Accurate Estimate of the Critical Exponent ν for Self-Avoiding Walks via a Fast Implementation of the Pivot Algorithm

Nathan Clisby*

ARC Centre of Excellence for Mathematics and Statistics of Complex Systems, Department of Mathematics and Statistics, The University of Melbourne, Victoria 3010, Australia

(Received 8 March 2009; revised manuscript received 14 January 2010; published 1 February 2010)

We introduce a fast implementation of the pivot algorithm for self-avoiding walks, which we use to obtain large samples of walks on the cubic lattice of up to 33×10^6 steps. Consequently the critical exponent ν for three-dimensional self-avoiding walks is determined to great accuracy; the final estimate is $\nu = 0.587\,597(7)$. The method can be adapted to other models of polymers with short-range interactions, on the lattice or in the continuum.

DOI: 10.1103/PhysRevLett.104.055702 PACS numbers: 64.60.De, 05.10.−a, 64.70.km, 68.35.Rh

The self-avoiding walk (SAW) model is an important model in statistical physics [1]. It models the excluded-volume effect observed in real polymers, exactly capturing universal features such as critical exponents. It is also the $n \to 0$ limit of the n-vector model, which includes the Ising model ($n = 1$) as another instance, thus serving as an important model in the study of critical phenomena. Exact results are known for self-avoiding walks in two dimensions [2,3] and for $d \geq 4$ (mean-field behavior has been proved for $d \geq 5$ [4]), but not for the most physically interesting case of $d = 3$.

We have efficiently implemented the pivot algorithm via a data structure we call the SAW tree, which allows rapid Monte Carlo simulation of SAWs of millions of steps. We discuss this implementation in general terms here, and then use this implementation to accurately calculate the critical exponent ν for $\mathbb{Z}^3$. More details about the implementation can be found in a companion article [5]. This new algorithm can also be adapted to other models of polymers with short-range interactions, on the lattice and in the continuum, and hence promises to be widely useful.

An N-step SAW on $\mathbb{Z}^d$ is a mapping ω: $\{0, 1, \ldots, N\} \to \mathbb{Z}^d$ with $|\omega(i+1) - \omega(i)| = 1$ for each i ($|x|$ denotes the Euclidean norm of x), and with $\omega(i) \neq \omega(j)$ for all $i \neq j$. We generate three-dimensional SAWs via the pivot algorithm, and calculate various observables which characterize the size of the SAWs: the squared end-to-end distance R_e^2, the squared radius of gyration R_g^2, and the mean-square distance of a monomer from its end points R_m^2, where

$$R_e^2 = |\omega(N) - \omega(0)|^2,$$

$$R_g^2 = \frac{1}{2(N+1)^2} \sum_{i,j=0}^{N} |\omega(i) - \omega(j)|^2,$$

$$R_m^2 = \frac{1}{2(N+1)} \sum_{i=0}^{N} [|\omega(i) - \omega(0)|^2 + |\omega(i) - \omega(N)|^2].$$

We seek to calculate the mean values of these observables over all SAWs of N steps, where each SAW is given equal weight. Their asymptotic forms are expected to be described by

$$\langle R_x^2 \rangle_N = D_x N^{2\nu} \left[1 + \frac{a_1}{N} + \frac{a_2}{N^2} + \cdots + \frac{b_x}{N^{\Delta_1}} + \frac{b_1}{N^{2\Delta_1}} + \cdots + \frac{c_0}{N^{\Delta_2}} + \cdots \right] + \text{af}, \quad (1)$$

with $0 < \Delta_1 < \Delta_2 < \cdots$, and where additional terms of the form $c/N^{k_0 + k_1\Delta_1 + k_2\Delta_2 + k_3\Delta_3 + \cdots}$ ($k_0, k_1, k_2, k_3, \cdots \geq 0$) are not shown. In addition, af indicates terms arising from the antiferromagnetic singularity, which occurs in models on loose-packed lattices such as $\mathbb{Z}^d$; these terms are negligible compared with terms included in fits. The exponents ν, Δ_1, and Δ_2 are universal; i.e., they are dependent only on the dimensionality of the lattice and the universality class of the model, while the amplitudes D_x are observable dependent. However, amplitude ratios, such as D_g/D_e and b_g/b_e, are universal quantities.

The pivot algorithm is a powerful approach to the study of self-avoiding walks, invented by Lal [6] and later elucidated and popularized by Madras and Sokal [7]. From an initial SAW of length N, such as a straight rod, new N-step walks are successively generated by choosing a site of the walk at random, and attempting to apply a lattice symmetry operation, or pivot, to one of the parts of the walk; if the resulting walk is self-avoiding the move is accepted, otherwise the move is rejected and the original walk is retained. The group of lattice symmetries for $\mathbb{Z}^3$ has 48 elements, and we use all of them except the identity as potential pivot operations. Thus a Markov chain is formed in the ensemble of SAWs of fixed length; this chain satisfies detailed balance and is ergodic, ensuring that SAWs are sampled uniformly at random. Furthermore, as demonstrated by Madras and Sokal [7] through strong heuristic arguments and numerical experiments, the Markov chain has a short integrated autocorrelation time for global observables, thus making the pivot algorithm extremely efficient in comparison to Markov chains utilizing local moves. See [7,8] for detailed discussion.

The implementation of Madras and Sokal utilized a hash table to record the location of each site of the walk. They

showed that the pivot algorithm has integrated autocorrelation time $O(N^p)$, with p dimension-dependent but close to zero ($p \lesssim 0.2$), and argued convincingly that the CPU time per successful pivot is $O(N)$ for their implementation.

Madras and Sokal argued that $O(N)$ is best possible because it takes time of order N to merely write down an N-step SAW. However, Kennedy [9] recognized that it is *not* necessary to write down the SAW for each successful pivot, and from clever use of geometric constraints developed an algorithm that broke the $O(N)$ barrier. The CPU time for this implementation grows as a dimension-dependent fractional power of N (see Table I).

We have extended this idea to obtain a radical further improvement: for $\mathbb{Z}^2$ and $\mathbb{Z}^3$ the mean CPU time per attempted pivot, which we denote $T(N)$, is now only $O(\log N)$ for the range of N studied, and we have a theoretical argument that the large N behavior is $O(1)$. The key observation is that although there are typically $O(N)$ nearest neighbor contacts for a SAW of length N, the number of contacts between two halves of a SAW is typically $O(1)$, as shown via renormalization group [10] and Monte Carlo [11] methods. When we attempt to pivot part of a SAW, it is guaranteed that each of the two subwalks remain self-avoiding, and hence we only need to determine if the subwalks intersect. If the resulting walk is self-avoiding, then we expect, on average, that there will be a constant number of contacts between the two subwalks.

We will now briefly discuss the relevant data structure and algorithms; full details can be found in [5]. We implement a binary tree data structure (see, e.g., [12]) which we call a SAW tree. The root node of the SAW tree contains information about the whole walk, including R_e^2, R_g^2, R_m^2, and its minimum bounding box, which is the smallest rectangular prism with faces of the form $x_i = c$ which completely contains the walk. The two children of the root node are valid SAW trees, and contain bounding box information for the first and second halves of the SAW, etc., until the leaves of the tree store individual sites. The SAW tree is related to the R tree [13], a data structure used in the field of computational geometry, but with additional information encoding the state of the SAW. Thus far the SAW tree has been implemented for $\mathbb{Z}^d$, but can be straightforwardly adapted to other lattices and the continuum, as well as other polymer models with short-range interactions. To guarantee optimal performance, we implement the SAW tree so that it is balanced, i.e., so that the depth never exceeds some fixed constant times $\log N$. We define the *level* of a node as the number of generations between a node and the leaves.

TABLE I. $T(N)$, mean time per attempted pivot for N-step SAWs.

Lattice	Madras and Sokal	Kennedy	This Letter
Square	$O(N^{0.81})$	$O(N^{0.38})$	$O(1)$
Cubic	$O(N^{0.89})$	$O(N^{0.74})$	$O(1)$

Bounding boxes enable us to rapidly determine if two subwalks intersect after a pivot attempt: if two bounding boxes do not intersect, then the subwalks which they contain cannot intersect. If a pivot attempt is successful, then it is necessary to resolve all intersections between bounding boxes of different nodes in the tree on opposite sides of the pivot site. Our implementation ensures that intersection tests are typically performed between bounding boxes of nodes which are at the same level. We argue in [5] that the nodes at fixed level in the SAW tree form a renormalized walk, and the intersections between bounding boxes correspond to contacts in the original walk. This implies that at each level there are $O(1)$ intersections, and as the tree has $O(\log N)$ levels this leads to the conclusion that a successful pivot takes time $O(\log N)$. Successful pivots occur with probability $O(N^{-p})$, so overall mean time spent on successful pivots is $O(N^{-p} \log N)$. When a pivot attempt is unsuccessful, with high probability the first intersection occurs near the pivot site. Thus only a small fraction of the SAW tree needs to be traversed to find the intersection, and we argue in [5] that this takes mean time $O(1)$. Unsuccessful pivots occur with probability $O(1)$, and so the overall behavior is $T(N) = O(N^{-p} \log N + 1) = O(1)$. In Fig. 1 we show $T(N)$ for $\mathbb{Z}^2$ and $\mathbb{Z}^3$ from a separate data run, with maximum length $N = 2^{28} - 1 \approx 2.68 \times 10^8$. In both cases it is apparent there is a crossover due to the shorter latency of cache versus main memory. In [5] we argue that $O(1)$ behavior may be reached only for very large N, which makes interpretation of Fig. 1 difficult. For $\mathbb{Z}^2$ some curvature is visible, and the trend appears consistent with $T(N)$ approaching a constant for sufficiently large N. The exponent p is smaller for $\mathbb{Z}^3$ ($p \approx 0.11$) compared with $\mathbb{Z}^2$ ($p \approx 0.19$); hence, the approach to a constant is far slower, and in fact almost no curvature is visible for $\mathbb{Z}^3$. We believe the numerical evidence provides a strong case that $T(N)$ is at most $O(\log N)$, and is consistent with $T(N) = O(1)$; see [5] for more details.

$T(N)$ is shown for the various implementations in Table I. For SAWs of length $N = 10^6$ on the cubic lattice, the performance gain for our implementation is approxi-

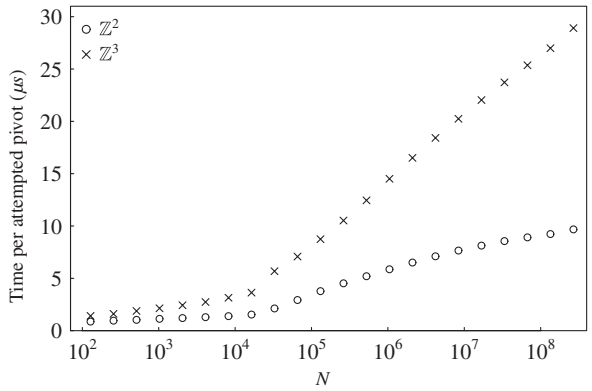

FIG. 1. $T(N)$ for $\mathbb{Z}^2$ and $\mathbb{Z}^3$. Note that these estimates were obtained in a separate data run on a different computer from the main experiment, with lengths from $N = 2^7 - 1$ to $N = 2^{28} - 1$.

mately 200 when compared with Kennedy's, and over a thousand when compared with that of Madras and Sokal. The dramatic performance gain from the new implementation not only makes it possible to obtain large samples of walks with millions of steps, it also makes the regime of very long walks, of up to 10^9 steps, accessible to computer experiments.

For SAWs of length N, it is expected that the exponential autocorrelation time is approximately $O(N/f)$ [7], where f is the fraction of pivot attempts which are successful. The first $20N/f$ configurations were discarded, ensuring that for all practical purposes SAWs were sampled from the uniform distribution. Batch estimates of $\langle R_e^2 \rangle_{33554431}$, using a batch size of 10^8, are shown in Sec. 4 of [5]; in this case the first 50 batches were discarded, while initialization bias is visually apparent for (at most) the first 10 batches.

The computer experiment was performed on a cluster of AMD Opteron Barcelona 2.3 GHz quad core processors, for a total of 16 500 CPU hours. Code was written in C, and compiled with gcc. 10^{11} pivot attempts were made on SAWs of length ranging from 15 to 3.36×10^7, for a grand total of 1.89×10^{13} pivot attempts. Data were collected from every pivot attempt for f and the Euclidean-invariant moments $R_x^{2k} = (R_x^2)^k$ with $x \in \{e, g, m\}$, $1 \le k \le 5$ [14]. The longest walks with $N = 3.36 \times 10^7$ required 3 GB of memory; much longer walks could conceivably be simulated in the future. By comparing fits from the whole data set ($N \le 3.36 \times 10^7$) with fits from a reduced data set ($N \le 2.1 \times 10^6$), we confirmed that data from the longest walks were indeed highly useful in tying down the various estimates (see Fig. 1 in [15]). However, the greatest benefit from the simulation of truly long walks, of say 10^9 steps, may be the ability to directly simulate properties of realistic systems, such as DNA knotting, rather than determination of universal parameters.

Monte Carlo estimates of global parameters are collected in Tables II–V, Sec. 2 of [15], with confidence intervals calculated using the standard binning technique. For all lengths studied the integrated autocorrelation time of the Markov chain is much less than the batch size of 10^8. We confirm the accuracy of the confidence interval estimates by studying the effect of batch size in Sec. 5 of [5].

We estimated the critical exponents ν and $\Delta_1 \approx 1/2$ and associated amplitudes D_x and b_x by fitting the leading term and leading correction of Eq. (1) via weighted nonlinear regression. We truncated the data set by requiring $N \ge N_{\min}$, with $N_{\min}$ a free parameter. We shifted the value of N of Eq. (1) by an amount δN_x to obtain smoother convergence by altering the subleading corrections (see, e.g., [16]); estimates for ν, Δ_1, D_x, and b_x are unaffected in the limit $N_{\min} \to \infty$. With $\delta N_e = 0.35$, $\delta N_g = 1$, $\delta N_m = 0.4$, our final model was

$$\langle R_x^2 \rangle_N = D_x (N + \delta N_x)^{2\nu} \left[1 + \frac{b_x}{(N + \delta N_x)^{\Delta_1}} \right]. \quad (2)$$

Unfortunately we cannot fit the next-to-leading corrections with exponents 1, $\Delta_2 \approx 1$, and $2\Delta_1 \approx 1$ as the differences between them are far too small to resolve. For sufficiently large $N_{\min}$ we found that reduced χ^2 values for all fits approached 1 from above, indicating Eq. (2) is asymptotically correct.

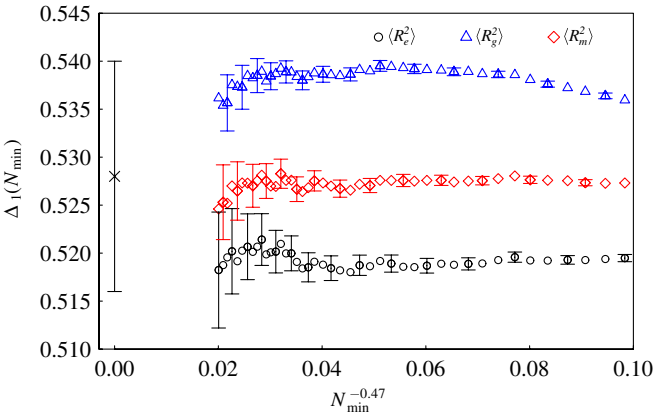

FIG. 2 (color online). Estimates for Δ_1.

Final estimates of parameters have been made directly from Figs. 2 and 3, combining multiple sources visually in an attempt to make estimates robust, and allow the reader to critically evaluate our final results. We do not distinguish between subjective and statistical errors, as we believe that in this context the distinction is itself quite subjective [17]. We provide here some guidance for the interpretation of Figs. 2 and 3, and refer the interested reader to [18] for (much) more information on series analysis.

(i) We plot estimates against $N_{\min}^{-y}$, where y is chosen such that $N_{\min}^{-y}$ is of the same order as the residual error from the fit. The estimates for Δ_1 and b_x have $y = 1 - \Delta_1 \approx 0.47$, and $y = 1$ for ν and D_x.

(ii) We seek to extrapolate the fits to $N_{\min} = \infty$, or $N_{\min}^{-y} = 0$. Depending upon whether the true value y_{exact} is less than, equal to, or greater than y, the estimates would approach a limiting value at $N_{\min}^{-y} = 0$ with infinite, finite, or zero slope, respectively.

(iii) Successive estimates are highly correlated, and so any trend which lies within the error bars should be dis-

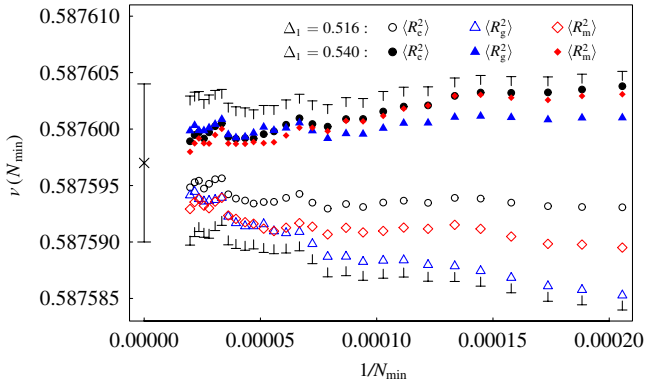

FIG. 3 (color online). Estimates for ν from fits with biased Δ_1. We show the envelope of maximum and minimum values of end points of error bars for all observables.

TABLE II. Comparison of parameter estimates.

Source[a]	ν	Δ_1	D_e	D_g
This Letter	0.587 597(7)	0.528(12)	1.220 35(25)	0.195 14(4)
[16][b] Series	0.587 74(22)		1.217 8(54)	
[19] MC	0.5874(2)			
[20][c] Series	0.587 55(55)		1.225	
[21] FT $d = 3$	0.5882(11)	0.478(10)		
[21] FT ϵ bc	0.5878(11)	0.486(16)		
[22] MCRG	0.587 56(5)	0.5310(33)		
[8][d] MC	0.5877(6)	0.56(3)	1.216 67(50)	0.194 55(7)

[a]Abbreviations: MC ≡ Monte Carlo, FT = Field theory, $d = 3 \equiv d = 3$ expansion, ϵ bc ≡ ϵ expansion with boundary conditions, MCRG ≡ Monte Carlo renormalization group.
[b]Using Eqs. (74) and (75) with $0.516 \leq \Delta_1 \leq 0.54$.
[c]No error estimates were made in [20], but estimates for ν were in the range $0.5870 \leq \nu \leq 0.5881$.
[d]In addition $b_e = -0.483(39)$, $b_g = -0.1143(47)$. D_e, D_g, b_e, and b_g estimates were biased with $\nu = 0.5877$, $\Delta_1 = 0.56$; the confidence intervals were not intended to be taken seriously.

regarded. Only some of the error bars are plotted in order to reduce visual noise.

(iv) There are no bounds on the errors of the truncated asymptotic formulas, and hence the interpretation of the graphs is subjective. The underlying systematic error is observable dependent, and so combining estimates from a variety of observables improves robustness.

In Fig. 2 we plot estimates of Δ_1 with our final result plotted at 0; the error bar reflects the scatter between observables. In Fig. 3 we plot estimates for ν, biased with the lower and upper limits of our range for Δ_1, with our final result at 0. Similar plots for the amplitudes are given in Figs. 5 and 6 of [15]. We have conservatively chosen the error bar for the final result to encompass estimates from all observables $\langle R_x^2 \rangle$. As the amplitudes D_x are highly correlated with estimates for Δ_1, the biasing of Δ_1 greatly extends the range over which stable fits can be obtained. This is the reason the biased fits are preferred over the unbiased fits shown in Fig. 3 of [15].

We report our final results in Table II, and in addition we have $D_m = 0.586\,87(12)$, $b_e = -0.49(5)$, $b_g = -0.1125(125)$, and $b_m = -0.295(30)$. If one assumes the hyperscaling relation $d\nu = 2 - \alpha$, then one also obtains $\alpha = 0.237\,209(21)$. The estimates of ν, D_e, and D_g are in accordance with previous results, although more accurate. The estimate for Δ_1 is more accurate than the previous Monte Carlo value [8], but less accurate than the Monte Carlo renormalization group estimate of [22], which relies upon an uncontrolled although accurate approximation. The claimed accuracy of the field theory estimates [21] for Δ_1 is also comparable, but as discussed by Li *et al.* [8] these calculations have underlying systematic errors of uncertain magnitude. Any desired amplitude ratios can be calculated from the amplitude estimates. The rational number with smallest denominator within 3 standard deviations of $\nu = 0.587\,597$ is $161/274$, suggesting that ν cannot be expressed as a rational number with small denominator.

We would like to stress that, due to the neglect of subleading terms, there are underlying systematic errors in our estimates which are not and cannot be controlled. We have the luxury of high quality data from long walks, and have attempted to be conservative with our claimed errors, but acknowledge there is a risk that the (subjective) confidence intervals may not be sufficiently large.

In summary, an efficient version of the pivot algorithm for SAWs has been implemented and used to calculate ν; the algorithms developed promise to be widely useful in the Monte Carlo simulation of SAWs and related models.

I thank I. G. Enting, A. J. Guttmann, G. Slade, A. Sokal, an anonymous referee, VPAC, and the Australian Research Council.

*n.clisby@ms.unimelb.edu.au

[1] N. Madras and G. Slade, *The Self-Avoiding Walk* (Birkhaüser, Boston, 1993).
[2] B. Nienhuis, Phys. Rev. Lett. **49**, 1062 (1982).
[3] G. F. Lawler, O. Schramm, and W. Werner, in *Proceedings of Symposia in Pure Mathematics* (American Mathematical Society, Providence, 2004), Vol. 72, Part 2, pp. 339–364.
[4] T. Hara and G. Slade, Commun. Math. Phys. **147**, 101 (1992).
[5] N. Clisby (unpublished).
[6] M. Lal, Mol. Phys. **17**, 57 (1969).
[7] N. Madras and A. D. Sokal, J. Stat. Phys. **50**, 109 (1988).
[8] B. Li, N. Madras, and A. D. Sokal, J. Stat. Phys. **80**, 661 (1995).
[9] T. Kennedy, J. Stat. Phys. **106**, 407 (2002).
[10] S. Müller and L. Schäfer, Eur. Phys. J. B **2**, 351 (1998).
[11] M. Baiesi, E. Orlandini, and A. L. Stella, Phys. Rev. Lett. **87**, 070602 (2001).
[12] R. Sedgewick, *Algorithms in C* (Addison-Wesley, Reading, MA, 1998), 3rd ed., Pts. 1–4.
[13] A. Guttman, in *Proceedings of SIGMOD '84*, edited by B. Yormark (ACM Press, New York, 1984), pp. 47–57.
[14] Data from R_x^{2k} with $2 \leq k \leq 5$ have not been analyzed here. These data are inferior for critical exponent estimates, but may in the future be used to calculate universal amplitude ratios.
[15] See supplementary material at http://link.aps.org/supplemental/10.1103/PhysRevLett.104.055702 for additional graphs and raw data.
[16] N. Clisby, R. Liang, and G. Slade, J. Phys. A **40**, 10 973 (2007).
[17] E.g., what value of $N_{\min}$ should be chosen for the statistical error?
[18] A. J. Guttmann, in *Phase Transitions and Critical Phenomena*, edited by C. Domb and J. L. Lebowitz (Academic Press, New York, 1989), Vol. 13, pp. 1–234.
[19] T. Prellberg, J. Phys. A **34**, L599 (2001).
[20] D. MacDonald, S. Joseph, D. L. Hunter, L. L. Moseley, N. Jan, and A. J. Guttmann, J. Phys. A **33**, 5973 (2000).
[21] R. Guida and J. Zinn-Justin, J. Phys. A **31**, 8103 (1998).
[22] P. Belohorec, Ph.D. thesis, University of Guelph, 1997.