



# High-precision estimate of the hydrodynamic radius for self-avoiding walks

Nathan Clisby<sup>1,\*</sup> and Burkhard Dünweg<sup>1,2,3,4</sup>

<sup>1</sup>*School of Mathematics and Statistics, University of Melbourne, Victoria 3010, Australia*

<sup>2</sup>*Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany*

<sup>3</sup>*Condensed Matter Physics, TU Darmstadt, Karolinenplatz 5, 64289 Darmstadt, Germany*

<sup>4</sup>*Department of Chemical Engineering, Monash University, Clayton, Victoria 3800, Australia*

(Received 23 August 2016; published 1 November 2016)

The universal asymptotic amplitude ratio between the gyration radius and the hydrodynamic radius of self-avoiding walks is estimated by high-resolution Monte Carlo simulations. By studying chains of length of up to  $N = 2^{25} \approx 34 \times 10^6$  monomers, we find that the ratio takes the value  $R_G/R_H = 1.5803940(45)$ , which is several orders of magnitude more accurate than the previous state of the art. This is facilitated by a sampling scheme which is quite general and which allows for the efficient estimation of averages of a large class of observables. The competing corrections to scaling for the hydrodynamic radius are clearly discernible. We also find improved estimates for other universal properties that measure the chain dimension. In particular, a method of analysis which eliminates the leading correction to scaling results in a highly accurate estimate for the Flory exponent of  $\nu = 0.58759700(40)$ .

DOI: [10.1103/PhysRevE.94.052102](https://doi.org/10.1103/PhysRevE.94.052102)

## I. INTRODUCTION

A few years ago [1,2], one of the present authors demonstrated significant progress in calculating universal properties of self-avoiding walks (SAWs) [3] on a lattice, which is the standard model to describe the static equilibrium properties of isolated polymer chains in good solvent. These advances were made possible through the use of a recursive data structure called the SAW-tree, which allows for very fast checking of self-overlaps in Monte Carlo (MC) simulations based upon the pivot algorithm [4–7], such that chains with up to  $N \approx 34 \times 10^6$  monomers (repeat units) could be studied. Universal quantities that are accessible include critical exponents such as the Flory exponent  $\nu = 0.587597(7)$  [1], which connects the mean polymer size  $R$  with the degree of polymerization  $N$  via the scaling law  $R \propto bN^\nu$ , where  $b$  is the typical monomer size, and universal amplitude ratios such as the ratio of two different ways to define the size of the coil. The two most popular measures are the mean squared radius of gyration,  $\langle R_G^2 \rangle$ , and the mean squared end-to-end distance,  $\langle R_E^2 \rangle$ . Denoting the coordinates of the monomers by  $\vec{r}_i$ ,  $i = 1, \dots, N$ , the corresponding observables are defined as

$$R_G^2 = \frac{1}{2N^2} \sum_{i,j} |\vec{r}_i - \vec{r}_j|^2 = \frac{1}{N} \sum_i |\vec{r}_i - \vec{R}_{\text{CM}}|^2, \quad (1)$$

$$\vec{R}_{\text{CM}} = \frac{1}{N} \sum_i \vec{r}_i, \quad (2)$$

$$R_E^2 = |\vec{r}_N - \vec{r}_1|^2. \quad (3)$$

In the limit of infinite chain length, Ref. [1] found the universal ratio  $\lim_{N \rightarrow \infty} \langle R_E^2 \rangle / \langle R_G^2 \rangle \approx 6.254$ .

Besides  $R_G^2$  and  $R_E^2$ , the hydrodynamic radius is a third important measure of the coil dimension, which is measured in dynamic light scattering experiments [8]. The inverse

hydrodynamic radius is defined as

$$R_H^{-1} = \frac{1}{N^2} \sum_{i \neq j} \frac{1}{r_{ij}}, \quad (4)$$

$$r_{ij} = |\vec{r}_i - \vec{r}_j|, \quad (5)$$

with corresponding mean value  $\langle R_H^{-1} \rangle$ . This gives rise to another interesting amplitude ratio,  $\langle R_G^2 \rangle^{1/2} \langle R_H^{-1} \rangle$ , which is a universal constant in the limit of infinitely long chains that we denote as  $R_G/R_H$  by abuse of notation. In the present paper we will utilize the efficient algorithm of Refs. [1,2] to accurately calculate this universal quantity.

Only two recent high-resolution simulation studies have attempted to calculate the asymptotic ratio  $R_G/R_H$  with good accuracy: Dünweg *et al.* [9] find a value  $R_G/R_H = 1.591(7)$ , while Caracciolo *et al.* [10] quote  $R_G/R_H = 1.581(1)$ . These values are compatible with each other and also agree nicely with the prediction of renormalization-group calculations [11],  $R_G/R_H \approx 1.595$ . Mansfield and Douglas [12] have recently calculated the hydrodynamic radius in the infinite-chain length limit. However, while we calculate  $R_H^{-1}$  according to the definition in Eq. (4), they define a related quantity  $R_H^*$  (which is an expectation value) via the Stokes-Einstein relation

$$D = \frac{k_B T}{6\pi\eta R_H^*}, \quad (6)$$

where  $D$  is the translational diffusion coefficient of the molecule in infinitely diluted solution,  $k_B$  is the Boltzmann constant,  $T$  is the absolute temperature, and  $\eta$  is the solvent viscosity.  $R_H^{-1}$  according to Eq. (4) gives rise to the short-time (or Kirkwood) approximation to the diffusivity, while the true long-time value differs somewhat from the Kirkwood value [12–14]. Therefore their result is not directly comparable with ours. It will be shown that the present study has been able to obtain  $\langle R_H^{-1} \rangle$  according to Eq. (4) with substantially increased accuracy, and our estimate,  $R_G/R_H = 1.5803940(45)$ , is again in good agreement with Refs. [9,10].

A crucial aspect of the analysis of MC data is the observation that such simulations necessarily deal with

\*nclisby@unimelb.edu.au

finite chains of length  $N$ , while the abovementioned values for the universal numbers hold in the asymptotic limit  $N \rightarrow \infty$ . For this reason, a good understanding of the finite-chain-length effects (or corrections to scaling) is imperative for a correct and meaningful extrapolation. This is particularly true for the hydrodynamic radius since the corrections to scaling are very strong [9,10,12]. While for  $\langle R_G^2 \rangle$  and  $\langle R_E^2 \rangle$  the corrections are given by [1]

$$\langle R_G^2 \rangle = D_G N^{2\nu} (1 + a_G N^{-\Delta_1} + \dots), \quad (7)$$

$$\langle R_E^2 \rangle = D_E N^{2\nu} (1 + a_E N^{-\Delta_1} + \dots), \quad (8)$$

where the correction-to-scaling exponent  $\Delta_1 \approx 0.53$  [1], the hydrodynamic radius has an additional correction of order  $N^{-(1-\nu)}$ , with an exponent that is fairly close to  $\Delta_1$ , but which will ultimately be the *dominant* correction:

$$\langle R_H^{-1} \rangle = D_H N^{-\nu} (1 + a_H N^{-\Delta_1} + b_H N^{-(1-\nu)} + \dots); \quad (9)$$

here  $D_G$ ,  $D_E$ ,  $D_H$ ,  $a_G$ ,  $a_E$ ,  $a_H$ ,  $b_H$  are nonuniversal amplitudes. The origin of the  $N^{-(1-\nu)}$  term has been discussed in detail in Ref. [9]. These arguments shall not be repeated here; we rather refer the interested reader to that paper.

It turns out that the Monte Carlo sampling of  $R_H^{-1}$  with the algorithm of Refs. [1,2] is somewhat more tricky than one might expect at first glance. The reason for that problem is intricately related to the underlying recursive data structure, and it will be outlined in Sec. II. We have found a solution to the problem by inventing a sampling strategy, which will be elucidated in Sec. III. We then proceed in Sec. IV by outlining computational details of our study. In Sec. V we analyze our data and present a summary of results, including our estimate for  $R_G/R_H$  and a much improved estimate for  $\nu$  obtained by eliminating the leading correction to scaling. Our simulations are more accurate than those of Ref. [1] and hence allow us to also present improved estimates for the universal amplitude ratio  $\lim_{N \rightarrow \infty} \langle R_E^2 \rangle / \langle R_G^2 \rangle = D_E/D_G$  and  $\Delta_1$ . Finally, we conclude in Sec. VI.

## II. THE COMPUTATIONAL CHALLENGE

For our polymer simulations we utilize the pivot algorithm [4,5], which is the most powerful known method for sampling self-avoiding walks at fixed length. For SAWs on the simple cubic lattice with  $N$  monomers, the probability of a pivot move being successful decays as  $N^{-p}$  with  $p \approx 0.11$ . The standard hash table implementation [5] then requires mean processing time  $O(N)$  to generate an essentially new configuration with respect to global observables such as  $R_E^2$ . Recent algorithmic improvements [1,2,7] have further increased the relative advantage of the pivot algorithm over other methods. We utilize the SAW-tree data structure of Ref. [2] which allows us to perform pivot moves for an  $N$ -step SAW in mean processing time  $O(\log N)$ , resulting in mean processing time  $O(N^p \log N)$  to generate an essentially new configuration with respect to global observables.

The main ingredient of this implementation is a binary tree data structure that recursively decomposes a chain into subchains of decreasing length, until finally the monomer level is reached. Each node on the tree stores aggregate information

about its respective subchain, such as the coordinates of its center of mass, its end-to-end vector, its squared radius of gyration, and its bounding box (the smallest rectangular parallelepiped aligned with the lattice that completely encloses the subchain). Each geometric object within a bounding box is stored not in terms of absolute coordinates but rather in terms of coordinates *relative* to the origin and the orientation of the box. Now, a pivot move will always mean that a geometric transformation (combination of rotation, reflection, and translation) is applied to some monomers. Instead of moving all these monomers individually, the algorithm just moves those bounding boxes that need to be moved. Some bounding boxes will be big, some small, but the algorithm will always pick those boxes that are as big as possible. For example, in the simple case that the algorithm happens to just move the monomers number  $1, 2, \dots, N/2$ , only one single bounding box, corresponding to these monomers, is being transformed. Because of the storing of relative coordinates, all the data within such a box can be left as is and do not need to be updated. In other words, the algorithm always attempts to work at the highest possible levels of the tree and to avoid the data-intensive low levels as much as possible. Furthermore, since the coordinates of a box are known both from the outside and from the inside, this information makes it possible to recursively retrieve, starting from the top, the absolute coordinates of any geometric object if they are needed.

After a node has been updated, it needs to pass information to its higher-level node. For example, the end-to-end vector, the center of mass, and the gyration radius at the higher level will be changed, and so will be the bounding box. From there this passing will be done recursively all the way to the very top. However, information passing to lower levels is *not* needed, and this is what makes the method fast. It can thus be shown that the number of nodes that need to be updated is  $O(\log N)$ . The check for overlaps can also be done with average case  $O(\log N)$  computational complexity. The crucial observation is here that if two bounding boxes do not overlap, then this is also true for all monomers that they contain. Only in case of box overlap further investigation is needed, and this is again done in a top-down recursive fashion.

It is also clear that the evaluation of the end-to-end vector and of the center of mass are compatible with that approach. The end-to-end vector of a subchain that is decomposed into two sub-subchains is the sum of the end-to-end vectors of those sub-subchains, and therefore it is sufficient to pass information just to the higher-level node. Exactly the same statement holds for the center of mass, where instead of a sum we have an appropriately weighted average.

Although the method is slightly less obvious, the gyration radius may also be calculated in such a recursive fashion, as a few lines of straightforward algebra show that the following decomposition holds:

$$R_G^2 = \frac{N_1}{N} \{ R_{G1}^2 + |\vec{R}_{CM1} - \vec{R}_{CM}|^2 \} + \frac{N_2}{N} \{ R_{G2}^2 + |\vec{R}_{CM2} - \vec{R}_{CM}|^2 \}. \quad (10)$$

Here  $R_G^2$  is the squared gyration radius of the subchain with  $N$  monomers, while  $R_{G1}^2$  and  $R_{G2}^2$  are the corresponding squared

gyration radii of the two sub-subchains, with  $N_1$  and  $N_2$  monomers, respectively, while  $\vec{R}_{CM}$  is the center of mass of the subchain, and  $\vec{R}_{CM1}$ ,  $\vec{R}_{CM2}$  are the corresponding centers of mass of the sub-subchains. Thus, Eq. (10) allows us to calculate the gyration radius recursively as well.

However, the hydrodynamic radius is an observable that cannot be decomposed into subobservables of subchains. The reason is that  $R_H^{-1}$  involves interactions between distinct monomers and cannot be written in a form that involves only one-body terms (meaning that only sums of the form  $\sum_i \dots$  occur, but not terms of the form  $\sum_{ij} \dots$ ,  $\sum_{ijk} \dots$  and the like). In contrast,  $\vec{R}_E$  and also  $R_G^2$  can straightforwardly be written in such a form.

Therefore, calculating  $R_H^{-1}$  is in principle much harder than  $R_E$  or  $R_G$ , because a recursive evaluation cannot be done. The brute-force approach, in which one would evaluate the full double sum  $\sum_{i \neq j} r_{ij}^{-1}$  for each generated chain conformation, will obviously not work: The computational complexity of the sum, if done exactly, scales as  $O(N^2)$  (perhaps with an additional factor of  $O(\log N)$  depending on the details of the implementation). This could be improved to  $O(N)$  if the hydrodynamic radius were evaluated via the fast multipole method [15]. If we were using the hash table implementation of the pivot algorithm then this would indeed be a very effective approach, as the mean processing time to generate a new SAW would also be  $O(N)$ . However, both the naive and fast multipole methods would dominate the mean processing time required to generate a new SAW for the SAW-tree implementation of  $O(N^p \log N)$ . In other words, evaluation of the full sum for the hydrodynamic radius would lead to an algorithm for which nearly all advantages of the SAW-tree implementation would be lost.

Our simple solution, whose computational complexity is logarithmic in  $N$ , shall be outlined in the next section. From the structure of the method as explained below, it is clear that it can be applied to any observable that has the form  $\sum_i A_1(\vec{r}_i)$ ,  $\sum_{ij} A_2(\vec{r}_i, \vec{r}_j)$ ,  $\sum_{ijk} A_3(\vec{r}_i, \vec{r}_j, \vec{r}_k)$ , and so on, as well as combinations of these, and is thus quite general. However, it may fail if one is interested in complex observables such as knot types.

### III. SAMPLING STRATEGY FOR CALCULATION OF THE HYDRODYNAMIC RADIUS

The key to our approach to solve the abovementioned problem is the following simple observation: We write

$$\begin{aligned} R_H^{-1} &= \frac{1}{N^2} \sum_{i \neq j} \frac{1}{r_{ij}} = \left(1 - \frac{1}{N}\right) \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{r_{ij}} \\ &= \left(1 - \frac{1}{N}\right) \left[ \frac{1}{r} \right], \end{aligned} \quad (11)$$

where  $[\dots]$  denotes an average over all pairs. This means that, for a given conformation of the chain, we can find the observable  $R_H^{-1}$  not only by brute-force calculation of the sum but also by Monte Carlo sampling: We simply pick a pair of monomers  $(i, j)$  uniformly at random from the set of all monomer pairs and evaluate  $r_{ij}^{-1}$ . If we do this often and average over the results, this will stochastically converge

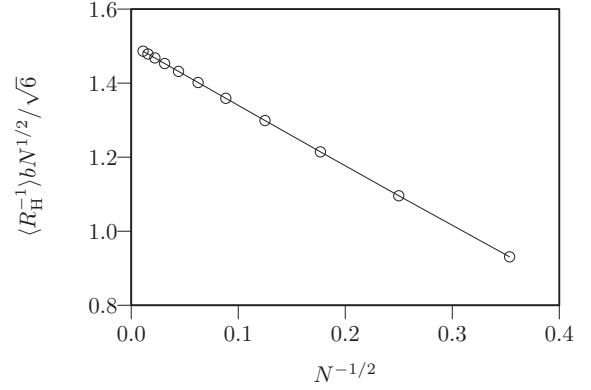


FIG. 1. Exact results for  $\langle R_H^{-1} \rangle$  for a Gaussian chain. Instead of  $\langle R_H^{-1} \rangle$  we rather plot the dimensionless ratio  $\langle R_H^{-1} \rangle b N^{1/2} / \sqrt{6}$ , where  $b N^{1/2} / \sqrt{6}$  is the asymptotic long-chain value for the gyration radius  $\langle R_G^2 \rangle^{1/2}$ . In other words, corrections to scaling are taken into account only for the hydrodynamic radius but not for the gyration radius. The argument on the abscissa,  $N^{-1/2}$ , reflects the leading-order correction to scaling. Note also that the asymptotic value for  $N \rightarrow \infty$  is  $R_G/R_H = 8/(3\sqrt{\pi}) \approx 1.5045$ .

toward  $R_H^{-1}/(1 - N^{-1})$ . Actually, it is sufficient to do this only *once* per generated chain conformation, since the average over pairs will be automatically included in the overall sampling. We thus write

$$\langle R_H^{-1} \rangle = \left(1 - \frac{1}{N}\right) \left[ \left[ \frac{1}{r} \right] \right], \quad (12)$$

where the average  $\langle \dots \rangle$  means the average over chain conformations, and  $[\dots]$  is the average over monomer pairs; these averaging operations are interchangeable.

This strategy gives rise to  $O(\log N)$  computational complexity for the operations being done for one chain conformation, since finding the actual coordinates of monomers  $i$  and  $j$  involves a recursive search along the binary tree. In other words, the computational complexity of the observable evaluation is comparable to the computational complexity to perform a single update by attempting to perform a pivot move.

In order to test this idea, we first studied a Gaussian chain in three-dimensional continuous space, with  $\langle r_{ij}^2 \rangle = b^2 |i - j|$ , as a simple toy model. For this model one finds analytically by a Gaussian integral  $\langle r_{ij}^{-1} \rangle = \sqrt{6/\pi} b^{-1} |i - j|^{-1/2}$ , and the remaining double sum is easily numerically evaluated to yield an exact value for  $\langle R_H^{-1} \rangle$  for any reasonable chain length (including all corrections to scaling). The result is shown in Fig. 1.

It is also very easy to stochastically generate such a chain using Gaussian random numbers, based upon the Box-Muller transformation. We therefore studied chains of length  $8 \leq N \leq 8192$  and sampled  $\langle R_H^{-1} \rangle$  from  $10^6$  stochastic realizations. We first calculated  $R_H^{-1}$  in the conventional way by brute-force evaluation of the double sum. Using the same kind of plot as in Fig. 1, the results are indistinguishable from the exact values. We hence rather show the deviation from the exact result, using the same normalization as in Fig. 1 (i.e., we study  $\langle R_H^{-1} \rangle$  normalized by the asymptotic gyration radius of a chain with the same  $N$ ). The result is shown in Fig. 2. As

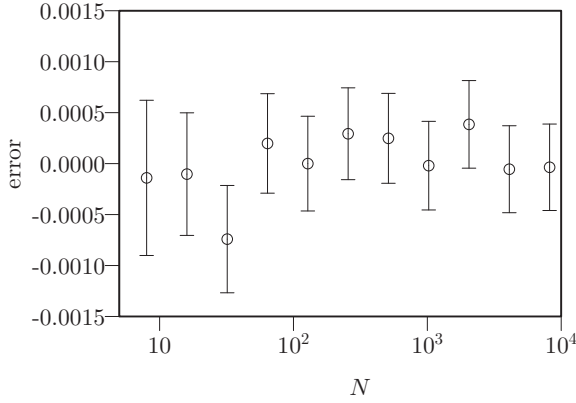


FIG. 2. Difference between  $\langle R_H^{-1} \rangle bN^{1/2}/\sqrt{6}$  (sampled value) and  $\langle R_H^{-1} \rangle bN^{1/2}/\sqrt{6}$  (exact value), as a function of chain length  $N$ . Here the sampled value for  $\langle R_H^{-1} \rangle$  results from averaging over  $10^6$  independent chains, using a full evaluation of the double sum  $\sum_{i \neq j} r_{ij}^{-1}$ . The error bars have been estimated as three times the standard error of mean.

it should be, the sampled results are well compatible with the exact values within error bars.

Using the same chains, we then sampled  $\langle R_H^{-1} \rangle$  by the “one pair of monomers per chain” sampling strategy as outlined above. As seen in Fig. 3, again the results are nicely compatible with the exact values within error bars. The important point to notice is that the latter are only roughly a factor of 10 larger than in the case of full evaluation, and this ratio varies only very weakly (possibly logarithmically) with chain length, as shown in Fig. 4. This, however, means quite clearly that the immense computational effort to evaluate the double sum does not pay off in terms of a substantially increased statistical accuracy and that rather the “one pair of monomers per chain” method is a much more efficient overall sampling strategy. One may think of a variant of this scheme, where one rather picks pairs  $(i, j)$  not uniformly, but rather with a probability  $\propto |i - j|^{-\alpha}$  for some  $\alpha$ . However, we expect such a change to only slightly improve the statistical accuracy, compared to the tremendous gain obtained by discarding the double sum. We hence did not try such a refinement and kept using simple uniform sampling.

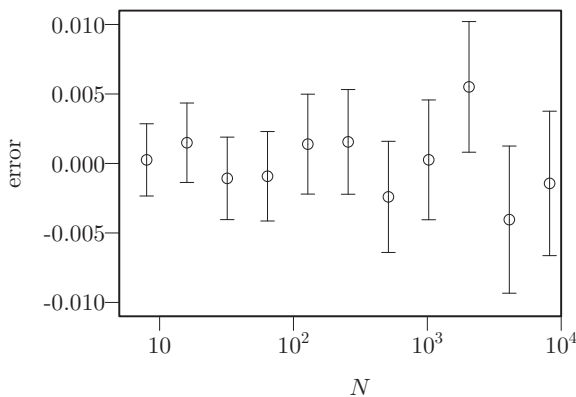


FIG. 3. Same as Fig. 2, but now applying the refined sampling strategy where the interparticle distance is only evaluated for one randomly selected pair of monomers.

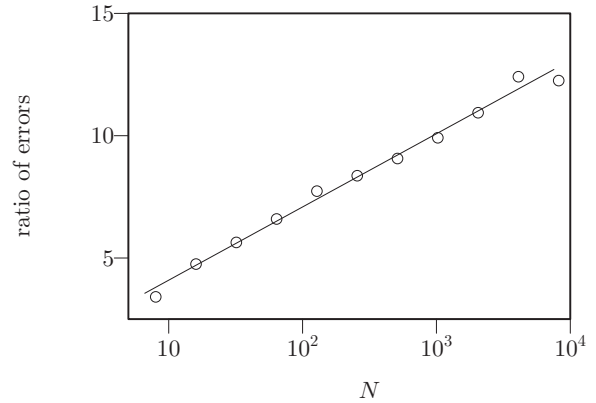


FIG. 4. Ratio of the error bars from Figs. 3 and 2, as a function of chain length  $N$ . As the statistical uncertainty of these data was not sampled, we do not show statistical error bars. The straight line is the function  $1.1 + 1.3 \log N$ .

At this point, we wish to remark that it may also be useful to pick more than just one pair of monomers per chain. This, of course, helps to improve the statistical accuracy somewhat. More importantly, however, this is needed if one is interested not only in the average value but also in higher moments of the distribution or in time correlation functions that characterize the efficiency of the algorithm. Let us discuss this in more detail for the variance of the inverse hydrodynamic radius. Obviously, we have

$$\begin{aligned} \text{var}(R_H^{-1}) &= \langle R_H^{-2} \rangle - \langle R_H^{-1} \rangle^2 \\ &= \frac{1}{N^4} \sum_{i \neq j} \sum_{k \neq l} \left\{ \left\langle \frac{1}{r_{ij}} \frac{1}{r_{kl}} \right\rangle - \left\langle \frac{1}{r_{ij}} \right\rangle \left\langle \frac{1}{r_{kl}} \right\rangle \right\} \\ &= \left( 1 - \frac{1}{N} \right)^2 \left\{ \left\langle \left[ \frac{1}{r} \right]^2 \right\rangle - \left\langle \left[ \frac{1}{r} \right] \right\rangle^2 \right\}, \end{aligned} \quad (13)$$

where the last step is performed by using the same trick as in Eq. (11) to convert the sum over monomers to an average. To sample this by a one-pair-per-chain strategy is impossible, however, since the form  $\langle [1/r]^2 \rangle$  no longer permits us to just exchange the averages  $\langle \dots \rangle$  and  $[\dots]$ . Rather, we have

$$\left[ \frac{1}{r} \right]^2 = \left[ \left[ \frac{1}{r} \frac{1}{r'} \right] \right], \quad (14)$$

where  $[[\dots]]$  is now an average involving *four* monomers  $i, j, k, l$  with  $i \neq j$  and  $k \neq l$ . To obtain this average, one needs to randomly pick such four monomers and calculate  $r_{ij}^{-1} r_{kl}^{-1}$ . This latter average is again interchangeable with  $\langle \dots \rangle$  and hence is in accord with our general strategy. Similar considerations apply for even higher moments or time correlation functions. These considerations have motivated us to run the simulation by not sampling one but rather two monomer pairs per chain.

In practice, for the main computer experiment of self-avoiding walks, the observable we sample is

$$Q = \frac{1}{2} \left( 1 - \frac{1}{N} \right) \left( \frac{1}{r} + \frac{1}{r'} \right), \quad (15)$$

which satisfies  $\langle [[Q]] \rangle = \langle R_H^{-1} \rangle$ .

#### IV. DETAILS OF COMPUTER EXPERIMENT

We now briefly describe the details of the computer experiment, which involved the pivot algorithm sampling of self-avoiding walks for which the number of monomers  $N$  varied from 512 to 33 554 432 ( $2^{25}$ ).

The pivot algorithm is ergodic and satisfies the detailed balance condition [5] and so samples self-avoiding walks uniformly at random. However, to avoid initialization bias, it is necessary to run the pivot algorithm until the Markov chain is indistinguishably close to equilibrium. In each case, the seed self-avoiding walk was generated using the pseudodimerize algorithm described in Ref. [2]; the system was then equilibrated by performing approximately  $20N$  successful pivots (no data were collected during the initialization stage).

Now that an appropriate initial SAW configuration had been generated, the computer experiment to collect data was begun. At each time step, various observables were sampled: The exact values for the squared end-to-end distance and the radius of gyration were used, while the inverse hydrodynamic radius and the square of the inverse hydrodynamic radius were estimated using an unbiased estimator, as described in Sec. III.

The computer experiment was run for 195 000 hours on Dell PowerEdge FC630 machines with Intel Xeon E5-2680 processing units (these were run in hyperthreaded mode, which gave a modest performance boost; 390 000 thread hours were used). In total there were  $1.70 \times 10^6$  batches of  $10^8$  attempted pivots, and thus there were a grand total of  $1.70 \times 10^{14}$  attempted pivots across all walk sizes.

We confirmed that the batching method of error estimation was reliably converging even for the largest values of  $N$ . This indicates that the degree of correlation between consecutive batches of  $10^8$  pivot attempts was minimal for each of our global observables  $R_E^2$ ,  $R_G^2$ , and  $R_H^{-1}$ , even for the largest size where  $N = 2^{25}$ .

The raw data that have been produced in this way are given in the tables of Appendix B. We include estimates of the amplitude ratios  $\langle R_E^2 \rangle / \langle R_G^2 \rangle$  and  $\langle R_G^2 \rangle^{1/2} \langle R_H^{-1} \rangle$  as they have smaller confidence intervals than might naively be expected from the estimates of  $\langle R_E^2 \rangle$ ,  $\langle R_G^2 \rangle$ , and  $\langle R_H^{-1} \rangle$  due to correlations between the observables  $R_E^2$ ,  $R_G^2$ , and  $R_H^{-1}$ , which reduce the variance of the ratio estimates.

We now briefly consider the properties of our Markov chain sampling method, with a view to gauging the relative effectiveness of our method for  $R_H^{-1}$  versus the observable  $R_E^2$ , which can be calculated exactly in an efficient manner.

Given an observable  $A$  with variance  $\text{var}(A) = \langle A^2 \rangle - \langle A \rangle^2$ , we follow Ref. [6] and define the autocorrelation function for this observable as

$$\rho_{AA}(t) = \frac{\langle A_s A_{s+t} \rangle - \langle A \rangle^2}{\text{var}(A)}. \quad (16)$$

The key quantity which measures the efficiency with which  $A$  is sampled is the integrated autocorrelation time  $\tau_{\text{int}}$ , defined as

$$\tau_{\text{int}}(A) = \frac{1}{2} + \sum_{t=1}^{\infty} \rho_{AA}(t). \quad (17)$$

$\tau_{\text{int}}$  may be thought of as the number of Markov chain steps required before the state is effectively new with respect to

the observable  $A$ . For a sampling scheme where consecutive estimates are completely uncorrelated, we would have  $\tau_{\text{int}}(A) = 1/2$ . While  $\tau_{\text{int}}$  may well be different for different observables, for the pivot algorithm we expect that global observables such as  $R_G^2$ ,  $R_E^2$ , and  $R_H^{-1}$  should decorrelate after a constant number of successful pivots.

We can then calculate an *a priori* estimate of the expected error on our estimate of the sample mean  $\bar{A}$  for  $n_{\text{sample}}$  Markov chain time steps:

$$\text{stdev}(\bar{A}) = \left( \frac{2\tau_{\text{int}}(A)\text{var}(A)}{n_{\text{sample}}} \right)^{\frac{1}{2}}. \quad (18)$$

Our goal in performing our Monte Carlo simulation is to estimate  $\langle A \rangle$  as accurately as possible for a given amount of computer time. Usually, this entails either finding an observable  $A'$  for which  $\langle A' \rangle = \langle A \rangle$  but  $\text{var}(A') < \text{var}(A)$ , thus allowing for more efficient sampling (variance reduction), or finding a Markov chain with an improved move set which reduces  $\tau_{\text{int}}(A)$ , or improving the efficiency of the computer implementation which allows  $n_{\text{sample}}$  to be increased for the same computational effort.

Our situation is a unique mix of these: We instead estimate an observable  $Q$  from Eq. (15) which can be much more efficiently evaluated, thus increasing  $n_{\text{sample}}$ , but at the expense of increasing the variance. The key question is what is the performance penalty from doing this, relative to an efficient exact method?

We examine this question by calculating the ratio of relative errors in the estimates of  $\langle R_H^{-1} \rangle$  and  $\langle R_E^2 \rangle$ , which we plot in Fig. 5. There we see that the relative error for  $\langle R_H^{-1} \rangle$  is substantially below that for  $\langle R_E^2 \rangle$ , although the ratio is growing with  $N$ , perhaps logarithmically. This behavior is qualitatively the same as the situation for a Gaussian chain as shown in Fig. 4. In fact, we expect that the relative performance penalty should be somewhat less than that case, because pivot moves are only successful on average once every  $O(N^p)$  attempts ( $p \approx 0.11$  for the simple cubic lattice), and so  $Q$  is sampled on  $O(N^p)$  occasions over a time period for which  $R_H^{-1}$  remains frozen.

Thus it seems that the performance penalty is quite modest. Whether there exist alternatives to the observable  $Q$  which

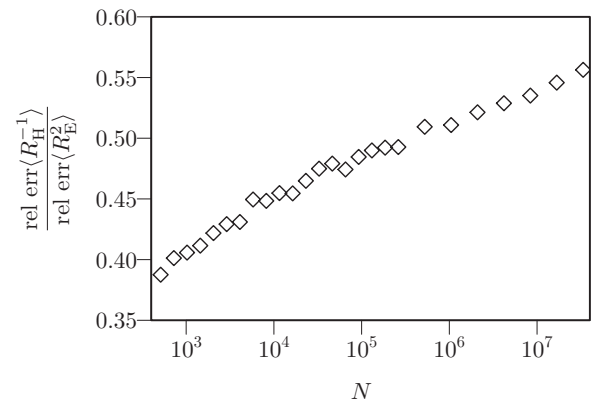


FIG. 5. Plot of the ratio of relative errors for  $\langle R_H^{-1} \rangle$  and  $\langle R_E^2 \rangle$ .

could significantly improve sampling performance is an open research question.

**V. ANALYSIS AND RESULTS**

In this section we describe the analysis of data collected in the tables of Appendix B. We initially fit the data for standard observables with a model derived from their expected asymptotic behavior; this is the conventional method. We then describe a method which has been used previously for the Ising model [16,17], which eliminates the leading-order correction to scaling term and allows for a much improved estimate for  $\nu$ . Next we analyze our data for the hydrodynamic radius and present a summary of our results together with estimates from the literature in Table I.

We first study the data for  $\langle R_G^2 \rangle$ . Starting from Eq. (7), we apply four-parameter fits to the data, where  $D_G$ ,  $a_G$ ,  $\nu$ , and  $\Delta_1$  are considered as fit parameters, while the higher-order corrections to scaling are neglected. Because of the large range of chain lengths and the high resolution accessible to our simulation, these higher-order corrections cause systematic errors in the fits at a comparable level to the statistical error. For this reason, we do the fits for various ranges of chain lengths ( $N \geq N_{\min}$ , where  $N_{\min}$  is varied systematically). The effect of the higher-order corrections is then a systematic dependence of the fit parameters on  $N_{\min}$ . In fact, the deviations for  $D_G$  and  $\nu$  are expected to scale as  $N_{\min}^{-y}$ , where  $y$  is the correction-to-scaling exponent corresponding to the first neglected term (for a derivation, see Appendix A). In Eq. (7) it is believed that there are in fact three competing next-to-leading correction terms with exponents 1 (analytic),  $2\Delta_1 \approx 1.06$ , and  $\Delta_2 \approx 1$

TABLE I. Summary of estimates of  $\nu$ ,  $\Delta_1$ , and  $R_G/R_H$ . In addition we have  $D_E/D_G = 6.253531(10)$  (cf.  $6.2537(18)$  [1]),  $D_G = 0.1951413(26)$  (cf.  $0.19514(4)$  [1]), and  $D_E = 1.220322(18)$  (cf.  $1.22035(25)$  [1]). Note that results in the table are listed in reverse chronological order, i.e., the most recently published work is at the top.

Source <sup>a</sup>	$\nu$	$\Delta_1$	$R_G/R_H$
Present work	0.58759700(40)	0.528(8)	1.5803940(45)
[19] CB	0.58775(83)		
[21] Series	0.58772(17)		
[1] MC	0.587597(7)	0.528(12)	
[22] <sup>b</sup> Series	0.58774(22)		
[10] MC			1.581(1)
[9] MC			1.591(7)
[23] MC	0.5874(2)		
[24] <sup>c</sup> Series	0.58755(55)		
[25] FT $d = 3$	0.5882(11)	0.478(10)	
[25] FT $\epsilon$ bc	0.5878(11)	0.486(16)	
[26] MCRG	0.58756(5)	0.5310(33)	
[6] MC	0.5877(6)	0.56(3)	
[11] FT			$\approx 1.595$

<sup>a</sup>Abbreviations: MC, Monte Carlo; CB, conformal bootstrap; FT, field theory; MCRG, Monte Carlo renormalization group.

<sup>b</sup>Using Eqs. (74) and (75) of Ref. [22] with  $0.516 \leq \Delta_1 \leq 0.54$ .

<sup>c</sup>No error estimates were made in Ref. [24], but estimates for  $\nu$  were in the range  $0.5870 \leq \nu \leq 0.5881$ .

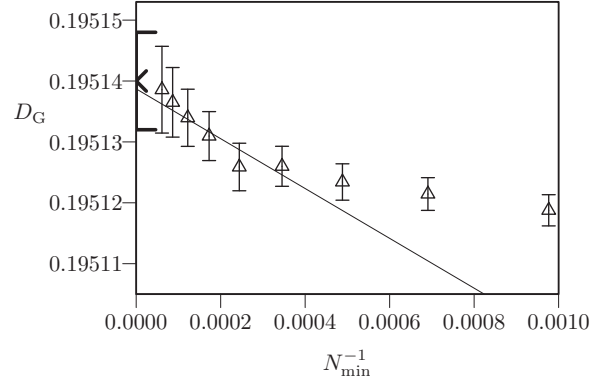


FIG. 6. Systematic variation of the fitted amplitude of  $\langle R_G^2 \rangle$  with  $N_{\min}$ . The line of best fit to the final six values is shown, and we plot our best estimate from these data of  $D_G = 0.1951400(80)$ .

( $\Delta_2$  is not known with any precision). Assuming a value  $y \approx 1$  we thus plot the estimates for  $D_G$  and  $\nu$  as a function of  $N_{\min}^{-1}$ . For  $\langle R_E^2 \rangle$  we can apply the same analysis to Eq. (8).

We perform one further trick to reduce the influence of unfitted correction to scaling terms and make extrapolation easier. We multiply our raw data by  $1 - c/N$ , where  $c$  is an arbitrary constant chosen to reduce the curvature observed in fits. Note that this trick does not change the leading or next-to-leading asymptotic behavior of the observables, and so if extrapolation is performed carefully this will not affect our final estimates. We found that a good choice for  $\langle R_G^2 \rangle$  was  $c = 0.0$ , for  $\langle R_E^2 \rangle$  we had  $c = 0.6$ , for  $\langle R_E^2 \rangle / \langle R_G^2 \rangle$  we had  $c = 0.2$ , for  $\langle R_H^{-1} \rangle$  we had  $c = -0.2$ , and for  $\langle R_G^2 \rangle^{1/2} \langle R_H^{-1} \rangle$  we had  $c = -0.5$ .

We plot the resulting estimates in Figs. 6, 7, and 8. Note that all error bars shown are statistical and arise from the fitting procedure. To take into account the systematic error from corrections to scaling we extrapolate to the left-hand side of the plots where  $N_{\min} \rightarrow \infty$ . We choose our final extrapolated value for the parameters by performing linear fits of subsequent estimates, with an error bar which is sufficiently large so as to account for both the observed statistical error and unobserved systematic error which manifests itself in the plots as nonlinear convergence. In the case of Fig. 8 we have the

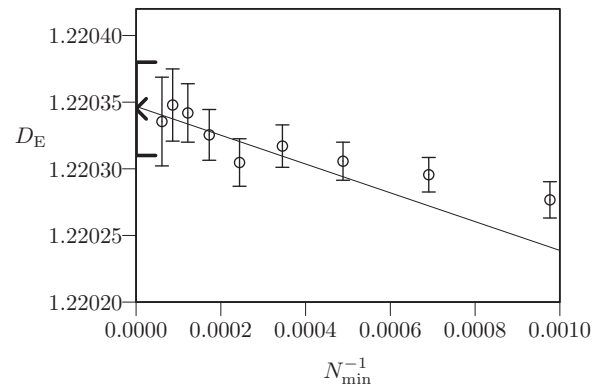


FIG. 7. Systematic variation of the fitted amplitude of  $\langle R_E^2 \rangle$  with  $N_{\min}$ . The line of best fit to the final six values is shown, and we plot our best estimate from these data of  $D_E = 1.220345(35)$ .

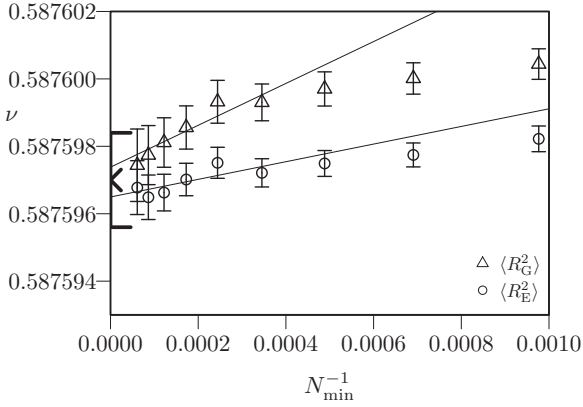


FIG. 8. Systematic variation of the fitted value of  $\nu$  with  $N_{\min}$ , using both  $\langle R_G^2 \rangle$  and  $\langle R_E^2 \rangle$  data. The line of best fit to the final six values is shown, and we show our best estimate from these fits of  $\nu = 0.5875970(14)$ .

benefit of two observables giving estimates for  $\nu$  which have different unfitted corrections, which increases the reliability of the extrapolation procedure.

The fit in Fig. 8 gives  $\nu = 0.5875970(14)$ , which improves significantly on the literature, but we can do better, as we show later in this section. Note that throughout this work we usually report two significant figures for our confidence intervals. This is not because we claim that these confidence intervals are so precise, but because information is lost when only one significant figure is used. For example, confidence intervals of  $35 \times 10^{-8}$  and  $44 \times 10^{-8}$  would both be reported as a confidence interval of  $4 \times 10^{-7}$  if only one significant figure were used.

Similarly, we can also study the ratio  $\langle R_E^2 \rangle / \langle R_G^2 \rangle$ , which converges toward the universal amplitude ratio  $D_E / D_G$ . Taking the ratio reduces the fits from four to three parameters, as the powers of  $N^{2\nu}$  cancel out, and for this reason the estimate  $D_E / D_G$  is more accurate than for the individual amplitudes  $D_E$  and  $D_G$ . The estimated values should again vary systematically like  $N_{\min}^{-1}$ , and the corresponding plot is Fig. 9. The universal ratio is therefore found to take the asymptotic value  $D_E / D_G = 6.253531(10)$ .

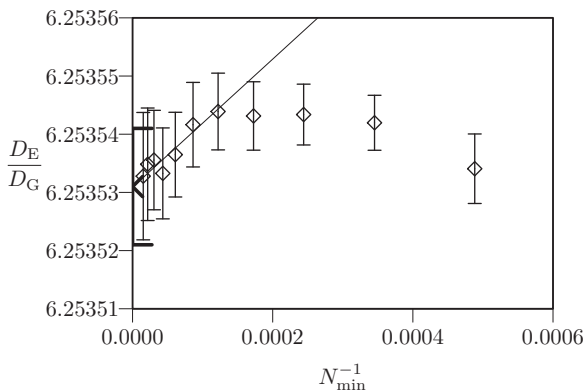


FIG. 9. Systematic variation of our estimates of  $D_E / D_G$  with  $N_{\min}$ . The line of best fit to the final six values is shown, and we show our best estimate from these fits of  $D_E / D_G = 6.253531(10)$ .

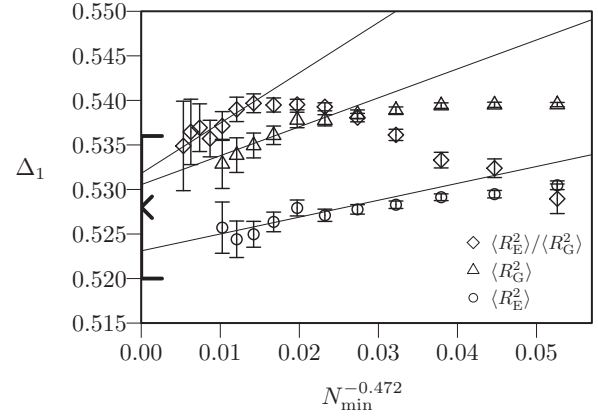


FIG. 10. Systematic variation of the fitted value of  $\Delta_1$  with  $N_{\min}$ , using data for  $\langle R_G^2 \rangle$ ,  $\langle R_E^2 \rangle$  and their ratio. The line of best fit to the final six values is shown, and we show our best estimate from these fits of  $\Delta_1 = 0.528(8)$ .

Finally, we can also use these data to determine  $\Delta_1$ , whose value is found to be  $\Delta_1 = 0.528(8)$ . Again taking the next to leading correction exponent as  $-1$ , the fitted value should vary with  $N_{\min}$  like  $N_{\min}^{\Delta_1 - 1} \approx N_{\min}^{-0.472}$ . The results are shown in Fig. 10.

We now describe a method of analysis which allows us to eliminate the leading correction to scaling and obtain a much improved estimate for  $\nu$ .

It is a standard technique to use improved models for simulations in statistical mechanics, where typically a parameter is chosen so that the leading correction to scaling term for all observables is reduced sufficiently so that their contributions are below the level of statistical error. For models in the self-avoiding walk universality class, two such improved models are the Domb-Joyce model [10] and the bead model [18].

The basic idea of the method is very simple: Instead of attempting to find an improved model, we find an improved observable instead. This technique was previously used for the three-dimensional dilute Ising model [16] and models in the universality class of the three-dimensional Ising model [17]. Since  $\langle R_G^2 \rangle$  and  $\langle R_E^2 \rangle$  are independent measures of the size of a polymer, the relative size of the leading correction to scaling term for each of these observables is different. By forming an improved observable  $R_{\text{imp}}^2$  via the linear combination

$$R_{\text{imp}}^2 = R_E^2 - 4.478 R_G^2, \quad (19)$$

we find that we are able to reduce the amplitude of the leading correction to scaling to a level below the statistical noise.

We are then able to fit  $\langle R_{\text{imp}}^2 \rangle$  by the truncated model

$$\langle R_{\text{imp}}^2 \rangle = D_{\text{imp}} N^{2\nu} \left[ 1 + \frac{\varepsilon}{N^{\Delta_1}} + O\left(\frac{1}{N}\right) \right], \quad (20)$$

where we only fit  $D_{\text{imp}}$  and  $\nu$ , neglecting the  $O(\varepsilon)$  term. We confirm that this is indeed an excellent model for the data for  $N_{\min} \geq 8192$  as the reduced  $\chi^2$  of the fits is approximately 1. By reducing the order of the fits from four parameters to two, we obtain sensible fits even for  $N_{\min}$  up to 262 144, which are

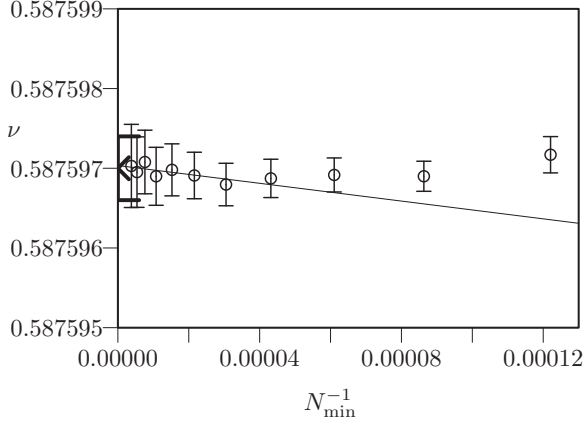


FIG. 11. Systematic variation of the fitted value of  $\nu$  with  $N_{\min}$ , using data for the improved combination  $\langle R_{\text{imp}}^2 \rangle = \langle R_E^2 \rangle - 4.478 \langle R_G^2 \rangle$ . The line of best fit to the final six values is shown, and we show our best estimate from these fits of  $\nu = 0.58759700(40)$ .

far more accurate than the estimates from fits of  $\langle R_E^2 \rangle$  and  $\langle R_G^2 \rangle$ . We plot the resulting estimates for  $\nu$  against  $N_{\min}^{-1}$  in Fig. 11, where it can be seen that convergence in the limit  $N_{\min} \rightarrow \infty$  is smooth.

Note that in this case we did not use the additional trick of multiplying by  $1 - c/N$ . We have also checked the stability of the method by varying the constant in Eq. (19), and find that within the interval (4.473, 4.483) the plot in Fig. 11 is quite linear and can be extrapolated easily.

Note the substantial decrease in range and domain for the plots from the standard approach in Fig. 8 as compared to the new approach in Fig. 11. Purely from this method of analysis we have managed to decrease the error by more than a factor of three, from  $14 \times 10^{-7}$  to  $4 \times 10^{-7}$ . Our central estimate was not changed, and our final estimate is  $\nu = 0.58759700(40)$ .

We now perform one final trick to obtain improved estimates for  $D_E$  and  $D_G$ . We first plot the estimates for  $D_{\text{imp}}$  obtained from our two-parameter fits in Fig. 12. We then use the fact that our estimate of  $D_E/D_G$  is more accurate

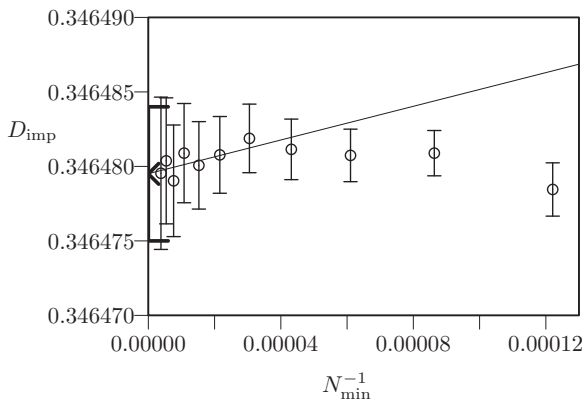


FIG. 12. Systematic variation of the fitted value of  $D_{\text{imp}} = D_E - 4.478 D_G$  with  $N_{\min}$ , using data for the improved combination  $\langle R_{\text{imp}}^2 \rangle = \langle R_E^2 \rangle - 4.478 \langle R_G^2 \rangle$ . The line of best fit to the final six values is shown, and we show our best estimate from this fit of  $D_{\text{imp}} = 0.3464795(45)$ .

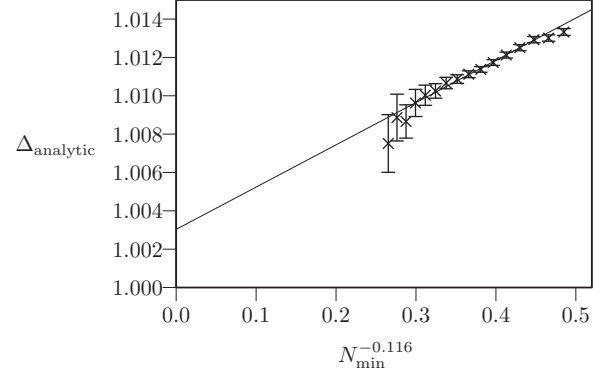


FIG. 13. Systematic variation of the fitted value of  $\Delta_a$  with  $N_{\min}$ , using data for  $\langle R_H^{-1} \rangle$ .

than the estimates of  $D_E$  and  $D_G$  individually, and form the combinations

$$D_E = \frac{D_{\text{imp}}}{1 - 4.478 D_G/D_E}, \quad (21)$$

$$D_G = \frac{D_{\text{imp}}}{D_E/D_G - 4.478}. \quad (22)$$

We combine the errors from  $D_{\text{imp}}$  and  $D_E/D_G$  as if they were independent and obtain the improved estimates  $D_G = 0.1951413(26)$  and  $D_E = 1.220322(18)$ .

We now turn to the  $R_H^{-1}$  data, where Eq. (9) applies. Again, we start with a four-parameter fit, where we take the leading order into account, plus the dominant correction to scaling. The latter should be the analytic term, which is absent for  $\langle R_G^2 \rangle$  and  $\langle R_E^2 \rangle$ . If only those two terms are present, the fit function can be written as

$$\langle R_H^{-1} \rangle = D_H N^{-\nu} + E_H N^{-\Delta_a}, \quad (23)$$

where the analytic value  $\Delta_a$  is one. This contribution is difficult to distinguish from the next-order contribution, which scales as  $N^{-(\nu+\Delta_1)} \approx N^{-1.116}$ , where the exponent is only slightly different. However, our data are accurate enough that this is actually possible. We therefore apply a four-parameter fit to the data according to Eq. (23), where  $\Delta_a$  is left as a fit parameter. Using the results of Appendix A, these data should then vary with  $N_{\min}$  according to  $\Delta_a \propto N_{\min}^{\Delta_a - \nu - \Delta_1} = N_{\min}^{-0.116}$ . As seen in Fig. 13, they nicely extrapolate to  $\Delta_a \approx 1$ , with a value that is clearly distinguishable from the next order (1.116).

Finally, we focus on the universal amplitude ratio  $R_G/R_H$ , which was the original motivation to perform the present study. This can be written as

$$\langle R_G^2 \rangle^{1/2} \langle R_H^{-1} \rangle = \frac{R_G}{R_H} + B N^{-(1-\nu)} + C N^{-\Delta_1} + \dots, \quad (24)$$

where the omitted leading-order correction is  $O(N^{-1})$ . We now use the value for  $\nu$  as obtained from the  $\langle R_G^2 \rangle$  and  $\langle R_E^2 \rangle$  data, and treat the parameters  $R_G/R_H$ ,  $B$ ,  $C$ , and  $\Delta_1$  in Eq. (24) as fit parameters in a four-parameter fit. The parameter  $R_G/R_H$  should then vary linearly with  $N_{\min}^{-1}$ . The data are shown in Fig. 14 and give rise to an estimate for the universal amplitude ratio of  $R_G/R_H = 1.5803940(45)$ .



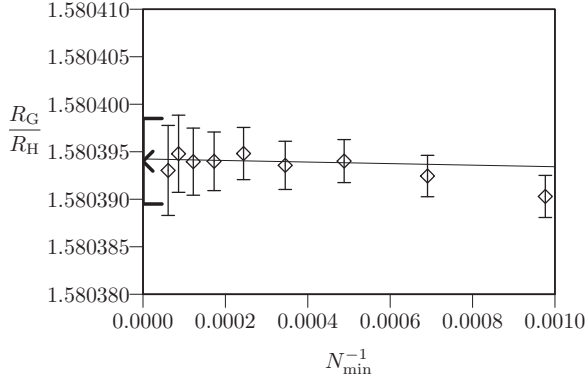


FIG. 14. Systematic variation of the fitted value of  $R_G/R_H$  with  $N_{\min}$ . The line of best fit to the final six values is shown, and we plot our best estimate from these data of  $R_G/R_H = 1.5803940(45)$ .

Table I summarizes our results, with a comparison with previous results from Monte Carlo, series expansion, field theoretic, and conformal bootstrap methods. We wish to highlight the recent conformal bootstrap estimate of  $\nu = 0.58775(83)$  [19] as this approach shows a great deal of promise. The method has been spectacularly successful for the three-dimensional Ising model giving  $\nu = 0.6299748(40)$  [20]; in this case it is far superior to Monte Carlo methods.

## VI. CONCLUSION

The combination of the pivot algorithm and the SAW-tree data structure of Refs. [1,2] provides an extremely efficient method to obtain the properties of long SAWs with high accuracy. The SAW-tree allows for the efficient computation of observables such as  $R_G^2$  and  $R_E^2$ , but not for other observables such as  $R_H^{-1}$ , which leads to a unique problem: How to efficiently sample an observable whose calculation would dominate the runtime of the Markov chain sampling algorithm? The key insight is that the observable does not need to be calculated exactly in order to obtain accurate estimates; instead we only need to find an unbiased estimator of the observable which can be calculated efficiently and which has moderate variance.

Starting from the observation that a large class of observables can be written as the sum of  $n$ -body terms involving  $n$  monomers, where this series typically stops at low (and in most cases at second) order, we propose a double sampling scheme, where not only the chain conformations are generated at random, but also the monomers that contribute to the  $n$ -body interactions are picked at random, such that this evaluation involving just a few monomers replaces an exhaustive sum over all sets of  $n$  monomers. This leads to an efficient Monte Carlo sampling for many observables, and the present work demonstrates its usefulness by applying it to the problem of sampling the hydrodynamic radius of three-dimensional SAWs. Using this technique we estimated with high accuracy the universal amplitude ratio  $R_G/R_H = 1.5803940(45)$  and discerned the competing corrections to scaling for  $(R_H^{-1})$ . Finally, we have constructed an improved observable for which the leading correction to scaling has negligible amplitude and

used it to obtain an improved estimate for the Flory exponent of  $\nu = 0.58759700(40)$ .

## ACKNOWLEDGMENTS

N.C. acknowledges funding by the Australian Research Council under the Future Fellowship scheme (Project No. FT130100972) and Discovery scheme (Project No. DP140101110). B.D. acknowledges hospitality of the University of Melbourne during two visits where some of this work was conducted.

## APPENDIX A: FITTING STRATEGY

Here we will describe some of the details of the fitting procedure used in the main text. The description will be quite general, but we will refer to specific examples from the analysis section.

Suppose we are interested in a certain observable, e.g.,  $\langle R_G^2 \rangle$ , as a function of chain length  $N$ . Let us denote this observable as  $R(N)$ . We expect that for  $R(N)$  there exists an infinite asymptotic expansion in  $N$ :

$$R(N) = \sum_{i=1}^{\infty} \alpha_i N^{x_i} \quad (\text{A1})$$

$$= \sum_{i=1}^M \alpha_i N^{x_i} + O(N^{x_{M+1}}), \quad (\text{A2})$$

where  $x_1 > x_2 > \dots > x_M > x_{M+1} > \dots$ , such that  $x_1$  describes the leading asymptotic power-law dependence, while the exponents  $x_2, x_3, \dots$  correspond to the corrections to scaling. The parameters  $\alpha_i$  are the corresponding amplitudes. Note that the procedure described here can be straightforwardly adapted to observables with different asymptotic behavior, e.g., exponential growth with power law corrections.

Now we perform a computer experiment which gives us  $R(N)$  for certain values of  $N$ . For an enumeration study this information would be exact but typically involve quite small  $N$ , while for a Monte Carlo computer experiment there would be sampling error associated with these values but one could reach large values of  $N$  and reduce the influence of corrections to scaling.

Our principal goal in performing the computer experiment is to estimate some of the quantities associated with this asymptotic expansion such as the leading exponent  $x_1$ , the leading-order correction to scaling exponent  $x_2$ , and the leading amplitude  $\alpha_1$ .

We obtain estimates by performing nonlinear fits of our data using Eq. (A1) by appropriately truncating the expansion after  $M$  terms. We cannot perform a fit with an arbitrarily large number of terms, as we only have data over a finite range for  $N$ . There may also be asymptotic corrections with comparable exponents, which makes it extremely difficult to reliably distinguish between them, and for Monte Carlo there is statistical error on  $R(N)$ . Each of these factors is relevant in our case: We have data up to  $N = 2^{25}$ , our data has statistical error, and next-to-leading corrections to scaling have comparable exponents which are all around 1:  $2\Delta_1 \approx 1.06$ ,  $\Delta_2 \approx 1$ , and 1. So, in practice we can only make reliable fits of the leading

correction to scaling. It is possible to fit the three competing next-to-leading corrections with a single “effective” term with exponent approximately one, but it is difficult to see how to sensibly interpret such a procedure.

In the general case, we attempt to simultaneously adjust all amplitudes  $\alpha_i$  and all exponents  $x_i$  by the direct application of a  $2M$ -parameter nonlinear fit routine. If we include in the fit  $2M$  data points for  $R(N)$  then the fitted function  $S(N)$  will be exact at those points, but more frequently we perform a nonlinear weighted least squares fit [weighting appropriately by the statistical error in our estimates of  $R(N)$ ] and so  $S(N)$  will instead be an approximation. Regardless, by design we have  $S(N) \approx R(N)$ , where we are careful to ensure that we can meaningfully interpret the fit by confirming that the model is appropriate and the reduced  $\chi^2$  value is approximately one. If the model is appropriate then  $S(N)$  will be the same as  $R(N)$  at the data points to within statistical accuracy, and so  $S(N) = R(N) + \Delta R(N)$  where  $\Delta R(N)$  is of the same order as the statistical accuracy of our estimate.

The truncation will result in somewhat distorted values for the amplitudes and exponents in the truncated model. We denote these errors as  $\Delta\alpha_i$  for the amplitudes and  $\Delta x_i$  for the exponents:

$$S(N) = \sum_{i=1}^M (\alpha_i + \Delta\alpha_i) N^{x_i + \Delta x_i}. \quad (\text{A3})$$

But we have

$$S(N) = R(N) + \Delta R(N), \quad (\text{A4})$$

$$\sum_{i=1}^M (\alpha_i + \Delta\alpha_i) N^{x_i + \Delta x_i} = \sum_{i=1}^M \alpha_i N^{x_i} + O(N^{x_{M+1}}) + \Delta R(N). \quad (\text{A5})$$

We restrict attention only to rather large values of  $N$ , where the truncated model accurately fits the data, and so  $|\Delta\alpha_i| \ll |\alpha_i|$  and  $|\Delta x_i| \ll |x_i|$ . In addition, we can expect that the neglected terms represented by  $O(N^{x_{M+1}})$  are small and dominated by the first neglected correction to scaling corresponding to exponent  $x_{M+1}$ . In this limit, we may linearize Eq. (A5) around  $\alpha_i$  and  $x_i$ ,

$$\sum_{i=1}^M (\Delta\alpha_i N^{x_i} + \alpha_i \Delta x_i N^{x_i} \log N) = O(N^{x_{M+1}}) + \Delta R(N). \quad (\text{A6})$$

We now perform fits according to Eq. (A3) in an interval  $N \geq N_{\min}$ , where  $N_{\min}$  is systematically varied but where it must be sufficiently large that the truncated model is accurate. The errors in estimates depend on this choice  $N_{\min}$ , and so  $\Delta\alpha_i$  and  $\Delta x_i$  should be understood to be implicit functions of  $N_{\min}$ . Equation (A6) is valid for any value of  $N$  in the fitting range, and therefore is valid for  $N_{\min}$ :

$$\begin{aligned} & \sum_{i=1}^M (\Delta\alpha_i N_{\min}^{x_i} + \alpha_i \Delta x_i N_{\min}^{x_i} \log N_{\min}) \\ & = O(N_{\min}^{x_{M+1}}) + \Delta R(N_{\min}). \end{aligned} \quad (\text{A7})$$

Neglecting logarithmic corrections, and assuming that all error terms on the left-hand side of Eq. (A7) are of the same order as the right-hand side, we thus find for the error in the exponents that

$$\Delta x_i \propto N_{\min}^{-(x_i - x_{M+1})} + N_{\min}^{-x_i} \Delta R(N_{\min}), \quad (\text{A8})$$

and similarly for the amplitudes

$$\Delta\alpha_i \propto N_{\min}^{-(x_i - x_{M+1})} + N_{\min}^{-x_i} \Delta R(N_{\min}). \quad (\text{A9})$$

How are we to interpret these expressions, and use them to obtain the most accurate estimates of  $\alpha_i$  and  $x_i$  possible? First, note that  $\Delta R(N_{\min})$  is the statistical error and is a known quantity. The corresponding statistical error in the estimates for  $\alpha_i$  and  $x_i$  are of order  $N_{\min}^{-x_i} \Delta R(N_{\min})$ . Typically, we expect that the statistical errors will increase as  $N_{\min}$  increases, but the rate of increase will be smallest for the leading term with  $i = 1$ . In contrast, the systematic errors, of order  $N_{\min}^{-(x_i - x_{M+1})}$  (neglecting logarithmic factors) are unknown, and decay with increasing  $N_{\min}$ . This decay is most rapid for the leading term. By definition, the systematic error from truncation is not fitted, and so the only way which it can be accounted for in the analysis is to extrapolate to  $N_{\min} \rightarrow \infty$ , where this error vanishes. Now, we expect that for sufficiently large  $N_{\min}$ , a plot of  $\alpha_i$  and  $x_i$  against  $N_{\min}^{-(x_i - x_{M+1})}$  would be linear. If we have an idea of the value of  $x_{M+1}$ —even if we do not know it exactly—plotting our estimates in this way can greatly facilitate extrapolation. These observations are the motivation for the various power laws appearing in plots in the main text. Then, to interpret these fits requires judgment to decide when  $N_{\min}$  is sufficiently large that a reliable extrapolation can be made, but as small as possible so as to reduce statistical error.

Interpretation of the fits is a balancing act between systematic error and statistical error. Acquiring more data at large values of  $N$  may reduce systematic error at the expense of increasing statistical error. One of us (N.C.) is perennially surprised at how subtle the interpretation of such fits is: In principle, being able to perform accurate computer experiments for extremely large systems should make it possible to reduce the influence of corrections to scaling until they are negligible, but what happens in practice is that the extremely accurate values make it necessary to incorporate the leading-order correction to scaling even for  $N$  of the order of tens of millions, and in order to get a good handle on this term it is necessary to perform computer experiments for  $N$  of the order of tens of thousands, where poorly controlled next-to-leading corrections make things extremely difficult. One circumstance where this trap has been avoided is the calculation of the growth constant  $\mu$  for SAWs in Ref. [27], but this relies on the fact that the asymptotic corrections for  $\mu$  are smaller than for critical exponents.

## APPENDIX B: MONTE CARLO DATA

The global observables  $R_G^2$ ,  $R_E^2$ , and  $R_H^{-1}$  are correlated; therefore calculating ratios may be viewed as form of variance reduction. Hence we report the ratios as well. All raw data are listed in Tables II–IV.

TABLE II. Raw data on the end-to-end distance and the gyration radius.

$N$	$\langle R_E^2 \rangle$	$\langle R_G^2 \rangle$
512	$1.8336722(58) \times 10^3$	$2.9152213(82) \times 10^2$
724	$2.7631843(94) \times 10^3$	$4.396899(14) \times 10^2$
1024	$4.1626998(41) \times 10^3$	$6.6290075(60) \times 10^2$
1448	$6.2667402(69) \times 10^3$	$9.986311(10) \times 10^2$
2048	$9.433354(11) \times 10^3$	$1.5040985(16) \times 10^3$
2896	$1.4192522(18) \times 10^4$	$2.2640087(26) \times 10^3$
4096	$2.1353085(28) \times 10^4$	$3.4076501(41) \times 10^3$
5792	$3.2112468(46) \times 10^4$	$5.1264340(69) \times 10^3$
8192	$4.8297971(73) \times 10^4$	$7.712466(11) \times 10^3$
11 584	$7.261391(12) \times 10^4$	$1.1598097(18) \times 10^4$
16 384	$1.0918781(19) \times 10^5$	$1.7443237(29) \times 10^4$
23 168	$1.6412837(31) \times 10^5$	$2.6224555(48) \times 10^4$
32 768	$2.4675807(49) \times 10^5$	$3.9432498(75) \times 10^4$
46 336	$3.7087199(80) \times 10^5$	$5.927288(12) \times 10^4$
65 536	$5.575269(13) \times 10^5$	$8.911266(20) \times 10^4$
92 672	$8.378786(18) \times 10^5$	$1.3393305(27) \times 10^5$
131 072	$1.2594736(32) \times 10^6$	$2.0133731(50) \times 10^5$
185 344	$1.8926972(46) \times 10^6$	$3.0258005(71) \times 10^5$
262 144	$2.8449071(51) \times 10^6$	$4.5482716(79) \times 10^5$
524 288	$6.425547(21) \times 10^6$	$1.0273486(32) \times 10^6$
1 048 576	$1.4512152(53) \times 10^7$	$2.3203899(83) \times 10^6$
2 097 152	$3.277454(13) \times 10^7$	$5.240600(21) \times 10^6$
4 194 304	$7.401657(33) \times 10^7$	$1.1835309(52) \times 10^7$
8 388 608	$1.6715288(79) \times 10^8$	$2.672847(13) \times 10^7$
16 777 216	$3.774819(19) \times 10^8$	$6.036144(31) \times 10^7$
33 554 432	$8.524591(30) \times 10^8$	$1.3631415(48) \times 10^8$

TABLE III. Raw data on the hydrodynamic radius.

$N$	$\langle R_H^{-1} \rangle$	$\langle R_H^{-2} \rangle$
512	$8.400655(10) \times 10^{-2}$	$7.124977(20) \times 10^{-3}$
724	$6.9369818(95) \times 10^{-2}$	$4.858005(16) \times 10^{-3}$
1024	$5.7174946(23) \times 10^{-2}$	$3.2998280(32) \times 10^{-3}$
1448	$4.7059008(21) \times 10^{-2}$	$2.2352724(25) \times 10^{-3}$
2048	$3.8678939(19) \times 10^{-2}$	$1.5099458(18) \times 10^{-3}$
2896	$3.1760618(17) \times 10^{-2}$	$1.0180323(14) \times 10^{-3}$
4096	$2.6052526(15) \times 10^{-2}$	$6.849483(10) \times 10^{-4}$
5792	$2.1356062(14) \times 10^{-2}$	$4.6023215(78) \times 10^{-4}$
8192	$1.7492303(12) \times 10^{-2}$	$3.0875098(56) \times 10^{-4}$
11 584	$1.4321075(11) \times 10^{-2}$	$2.0694197(42) \times 10^{-4}$
16 384	$1.17175446(93) \times 10^{-2}$	$1.3853382(31) \times 10^{-4}$
23 168	$9.5844326(85) \times 10^{-3}$	$9.268299(23) \times 10^{-5}$
32 768	$7.8358154(73) \times 10^{-3}$	$6.194747(17) \times 10^{-5}$
46 336	$6.4050264(66) \times 10^{-3}$	$4.138895(13) \times 10^{-5}$
65 536	$5.2334061(57) \times 10^{-3}$	$2.7631557(94) \times 10^{-5}$
92 672	$4.2756617(44) \times 10^{-3}$	$1.8443015(61) \times 10^{-5}$
131 072	$3.4920527(44) \times 10^{-3}$	$1.2302115(50) \times 10^{-5}$
185 344	$2.8519139(34) \times 10^{-3}$	$8.205089(33) \times 10^{-6}$
262 144	$2.3284895(21) \times 10^{-3}$	$5.469584(17) \times 10^{-6}$
524 288	$1.5518439(26) \times 10^{-3}$	$2.429360(14) \times 10^{-6}$
1 048 576	$1.0338339(19) \times 10^{-3}$	$1.0781790(76) \times 10^{-6}$
2 097 152	$6.885461(14) \times 10^{-4}$	$4.782487(40) \times 10^{-7}$
4 194 304	$4.584827(11) \times 10^{-4}$	$2.120460(20) \times 10^{-7}$
8 388 608	$3.0523993(78) \times 10^{-4}$	$9.39847(10) \times 10^{-8}$
16 777 216	$2.0319314(57) \times 10^{-4}$	$4.164729(53) \times 10^{-8}$
33 554 432	$1.3525101(26) \times 10^{-4}$	$1.845312(22) \times 10^{-8}$

TABLE IV. Raw data on amplitude ratios.

$N$	$\langle R_E^2 \rangle / \langle R_G^2 \rangle$	$\langle R_G^2 \rangle^{1/2} \langle R_H^{-1} \rangle$
512	6.289993(10)	1.4343295(15)
724	6.284393(11)	1.4546008(17)
1024	6.2795219(32)	1.47207524(52)
1448	6.2753307(35)	1.48711752(57)
2048	6.2717661(37)	1.50007398(61)
2896	6.2687579(39)	1.51122100(68)
4096	6.2662199(42)	1.52081824(73)
5792	6.2640947(45)	1.52907518(80)
8192	6.2623255(48)	1.53618530(86)
11 584	6.2608471(51)	1.54230033(94)
16 384	6.2596072(54)	1.5475694(10)
23 168	6.2585759(59)	1.5521027(11)
32 768	6.2577340(62)	1.5560062(12)
46 336	6.2570265(68)	1.5593691(13)
65 536	6.2564277(71)	1.5622629(14)
92 672	6.2559513(66)	1.5647580(13)
131 072	6.2555401(80)	1.5669058(16)
185 344	6.2551952(75)	1.5687600(15)
262 144	6.2549194(55)	1.5703535(11)
524 288	6.254495(10)	1.5729210(21)
1 048 576	6.254187(11)	1.5748209(23)
2 097 152	6.253967(12)	1.5762441(26)
4 194 304	6.253877(13)	1.5772942(29)
8 388 608	6.253739(15)	1.5780775(32)
16 777 216	6.253693(16)	1.5786606(35)
33 554 432	6.253636(10)	1.5791045(24)

- [1] N. Clisby, *Phys. Rev. Lett.* **104**, 055702 (2010).
- [2] N. Clisby, *J. Stat. Phys.* **140**, 349 (2010).
- [3] N. Madras and G. Slade, *The Self-Avoiding Walk* (Springer Science & Business Media, Berlin/Heidelberg, 2013).
- [4] M. Lal, *Mol. Phys.* **17**, 57 (1969).
- [5] N. Madras and A. D. Sokal, *J. Stat. Phys.* **50**, 109 (1988).
- [6] B. Li, N. Madras, and A. D. Sokal, *J. Stat. Phys.* **80**, 661 (1995).
- [7] T. Kennedy, *J. Stat. Phys.* **106**, 407 (2002).
- [8] M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics* (Oxford University Press, Oxford, UK, 1988).
- [9] B. Dünweg, D. Reith, M. Steinhauser, and K. Kremer, *J. Chem. Phys.* **117**, 914 (2002).
- [10] S. Caracciolo, B. M. Mognetti, and A. Pelissetto, *J. Chem. Phys.* **125**, 094904 (2006).
- [11] L. Schäfer and A. Baumgärtner, *J. Phys. (Paris)* **47**, 1431 (1986).
- [12] M. L. Mansfield and J. F. Douglas, *Phys. Rev. E* **81**, 021803 (2010).
- [13] B. Liu and B. Dünweg, *J. Chem. Phys.* **118**, 8061 (2003).
- [14] P. Sunthar and J. R. Prakash, *Europhys. Lett.* **75**, 77 (2006).
- [15] L. Greengard and V. Rokhlin, *J. Comput. Phys.* **73**, 325 (1987).
- [16] M. Hasenbusch, F. P. Toldin, A. Pelissetto, and E. Vicari, *J. Stat. Mech.: Theor. Exp.* (2007) P02016.
- [17] M. Hasenbusch, *Phys. Rev. B* **82**, 174433 (2010).
- [18] K. Kremer, A. Baumgärtner, and K. Binder, *Z. Phys. B* **40**, 331 (1981).
- [19] H. Shimada and S. Hikami, [arXiv:1509.04039](https://arxiv.org/abs/1509.04039) (unpublished).
- [20] F. Kos, D. Poland, D. Simmons-Duffin, and A. Vichi, *J. High Energy Phys.* **08** (2016) 036.
- [21] R. D. Schram, G. T. Barkema, and R. H. Bisseling, *J. Stat. Mech.: Theor. Exp.* (2011) P06019.
- [22] N. Clisby, R. Liang, and G. Slade, *J. Phys. A: Math. Theor.* **40**, 10973 (2007).
- [23] T. Prellberg, *J. Phys. A: Math. Gen.* **34**, L599 (2001).
- [24] D. MacDonald, S. Joseph, D. L. Hunter, L. L. Moseley, N. Jan, and A. J. Guttmann, *J. Phys. A: Math. Gen.* **33**, 5973 (2000).
- [25] R. Guida and J. Zinn-Justin, *J. Phys. A: Math. Gen.* **31**, 8103 (1998).
- [26] P. Belohorec, Ph.D. thesis, University of Guelph, Guelph, Canada, 1997 (unpublished).
- [27] N. Clisby, *J. Phys. A: Math. Theor.* **46**, 245001 (2013).