

Weighting Methods for Monte Carlo Calculation of Polymer Configurations

Frank L. McCrackin

Institute for Materials Research, National Bureau of Standards, Washington, D.C. 20234

(September 21, 1972)

In the Rosenbluth and Rosenbluth method of computing polymer configurations, the configurations are weighted in order to remove bias of the estimated parameters of the configurations. This weighting method is investigated and generalized for importance sampling and Boltzmann factors. The estimates are found to be unbiased in the limit for an infinite sample of configurations, but to have a bias for a finite sample. The standard deviations of the estimates are also derived.

Key words: Boltzmann factors; importance sampling; Monte Carlo; Nonself-intersecting walks; polymer configurations; random walks.

Polymer molecules in solution have been simulated by non-self-intersecting random walks on a lattice by many investigators [1–7]¹. The procedure is to calculate many non-self-intersecting random walks of a given number of steps or segments, n , and calculate a parameter of each walk, such as the square of the end-to-end distance, r^2 . The values of the parameter are then averaged. If the generated walks are a random sample of all non-self-intersecting walks of length n , then the average value of the parameter, such as $\langle r^2 \rangle$, will be an estimate of the average value of the parameter over all non-self-intersecting random walks of length n . Of course, for very small values of n , all possible walks may be generated and the average value of a parameter of the walks directly calculated. However, for large values of n , the number of possible walks becomes too large to generate even on a computer, so it is possible to generate only a sample of the possible walks.

The direct method of generating a sample of non-self-intersecting random walks is to generate a sample of random walks and discard those that intersect themselves. However, for large n this method is impractical because almost all generated random walks will be self-intersecting and must be discarded. Three practical methods, chain enrichment [1], dimerization [6] and the method of Rosenbluth and Rosenbluth [2] have been used to generate non-self-intersecting random walks for Monte Carlo studies of polymer configurations. This paper investigates the accuracy and bias of estimates of the parameters of walks generated by the method of Rosenbluth and Rosenbluth. Rosenbluth and Rosenbluth gave an intuitive justification, but no proof that their method is unbiased. Formulas for the bias and variance of the estimates are given in the appendix of reference [4], but without detailed derivations. Also, importance sampling and Boltzmann factors that are used in a current Monte Carlo study [11] are not considered in [4]. This paper gives complete derivations of the bias and variance of the estimates and generalizes the derivations for importance sampling and Boltzmann factors.

Although random walks are usually generated on three dimensional lattices for large values of n , the methods of generation will be illustrated for $n=4$ on a square lattice, following Rosenbluth and Rosenbluth [2]. All 25 nonintersecting random walks are shown in figure 1.

In the Rosenbluth and Rosenbluth [2] method of generating random walks, only steps for which the walk does not intersect itself are taken. Thus, in the walk shown in figure 2, either

¹ Figures in brackets indicate the literature references at the end of this paper.

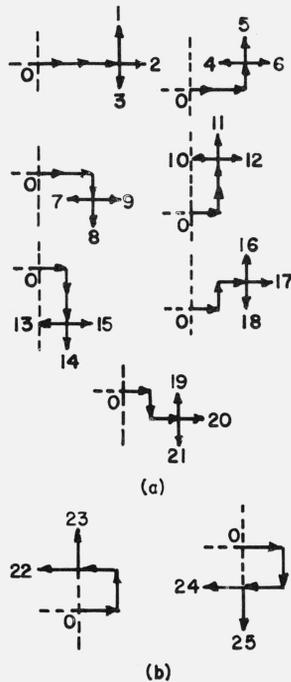


FIGURE 1. All nonintersecting walks of 4 steps on a square lattice. The direction of the first step is fixed.

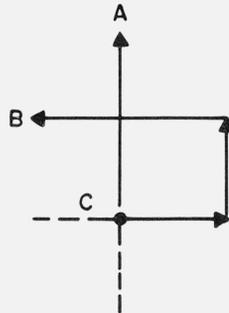


FIGURE 2. Three possible steps for a walk on a square lattice. Step C is not allowed for a nonintersecting walk.

step A or B would be chosen, each with probability $\frac{1}{2}$. Although there is still attrition of the walks due to trapping [3], this method allows efficient generation of long walks. However, different walks are generated with different probabilities. In the example of figure 1, walks 1 to 21 are each generated with probability $(\frac{1}{3})^3$ while walks 22 to 25 are each generated with probability of $(\frac{1}{3})^2 \frac{1}{2}$.

Consider a very large number, m , of random walks generated by this method. The average value of r^2 over the sample will be

$$\langle r^2 \rangle = \frac{1}{m} \sum_{i=1}^m r_i^2 \cong \sum_{j=1}^{25} r_j^2 P_j = 6.74 \quad (1)$$

where r_j^2 and P_j are the square end-to-end distance and probability for each of the 25 walks of figure 1. However, the correct average [2] obtained by averaging r^2 over the 25 configurations is 7.04. This method is seen to produce compact walks with too great a probability, so simple averages

of the walks produce incorrect, biased, results. In order to remove the bias, Rosenbluth and Rosenbluth [2] weighted each configuration to give, for any parameter v ,

$$\langle \hat{v} \rangle = \frac{\sum_{k=1}^m v_k w_k}{\sum_{k=1}^m w_k} \quad (2)$$

where the weights, w_k , are the reciprocals of the probabilities of the walks. They gave an intuitive argument, but no proof, for the weighting procedure. In the above example, $w_k = 3^3$ for walks 1 to 21 and $w_k = 3^2$ for walks 22 to 25.

Also, the number of walks, T , of n steps was estimated by

$$\hat{T} = \frac{1}{m} \sum_{k=1}^m w_k. \quad (3)$$

The carets placed over v and T indicate that eqs (2) and (3) give estimates for $\langle v \rangle$ and T rather than their true values.

We first prove that eq (3) gives a correct estimate. A complication arises due to trapping [2] of the walks, when the walk cannot be continued to n steps. In this case the walk is terminated and its weighting factor is defined to be zero. Let W_T be the number of trapped configuration of walks of less than n steps, so that every walk gives either one of the T n -step walks or one of the W_T trapped configurations. The expectation [9] of w is given by the summation of $w_i P_i$ over all walks where P_i is the probability of the walk. That is:

$$E(\hat{T}) = E(w) = \sum_{i=1}^T w_i \frac{1}{w_i} + \sum_{i=1}^{W_T} 0 \cdot P_i = T. \quad (4)$$

Thus eq (3) estimates the number of walks. This proof follows the method of Lehman and Weiss [8].

The variance and standard deviation of T are also of interest. The variance of w is

$$\text{Var } w = E(w^2) - (Ew)^2 \quad (5)$$

$$= \sum_{i=1}^T w_i^2 \frac{1}{w_i} - T^2$$

$$= \sum_{i=1}^T w_i - T^2. \quad (6)$$

Because \hat{T} is an average of m independent values of w , its variance is [9]

$$\text{Var } \hat{T} = \frac{1}{m} \text{Var } w = \left[\sum_{i=1}^T w_i - T^2 \right] / m \quad (7)$$

and its standard deviation is

$$\sigma(\hat{T}) = \left[\sum_{i=1}^T w_i - T^2 \right]^{1/2} / \sqrt{m} \quad (8)$$

The investigation of the estimate given by eq (2) that was outlined in reference [10] is now given in detail as follows:

For a parameter v , define:

$$R = \frac{1}{m} \sum_{k=1}^m v_k w_k \quad (9)$$

so that

$$\langle \hat{v} \rangle = R / \hat{T}. \quad (10)$$

Let $E(R)$ and $E(\hat{T})$ be the expectation values of R and \hat{T} , with

$$E(\hat{T}) = T. \quad (11)$$

The expectation of R is given by the expectation of vw as

$$E(R) = \sum_{i=1}^T v_i w_i \frac{1}{w_i} = \sum_{i=1}^T v_i = T \langle v \rangle \quad (12)$$

As the sample size m tends to infinity, R and \hat{T} tend to their expectation values and $\langle \hat{v} \rangle$ tends to the average of v over the walks. Thus, eq (2) is asymptotically unbiased, i.e., it is unbiased in the limit of infinite sample size.

To investigate eq (2) for large but finite values of m , $\langle \hat{v} \rangle$ is expanded about $E(R)$ and $E(\hat{T})$ to give

$$\begin{aligned} \langle \hat{v} \rangle = & \frac{E(R)}{E(\hat{T})} + \frac{1}{E(\hat{T})} [R - E(R)] - \frac{E(R)}{[E(\hat{T})]^2} [\hat{T} - E(\hat{T})] - \frac{1}{[E(\hat{T})]^2} [R - E(R)][\hat{T} - E(\hat{T})] \\ & + \frac{E(R)}{[E(\hat{T})]^3} [\hat{T} - E(\hat{T})]^2 + \text{higher order terms.} \end{aligned} \quad (13)$$

We now take the expectation value of $\langle \hat{v} \rangle$. The first term gives $\langle v \rangle$ and the expectation value of the two first order terms vanish. Expansion and substituting for $E(R)$ and $E(\hat{T})$ from eqs (12) and (11) in eq (13) gives the approximation

$$E(\langle \hat{v} \rangle) = \langle v \rangle + [\langle v \rangle E(\hat{T}^2) - E(R\hat{T})] / T^2 \quad (14)$$

By the definition of the variance,

$$\text{Var } \hat{T} = E(\hat{T}^2) - [E(\hat{T})]^2. \quad (15)$$

Substituting eqs 4 and 7 gives

$$E(\hat{T}^2) = \frac{1}{m} \sum_{i=1}^T w_i + \frac{m-1}{m} T^2 \quad (16)$$

To evaluate $E(R\hat{T})$, we substitute from eqs (3) and (9) and rearrange terms to give

$$\begin{aligned}
E(R\hat{T}) &= \frac{1}{m^2} E\left(\sum_{k=1}^m v_k w_k \sum_{k=1}^m w_k\right) \\
&= \frac{1}{m^2} E\left(\sum_{k=1}^m v_k w_k^2\right) + \frac{1}{m^2} E\left(\sum_{l \neq k}^m \sum_{l=1}^m v_k w_k w_l\right)
\end{aligned} \tag{17}$$

The first term on the right hand side gives $\frac{1}{m} \sum_{i=1}^T v_i w_i$. The second term is a summation of $m(m-1)$ quantities, and because $v_k w_k$ and w_l are independent, the expectation value of their product is the product of their expectation values, so

$$\begin{aligned}
E\left(\sum_{l \neq k}^m \sum_{k=1}^m v_k w_k w_l\right) &= m(m-1)E(v_k w_k)E(w_l) \\
&= m(m-1)E(R)E(w)
\end{aligned} \tag{18}$$

Substituting eqs (16), (17), and (18) in eq (14) finally gives

$$E[\langle \hat{v} \rangle] = \langle v \rangle - \frac{1}{mT^2} \sum_{i=1}^T (v_i - \langle v \rangle) w_i \tag{19}$$

This gives the approximate bias for the average given by eq (2). That is, $\langle \hat{v} \rangle$ calculated from a sample of m walks will, on the average, differ from the true average $\langle v \rangle$ by the right-hand term. However, for increasing sample size m , the bias will approach zero, so the average is asymptotically unbiased.

For the simple example shown in figure 1, the summation was evaluated to give

$$E[\langle \hat{r}^2 \rangle] = \langle r^2 \rangle - 0.233/m$$

Thus, for a sample size of 100 walks, the mean value of r^2 calculated by eq (2) would be on the average too low by about only 0.002 lattice units squared. For longer walks on three-dimensional lattices, it is not practical to calculate the bias from eq (19). Methods of estimating the bias for these cases will be given in a later publication [11].

Using the formulas of *Ku* [12], the approximate standard deviation of $\langle \hat{v} \rangle$ (from its average value, not from the true value $\langle v \rangle$) is derived:

$$\sigma(\langle \hat{v} \rangle) = \left[\frac{1}{T^2} \sum_{i=1}^T (v_i - \langle v \rangle)^2 w_i \right]^{1/2} / \sqrt{m} \tag{20}$$

The standard deviation decreases as the square root of the sample size. Therefore, for sufficiently large sample size, the bias will be much smaller than the standard deviation, so the bias may be neglected. For the example of figure 1,

$$\sigma(\langle \hat{r}^2 \rangle) = 3.58/\sqrt{m}$$

For $m=100$, the standard deviation of this estimate is about 0.36, which is much larger than the bias of the estimate.

To apply this method to long walks on various lattices, weighting factors, w , must be computed for each walk as the walk is generated. Let q be the maximum number of choices for a step of a walk on a lattice, i.e., one less than the coordination number of the lattice. Thus, $q=3, 5$ and 11

for square, simple cubic, and face-centered cubic lattices, respectively. Also, when two non-adjacent sites of a walk lie within one lattice distance of each other, they are said to form a contact. Thus in figure 1, walks 4, 7, 18, 19, and 22 to 25 contain a contact. Now let the i th site of a walk form C_i contacts, so the next step of the walk has $q - C_i$ possible directions. One of these directions is chosen randomly, so it is chosen with a probability of $\frac{1}{q - C_i}$. Therefore, the probability of generating a particular walk of n steps is $\prod_{i=1}^{n-1} \frac{1}{q - C_i}$ and $w = \prod_{i=1}^{n-1} (q - C_i)$.

During generation of the walks, some of the walks are trapped, i.e., all surrounding sites are occupied so the walk cannot be continued. Because w_i is defined to be zero for these walks, they should not be used in the calculation of eq (2).

Average for Walks With Boltzmann Factors

In calculations for polymer walks with nearest neighbor interaction energies, the averages over parameters of the walks multiplied by Boltzmann factors $\alpha_i = \exp(-p_i \epsilon / kt)$ are desired, where p_i is the number of contacts of the i th walk, ϵ is the energy per contact, k is the Boltzmann constant, and t is the temperature. That is, the quantities

$$S = \sum_{i=1}^T \alpha_i \quad (21)$$

and

$$\langle v \rangle = \sum_{i=1}^T \alpha_i v_i / S \quad (22)$$

are to be estimated. The derivation follows as previously. Thus corresponding to eq (4).

$$E(w\alpha) = \sum_{i=1}^T w_i \alpha_i \frac{1}{w_i} = S \quad (23)$$

so an estimate of S from a sample of m walks is

$$\hat{S} = \frac{1}{m} \sum_{k=1}^m w_k \alpha_k \quad (24)$$

to estimate $\langle v \rangle$ we propose

$$\langle \hat{v} \rangle = \frac{\sum_{k=1}^m v_k \alpha_k w_k}{\sum_{k=1}^m \alpha_k w_k} \quad (25)$$

The derivation is the same as previously with αw replacing w and S replacing T (except for limits of the summations). Equation (19) then becomes

$$E[\langle \hat{v} \rangle] = \langle v \rangle - \frac{1}{mS^2} \sum_{i=1}^T (v_i - \langle v \rangle) \alpha_i w_i \quad (26)$$

and eq (20) becomes

$$\sigma(\langle \hat{v} \rangle) = \left[\frac{1}{S^2} \sum_{i=1}^T (v_i - \langle v \rangle)^2 \alpha_i w_i \right]^{1/2} / \sqrt{m} \quad (27)$$

Importance Sampling

In the described method, all allowable steps from a given site have been given equal probabilities. However, walks are sometimes performed in which the allowable steps are taken with different probabilities. In this way, many walks for which a desired parameter is large (important walks) will be generated. The average of the parameter over this sample of generated walks will be calculated with a higher accuracy. This technique is called Importance Sampling [13]. One method used gives twice the probability to allowable steps that point toward the origin than to other steps. Thus, in figure 3a on the rectangular lattice, step A would be taken with a probability of $\frac{1}{2}$ and steps B and C would be taken with a probability of $\frac{1}{4}$ each. In figure 3b, step A would be taken with a probability $\frac{2}{3}$ and step B with probability $\frac{1}{3}$. With this method many more coiled walks containing large number of contacts are generated than when all steps have equal probability.

For walks with large nearest-neighbor interaction energies of attraction, the Boltzmann factors α_i are large for walks containing many contacts. Then, for a parameter v such as end-to-end distance, many walks for which $v_k \alpha_k$ is large are generated by the above method so that more accurate values for the average of v by eq (25) should be obtained than for walks generated in the ordinary way. This method of generating walks will be used in reference [11].

Many other methods of choosing unequal probabilities for the steps may be used. For another example, the step in the same direction as the preceding step may be given a higher probability than the other steps. For any such method, weighting factors w_i may be calculated so that eq (2) will apply. We now derive the weighting factors for a general method of chain generation with Importance Sampling.

For any step in a walk, let there be s allowable steps. Let a multiplicity m_k ($k=1$ to s) be assigned to each step proportional to the probability of the step. For example, let steps toward the start of the walk be given twice the probability of other steps. Thus, in figure 3a, $s=3$, step A has a multiplicity of 2 and steps B and C each have a multiplicity of 1; in figure 3b, $s=2$, and steps A and B have multiplicities of 2 and 1 respectively. In general, the probability of a step is

$$\frac{m_c}{\sum_{k=1}^s m_k} \quad (28)$$

where m_c is the multiplicity of the chosen step. The probability of a given N step walk is then the product of the factors (28) for all steps of the walk and the weighting factor is the reciprocal of the probability, or the products of the factors

$$\frac{\sum_{k=1}^s m_k}{m_c} \quad (29)$$

over all steps of the walk.

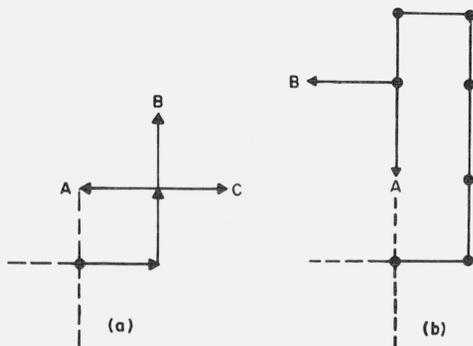


FIGURE 3. Typical configurations on a square lattice illustrating assignment of multiplicities to allowed steps.

Conclusions

The biases and standard deviations for parameters of walks with excluded volume generated by the method of Rosenbluth and Rosenbluth [2] have been derived and extended to the cases of importance sampling and Boltzmann factors. The direct calculation of the biases and standard deviations involve summations over all walks so is generally not feasible. However, the formulas will be used to estimate the biases and standard deviations from Monte Carlo calculations in a later paper.

I wish to thank Jacob Mazur and Charles Guttman for their help and discussions.

References

- [1] Wall, F. T., Windwer, S., and Gans, P. J., in *Methods in Computational Physics I* (Academic Press, Inc., New York, 1963).
- [2] Rosenbluth, M. N., and Rosenbluth, A. W., *J. Chem. Phys.* **23**, 356 (1955).
- [3] Mazur, J., and McCrackin, F. L., *J. Chem. Phys.* **49**, 648 (1968).
- [4] McCrackin, F., *J. Chem. Phys.* **47**, 1980 (1967).
- [5] Medalia, A. I., in *Surface and Colloid Science*, Vol. 4, edited by E. Matijevic (John Wiley & Sons, 1971).
- [6] Alexandrowicz, Z., *J. Chem. Phys.* **51**, 561 (1969).
- [7] Warvari, H. E., Knaell, K. K., and Scott, R. A., III, *J. Chem. Phys.* **55**, 2020 (1971).
- [8] Lehman, R. S., and Weiss, G. H., *J. Soc. Indust. Appl. Math.* **6**, 257 (1958).
- [9] A discussion of calculation of expectation values is given by H. Cramer, *Mathematical Methods of Statistics*, (Princeton University Press, 1946).
- [10] Rosenblatt, J., see Appendix of reference 4.
- [11] Submitted for publication in *Macromolecules*.
- [12] Ku, H. H., *J. Res. Nat. Bur. Stand. (U.S.)*, 70C (Eng. and Instr.) No. 4, 263 (Oct.-Dec. 1966). The formula of the third row of Table 1 is used.
- [13] Importance Sampling is discussed for example by H. Kahn in *Symposium on Monte Carlo Methods* p. 149, edited by H. A. Meyer (John Wiley & Sons, Inc. New York, 1956).

(Paper 76B3&4-371)