same data as in Fig. 1, but with two additional analytic curves. Both represent Eq. (7), one with $D = 0$ and the other with $D = 2$.
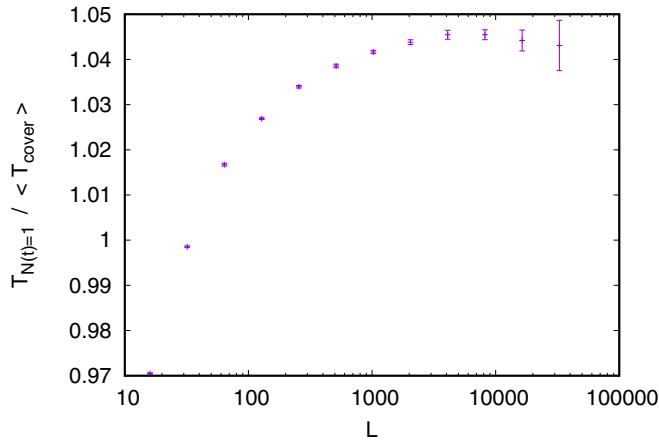
FIG. 9. Ratios $T_{N(t)=1}/\langle T_{\mathrm{cover}}\rangle$ plotted against L on a log-linear plot.

sites. Finally, at very large L, the ratio seems to decrease again, although this is not significant in view of the large error bars. Yet it suggests that the ratio converges to 1 for $L \to \infty$, which would completely reconcile our data with the mathematical proofs. This is supported by the fact that the $\ln \ln L / \ln L$ term is absent in $T_{N(t)=1}$.

In summary, our numerical data suggest *at face value* that $T_{N(t)=1}$ and $\langle T_{\mathrm{cover}}\rangle$ do not scale in the same way with L, in contrast to rigorous proofs. But they can be reconciled with the proofs if the (predicted) corrections to scaling are taken into account. As a result, the convergence towards the asymptotic behavior should be extremely slow (and nonmonotonic). Thus, without knowing the subleading terms, attempts to verify the leading behavior numerically would be futile.

The present paper can be seen as a warning that supposedly rigorous proofs *can* be wrong (and should thus be checked numerically), but more so as a warning that extrapolations of numerical data can be very subtle and misleading, even if they *look* completely benign and harmless. The vast number of wrong critical exponent estimates found in the literature bears ample witness to that. Combining rigorous mathematics and numerics can be useful if, as in the present case, the mathematics exclude too naive parametrizations, and the numerics can suggest the value(s) of constants that remain undetermined by the mathematical arguments.

[1] H. S. Wilf, The editor's corner: The white screen problem, Am. Math. Mon. **96**, 704 (1989).

[2] G. M. Viswanathan, S. V. Buldyrev, S. Havlin, M. G. E. Da Luz, E. P. Raposo, and H. E. Stanley, Optimizing the success of random searches, Nature (London) **401**, 911 (1999).

[3] M. C. Santos, G. M. Viswanathan, E. P. Raposo, and M. G. E. da Luz, Optimization of random searches on regular lattices, Phys. Rev. E **72**, 046143 (2005).

[4] O. Bénichou, C. Loverdo, M. Moreau, and R. Voituriez, Intermittent search strategies, Rev. Mod. Phys. **83**, 81 (2011).

[5] A. N. Mian, R. Beraldi, and R. Baldoni, On the coverage process of random walk in wireless ad hoc and sensor networks, in *The 7th IEEE International Conference on Mobile Ad-hoc and Sensor Systems (IEEE MASS 2010)* (IEEE, New York, 2010), p. 146.

[6] K. Li, Performance analysis and evaluation of random walk algorithms on wireless networks, Int. J. Found. Comput. Sci. **23**, 779 (2012).

[7] C. Avin and B. Krishnamachari, The power of choice in random walks: An empirical study, Comput. Networks **52**, 44 (2008).

[8] A. M. Nemirovsky, H. O. Mártin, and M. D. Coutinho-Filho, Universality in the lattice-covering time problem, Phys. Rev. A **41**, 761 (1990).

[9] M. J. A. M. Brummelhuis and H. J. Hilhorst, Covering of a finite lattice by a random walk, Physica A **176**, 387 (1991).

[10] M. J. A. M. Brummelhuis and H. J. Hilhorst, How a random walk covers a finite lattice, Physica A **185**, 35 (1992).

[11] K. R. Coutinho, M. D. Coutinho-Filho, M. A. F. Gomes, and A. M. Nemirovsky, Partial and Random Lattice Covering Times in Two Dimensions, Phys. Rev. Lett. **72**, 3745 (1994).

[12] J. R. G. Mendonça, Numerical evidence against a conjecture on the cover time of planar graphs, Phys. Rev. E **84**, 022103 (2011).

[13] M. Chupeau, O. Bénichou, and R. Voituriez, Cover times of random searches, Nat. Phys. **11**, 844 (2015).

[14] D. Aldous, Probability approximations via the poisson clumping heuristic, *Applied mathematical science*, Vol. 77 (Springer, New York, 1989).

[15] D. Aldous, Threshold limits for cover times, J. Theor. Probab. **4**, 197 (1991).

[16] A. Dembo, Y. Peres, J. Rosen, and O. Zeitouni, Cover times for Brownian motion and random walks in two dimensions, Ann. Math. **160**, 433 (2004).

[17] D. J. Amit, G. Parisi, and L. Peliti, Asymptotic behavior of the "true" self-avoiding walk, Phys. Rev. B **27**, 1635 (1983).

[18] H. Freund and P. Grassberger, How uniformly a random walker covers a finite lattice, Physica A **192**, 465 (1993).

[19] M. E. J. Newman and G. T. Barkema, *Monte Carlo Methods in Statistical Physics* (Oxford University Press, New York, 1999).

[20] R. M. Ziff, Four-tap shift-register-sequence random-number generators, Comput. Phys. **12**, 385 (1998).

[21] D. Belius, Gumbel fluctuations for cover times in the discrete torus, Probab. Theory Relat. Fields **157**, 635 (2013).

[22] A. Dvoretsky and P. Erdös, Some problems on random walk in space, in *Proceedings of the 2nd Berkeley Symposium of Mathematical Statistics and Probability* (Princeton University Press, Princeton, NJ, 1951), p. 353.

[23] D. Belius and N. Kistler, The subleading order of two dimensional cover times, Probab. Theory Relat. Fields **167**, 461 (2017).