A simple self-avoiding walking process as a reasonable non-conventional generator of polymeric linear chains.

David R. Avellaneda B.\textsuperscript{a,\ast}, Ramón E. R. González\textsuperscript{b}

\textsuperscript{a}Departamento de Estatística e Informática, Universidade Federal Rural de Pernambuco, Recife, Pernambuco, CEP 52171-900, Brazil
\textsuperscript{b}Departamento de Física, Universidade Federal Rural de Pernambuco, Recife, Pernambuco, CEP 52171-900, Brazil.

Abstract

In this work, we present a simple and efficient generator of polymeric linear chains, based on a random self-avoiding walk process. The chains are generated using a discrete process of growth, in cubic networks and in a finite time, without border limits and without exploring all the configurational space. First, we thoroughly describe the chains morphology exploring the statistics of two characteristic distances, the radius of gyration and the end-to-end distance. Moreover, we examine the dependence of mean characteristic distances with the number of steps (\(N\)). Despite the simplicity of our procedure, we obtain universal critical exponents, which are in very good agreement with previous values reported in the literature. Moreover, studying the balance between the monomer-monomer interaction and the bending energy, we find that initially, the chains develop by multiple doubling, forming a cluster and increasing its energy. After reaching a given number of steps, the chains stretch and flee from the cluster, which results in a reduction of its interaction energy. However, the behaviour of the bending energy reveals that the chains follow the same folding pathway in both regimes. Additionally, we also characterize the energy of the obtained chains, combining the local interaction energy with its corresponding bending energy but in a discrete version. This analysis is relevant because it allows differentiating between chains of equal interaction energy but with different structures.

\textbf{Keywords:} Self-avoiding random walk, polymer chains, critical exponents, interaction and bending energy, radius of gyration, end-to-end distance.

1. INTRODUCTION

Although the conformational properties of polymer chains in a good solvent have been the subject of intensive experimental, theoretical and numerical studies their full understanding is still an open challenge. Besides its simplicity, the natural self-avoiding random walk (SAW), proposed over half a century ago, is widely accepted as the principal model for dilute polymers \[1, 2, 3\]. SAW describes well a large spectrum of real systems with diverse details (such as bond angles and monomer-monomer potential). Moreover, the equivalence of SAW with the \(n \rightarrow 0\) limit of the
n-vector model [4] has provided an important connection with the theory of phase transitions and critical phenomena [5].

One of the first (but still widely used) theoretical approaches to this subject is the Flory theory [1], which with simple mean field arguments involving the concept of excluded volume brought about the understanding of underlying power laws and the role of dimensionality. The subsequent analytical approaches span rigorous methods that have achieved only limited success [6], approximate methods such as perturbation theory and self-consistent field theory, which break down for long chains [7], and renormalization group (RG) [8] which has yielded reasonably accurate estimates for critical exponents and some universal amplitude ratios.

In parallel with theoretical developments, numerical methods have shown to be a fundamental tool in establishing properties of long SAWs [9]. Exact enumeration methods have been used to find the number of all possible SAWs of finite length $N$, from which universal properties are estimated using techniques such as the ratio method, Padé approximants or differential approximants. Thus, results up to $N = 71$ steps for the square lattice [10, 11] and up to $N = 21$ steps for the cubic lattice [6] have been reported. A large number of Monte Carlo sampling techniques have been proposed ever since the 1950s (see, e.g., [9] for a comprehensive overview), where the pivot algorithm [2] has been shown to perform in time $O(N)$, and studies with SAWs of length up to $N = 80000$ have been reported [12].

Typically, an ensemble of self-repelling chains considers all possible configurations of a given length. Moreover, the growth process is directed by probability values, which are linked to the Boltzmann factor in terms of the potential energy of interaction and are proportional to the number of interactions [13].

Rather than attempting to improve the performance previous algorithms, in the current work we focus on different goals. Instead, in this paper we study relatively small SAW chain ensembles, which might be relevant describing polymer solutions where the polymer growth process has been going on for a limited time. We present a method to generate such ensembles using very simple and efficient numerical algorithm and, surprisingly, we obtain universal critical exponents, which are in very good agreement with the previous reported in the literature. In addition, we carefully examine the balance between the monomer-monomer interaction and the bending energy, addressing its relation with the structure of the chains. This analysis is relevant because it allows differentiating between chains of equal interaction energy but with different structures.

The paper is organized as follows. In the next section we briefly review the SAW and the principal conformation measures, the algorithm, and the energy measures. The subsequent section is devoted to the results of our simulations, and finally the conclusions are drawn.

2. METODOLOGY

Our aim is to efficiently generates an ensemble of linear homopolymer chains in a good solvent. For simplicity, we use the approximation that the solvent molecules are considered the same size as the monomers.
The chains are generated using the pathway of a particle that moves randomly in a cubic network with an unlimited boundary conditions. Following that approach, the particle is not allowed to occupy the sites it has visited before (a particle with property self-avoiding). We use the idea of a model known as “true” self-avoiding random walk [13] but in 3d. Thus, each steps given by the walker can be interpreted as a monomer (or a set of monomers of the same type), and the steps not visited by the walker (empty sites of the network) can be considered as molecules of the solvent, thus, the trajectory described by this particle defines a homopolymeric chain in a good solvent [13]. The probability of the walker taking a step in the direction $i$ in this type of chain depends on the number of times $n_j$ that the next time the site to be occupied is already “visited”, and is given by:

$$ p_i = \frac{e^{-gn_i}}{\sum_{j=1}^{3d} e^{-gn_j}} = \frac{1}{\sum_{j=1}^{3d} e^{-g(n_j-n_i)}}. $$

(1)

In general, the sum runs through all possible 3$d$ paths from the position occupied by the walker at each instant of time, including the address $i$, and $g$ is a positive parameter which measures the intensity with which the walk avoids it self. For the sake of simplicity, in this work we implement the limiting case $g = \infty$, which corresponds in the 3$d$ case to a discrete domain of probabilities $p_i(g = \infty) = [1/6, 1/5, 1/4, 1/3, 1/2, 1]$.

2.1. Numerical Algorithm

For a $d-$dimensional network with free boundaries the algorithm to generate a one chain is as follows:

1. Choose the number of attempts $N'$.
2. Choose the origin of the polymer, which in our case is the origin of the coordinate system.
3. Generate the first step randomly or choose it arbitrarily from a point in the cubic network.
4. Choose the following step randomly from one of the $2 \times d$ possible steps.
5. If the given step leads to self-intersection, go to item 4. and try again with another step. This step is most important to ensure the SAW.
6. If the step leads to an available location, add the step to the walk.
7. If the number of attempts is reached or if the number of possible steps is zero (the walker gets stuck), the simulation is accepted and saved.

Thus, the random chain is formed by $N$ steps generated from $N'$ attempts, being that, for the three-dimensional case $N' > N$, this is because the chains can get trapped before reaching the total number of attempts.

After generating the random chain, we store the positions of each of the monomers that constitute it and proceed to compute the characteristic measurements of its configuration. Starting by calculating the displacement $r_i$ of each monomer in order to obtain the mass center $r_c$ given in Eq. (3) and with which we calculate the radius of gyration, $R_g$, given in Eq. (4). Next, we compute the end-to-end vector module, $R_{ee}$, which can be easily derived from the distance of the $N$th monomer to the origin of the chain as described in Eq. (2).
2.2. Characteristic distances

As is proven below, this simple algorithm efficiently generates an ensemble of linear homopolymer chains in a good solvent; each of them is formed by \( N \) monomers in positions \( \{r_0, r_1, ..., r_N\} \) in the space of dimension \( d \). The separation distance between a monomer and its nearest neighbor is \( b = r_i - r_{i-1} \), for \( i = 1, 2, ..., N \), which would be equivalent to a Kuhn segment [1, 14].

Moreover, we thoroughly describe the chains morphology exploring the behavior of two characteristic distances, the end-to-end distances \( R_{ee} \) and the radius of gyration \( R_g \). The end-to-end distance is defined as the mean squared variance of the displacement and is reads as,

\[
R_{ee}^2 = \langle (r_N - r_0)^2 \rangle, \tag{2}
\]

the variables \( r_N \) and \( r_0 \) are the positions of the ends of the chain. The radius of gyration \( R_g \), whose square is the second moment around the center of mass \( r_c \) given by:

\[
r_c = \frac{1}{N+1} \sum_{i=0}^{N} r_i, \tag{3}
\]

so that, the radius of gyration takes the form [15, 14]:

\[
R_g^2 = \frac{1}{N+1} \sum_{i=0}^{N} \langle (r_i - r_c)^2 \rangle. \tag{4}
\]

For real (non-Gaussian, with excluded volume) chains, a relationship between these distances is [15]:

\[
\frac{6R_g^2}{R_{ee}^2} = 0.952. \tag{5}
\]

For this type of chain, the radius of gyration depends on \( N \), size of the Kuhn segment \( b \) and exponent \( \nu \): 

\[
R_g = bN^\nu. \tag{6}
\]

The exponent \( \nu \) is called the size exponent or the Flory exponent [1, 14, 16], which comes from Flory’s theory, which had remarkable success in explaining the experimental evidence in the swelling of real polymers.

2.3. Energy of linear chains

In our analysis, we use a definition of the interaction energy, which is based on the compactness of the chains [17, 18, 19]. Complementary, we also examine the bending energy of the chains, which in our case is a discrete variable. This magnitude characterizes the flexibility of the chains, as well as, its tangential correlations. Furthermore, the total energy of the chain is the sum of the bending energy and the interaction energy.
2.3.1. Interaction Energy Term

Interaction energy accounts for the energy of the chain due to its compactness and it quantifies the short-range interactions (Von Neumann neighborhood) for non-continuous monomers [17, 18, 19, 20]. It reads as,

\[ E = \sum_{i<j} e_{\mathbf{r}_i \cdot \mathbf{r}_j} \Delta (\mathbf{r}_i - \mathbf{r}_j), \]

where \( \Delta (\mathbf{r}_i - \mathbf{r}_j) = 1 \) if \( \mathbf{r}_i \) and \( \mathbf{r}_j \) are attached to the network, but \( i \) and \( j \) are not adjacent positions along the chain sequence and \( \Delta (\mathbf{r}_i - \mathbf{r}_j) = 0 \) for otherwise (see Fig. 1). The value of the factor, \( e_{\mathbf{r}_i \cdot \mathbf{r}_j} \), depends on the type of contact between the monomers and represents the potential energy of interaction between the monomers located in the position \( \mathbf{r}_i \) and \( \mathbf{r}_j \) respectively. In our case, it is \( e_{\mathbf{r}_i \cdot \mathbf{r}_j} = -1 \) because we consider an attractive monomer-monomer interaction in a homopolymeric linear chain.

![Figure 1: 14-step 2d-chain that shows the values that the interaction takes \( \Delta (\mathbf{r}_i - \mathbf{r}_j) \), where it adopts the values: 1, for the attachments but not adjacent sites (blue color) and 0 for the otherwise.](image)

2.3.2. Bending Energy

One of the basic characteristics of all macromolecules is their flexibility [21]. The polymer chains, in the pure state or in a dissolution, may adopt different conformations depending on their flexibility. When the flexibility is high, the chain may have large changes of direction within a few links. On the contrary, if the flexibility is low, the chain will be more rigid and will tend, in the limit, to behave as a hard stick. The flexibility of the polymer chain is related to the persistence length \( l_p \). This can be defined as the average value of the maximum linear length of the chain configuration (it is also related to the Kuhn segment as \( b = 2l_p \) [14]). At distances greater than \( l_p \), fluctuations in relation to itself or to the surroundings destroys the memory bound to the direction of the chain. Thus, the polymers are not completely flexible, some energy is required to fold them, which can happen up to at most the \( l_p \).

The correlation (in turn related to the flexibility) between \( \mathbf{u} \) and \( \mathbf{u}' \), two unit vectors that join three points of the chain (monomers in the Fig. 2a) and that are separated by a distance \( l \) is given by [15]:

\[ \langle \mathbf{u} \cdot \mathbf{u}' \rangle = [(1 - (b/l_p)^{1/2})^{1/2}] \exp (-l/l_p). \]
At the limit, the conformation of the chain is a smooth curve as described in Fig. 2b. Using Eq. (8) can obtain the correlation function between \( \mathbf{u}(s) \) and \( \mathbf{u}(s') \) of two segments of the chain, \( s \) and \( s' \), as a function of the persistence length \( l_p \), given by [15]:

\[
c(s, s') = \langle \mathbf{u}(s) \cdot \mathbf{u}(s') \rangle = \exp\left(-\frac{|s - s'|}{l_p}\right),
\]

this shows that the directional correlation of two segments of a macromolecule, decreases exponentially with the growth of the chain length [21, 22].

For bending energy we propose the following: we consider our polymer chain taking into account the interactions with other monomers of the same chain and the interactions with its environment. These interactions are described by an effective potential that represent the energy cost for its formation and whose stability is determined by two forces, one elastic with negative signal, that leads the chain to a collapse, and another repulsive of positive signal, that makes the chain is stretched. This energy cost is reflected in the chain in the form of free energy, for example, the number of conformations decreases with the increase of the vector end-to-end but increases the free energy of the same due to the high correlation that exists in the chain.

Chains with high flexibility experience changes of direction at a distance of few links tending to turn to itself, while low flexibility chains tend to become rigid, because the two-segment correlation function in the chain decreases exponentially with the distance between them as shown in Eq. (9). This correlated behavior occurs in the same way in a system of continuous mechanics, for the flexion model due to a force acting on a thin rod with stiffness constant \( k \) [23, 24]. Flexing generates a differentiable curve on the rod, where, at a point \( \mathbf{r}(s) \) of the curve there is a tangent vector \( \mathbf{u}(s) \) generating a behavior similar to that described in Fig. 2b for polymers. The Hamiltonian describing the internal energy of the rod of length \( l_c \) is given by,

\[
\mathcal{H} = \frac{k}{2} \int_0^{l_c} \left( \frac{\partial \mathbf{u}(s)}{\partial s} \right)^2 ds.
\]

By virtue of this, we propose a discretization of bending energy in order to be adapted to our chains, in this way, part of the internal energy of the chain which is described in terms of its
configuration and which is equivalent to Eq. (10) is given by the following relation:

\[ \mathcal{H} \simeq H = \frac{k}{b} \sum_{i,j=1}^{N} \varepsilon_{ij}p_i, \]  

(11)

where the weight function \( \varepsilon_{ij} \) can take values of (1) or (-1) depending on whether or not the direction of the \( i \)-th step changes as compared to the previous step (see Fig. 3) and \( p_i \) represents the probability that each step will find any of its accessible microstates (Eq. (1)), \( k \) is a constant of units of energy times distance and finally, our model adopts the \( b = 1 \) as the length of Kuhn.

Figure 3: 4-step 2d-chain that shows the value of the weight function \( \varepsilon_i \) used to compute the bending energy. Left: linear chain without deviation with its weight per step equal to (1). Right: chain with mixed deviations, when direction changes \( \varepsilon_i \) adopts a weight equal to (-1), as is the case with steps 2 and 4.

The bending energy, in its discrete version \( H \), describes the behavior of the polymer chain from its tangential correlations, for example, for a highly correlated polymer chain (Fig. 3, left) this energy will be purely positive (high free energy) what is expected from the Hamiltonian described by the rod and the correlation given in Eq. (9).

In our simulations the calculation of bending energy takes into account the term we call function weight as well as the relative probability of each step of the chain, which has the form of Eq. (1). This energy can be positive or negative depending on the winding of the chain.

3. RESULTS AND DISCUSSION

In order to obtain the results, eleven thousand three-dimensional chains was generated from 2400 attempts (\( N' = 2400 \)).

3.1. Chain Length, End-to-end Distance and Radius of Gyration

The behavior of the distribution of the number of steps for three-dimensional chains is shown in Fig. 4. The maximum of the distribution, is obtained for large chains (compared to \( N' \)). The maximum corresponds to \( N = 1775 \) steps for the chains generated from \( N' = 2400 \). This way, chains with approximately this number of steps appear with greater probability under the conditions established in the simulation.

In Fig. 5 and Fig. 6, the distribution of the characteristic distances (\( R_{se} \) and \( R_g \)) of the chains is shown. The graphs in Fig. 5 represent adjustments using Lhuillier’s proposal [25]:

\[ P(R) \sim \exp (-R^{-ad} - R^d). \]

(12)
The distribution behavior can be described separated in two regions that follow different exponential laws. For small values of characteristic distances, the distribution behavior is described by: $\sim \exp(-R^{-\alpha d})$, where $\alpha = (\nu d - 1)^{-1}$ and $d$ is a spatial dimension. For large values of the characteristic distances, the expression for the distribution is: $\exp(-R^\delta)$, $\delta = (1 - \nu)^{-1}$ is the Fisher exponent [26].

Figure 4: Normalized histogram of the number of steps $N$ with mean $\mu = 1775.42$ and standard deviation $\sigma = 33.10$. Behavior obtained for random chains in 3d generated from $N' = 2400$. The inset illustrates the histogram in a wide interval in which a Gaussian behavior occurs.

Figure 5: Normalized histograms of the radius of gyration $R_g$ and end-to-end distance $R_{ee}$. Behavior obtained for the ensemble of random chains generated from $N' = 2400$ in 3d. It is observed that a distribution agrees very well with the expression derived by Lhuillier [25, 26], both, the $R_g$ and $R_{ee}$ distribution.
The Fig. 6 shows the graphs of the $R_{ee}$ and $R_g$ distribution, adjusted according to the Fisher-McKenzie-Moore-des Cloizeaux law [14, 27, 28, 29]. This law is commonly used to describe the distribution of extreme-extreme distance and is of the following type:

$$P(R) \sim R^\theta \exp(-R^\delta).$$  \hspace{1cm} (13)

The exponent $\theta = (\gamma - 1)/\nu$ characterizes the shorts-distance intra-chain correlations between two segments of a long polymer in a good solvent. The total number of chain conformations is indirectly determined by the exponent $\gamma$. As opposed to ideal chains, where $\gamma = 1$, for real chains, $\gamma > 1$, i.e. there is a reduction of the probabilities in the distribution of $R$ for short chains.

![Figure 6: Normalized histograms of the radius of gyration $R_g$ and end-to-end distance $R_{ee}$. Behavior obtained for the ensemble of random chains generated from $N' = 2400$ in $3d$. It is observed that a distribution agrees very well with the function proposed by McKenzie and Moore [27] and des Cloizeaux [28, 29], both, the $R_g$ and $R_{ee}$ distribution.](image)

It is important to note that while Fisher-McKenzie-Moore-des Cloizeaux theoretical distribution parameters are contracted by fixing the number of steps of the chains and study the distance end-to-end of the different configurations generated, in our approach we generate a random ensemble of chains without fixing a priori the number of steps. However we have been demonstrated that the chains obtained follow the same distribution [14, 28].

The value of Flory exponent $\nu$, which describes the size of the polymer chain, was calculated by computing the mean value of the radius of gyration. The values of $\nu$ obtained from the simulation are incorporated in Fig. 7 and Fig. 8. These results closely approximate the expected theoretical value for the Flory exponent that is $\nu \approx 0.59$ for the three-dimensional case. The calculation of $\nu$ from the behavior of $R_g$ results in a value closer to the expected theoretical value than that calculated from the $R_{ee}$. 

9
The Flory exponent was also calculated indirectly from the delta ($\delta$) exponent that results from the distribution of characteristic distances, shown in Fig. 5 and Fig. 6. The values of the alpha ($\alpha$) and theta ($\theta$) exponents, the first indirectly and the second directly, were also obtained from these distributions.

Figure 7: Log-log scale representation of the end-to-end distance ($R_{ee}$) as a function of the number of steps ($N$) with its respective value of $\nu$.

Figure 8: Log-log scale representation of the radius of gyration ($R_g$) as a function of the number of steps ($N$) with its respective value of $\nu$. 

10
The Table 1 shows a comparison between the main critical exponents reported by various authors, both analytically and numerically, and the values reported by our simulations. The above results lead us to conclude that although the two behaviors, $R_{ee}$ and $R_g$, as a function of $N$ fit a power law, the value of the characteristic exponent of $R_g$ is closer to the theoretical value of Flory ($\nu = 0.60$ for 3d chains) and to the values reported in [14, 30, 31]($\nu = 0.588$).

Although both proposed functions fit well the distributions ($R_{ee}$ and $R_g$), the values of the critical exponents obtained for each function, specifically the delta exponent, are different and correspond to different laws. The $R_g$ distribution responds to the law proposed by Lhuillier and the distribution of $R_{ee}$, to the Fisher-McKenzie-Moore-des Cloiseaux law, as already reported in the literature.

| Author                  | $\nu$    | $\delta$ | $\alpha$ | $\theta$ |
|-------------------------|----------|----------|----------|----------|
| Rubinstein [14]         | 0.588    | 2.43     | 1.31     | 0.28     |
| Caracciolo et al. [30]  | 0.58758 ± 0.00007 | 2.4247 ± 0.0004 | 1.311 | 0.2680 ± 0.0011 |
| Vectoriel et al. [31]   | 0.588    | 2.38     | 1.31     |         |
| $R_g$ Vs $N$            | 0.578 ± 0.008 | -        | 1.36     |         |
| $R_{ee}$ Vs $N$         | 0.631 ± 0.011 | -        | 1.12     |         |
| $R_g$ (Lhuillier)       | 0.606    | 2.54     | 1.22     |         |
| $R_{ee}$ (Lhuillier)    | 0.724    | 3.63     | 0.853    |         |
| $R_g$ (DC)              | 0.225    | 1.29     | -        | 2.54     |
| $R_{ee}$ (DC)           | 0.574    | 2.35     | 1.295    | 0.91     |

3.2. Energy of the Polymeric Chains

To study the energy of the random polymer chains generated by our simulation, we analyzed the energy given by the interactions between monomer in each chain, as well as the energy associated with the bending of the chains. As generated chains are homopolymeric, they are expected to show a uniform behavior or have a low amount of metastable states [18, 32].

In the contact potential described in Eq. (7), the value of the constant $e_{inuj}$ has the information about of interaction energy between the non-continuous and adjacent monomers. This constant adopts the value of $-1$, for each contact, thus generating a “folding” force in the chain, known in proteins as a hydrophobic force [19]. The sum of all the interactions in chain defines the energy of the system, called interaction energy ($E$).
To take into account the flexibility of the chain and, consequently, its tangent correlations, we
proposed adding to the interaction term $E$, the bending energy ($H$) in its discrete version, proposed
in Eq. (11).

In the Fig. 9 we can see the characteristic histograms of each energy for three-dimensional
chains generated using our algorithm. The shape of the distribution is similar to the results ob-
tained previously for the distribution of number of steps.

![Histograms of Interaction and Bending Energies](image)

Figure 9: Interaction energy $E$ (left side of figure) and bending energy $H$ (right side) histograms, for three-dimensional chains. Gaussian distributions with standard deviations, $\sigma_E = 50.66$ and $\sigma_H = 30.52$, and mean values $\mu_E = -701.60$, and $\mu_H = -791.77$ respectively.

One of the most important results obtained here is that, even when considering both attractive
and repulsive behavior of the chains, the resulting chains have negative energy, because in the SAW
models for polymers in a good solvent, the attractions prevail over the repulsions [14, 15].

The total energy of the chain is described by the sum of interaction and bending energies, which
determine the structural configuration. In the Fig. 10, the total energy distribution is shown. The
most probable energy values are between $-1650$ and $-1350$ energy units, corresponding to long
chains. The drop in probability density at the tail of the distribution at energy values approaching
zero correspond to “small” chains. These small chains can also be identified in the in the probal-
ity density function of $N$, which shows the same drop on the left tail of the distribution. This
may be due to the fact that the total number of configurations, for real chains, is smaller than
for ideal chains because of the reduction of probabilities in the distributions of the characteristic
distances, mentioned above, for $R_{ee} \lesssim 40$ and $R_g \lesssim 10$ (see Fig. 5), which corresponds to $N \lesssim 500$.

Following the same procedure used to study the Flory exponent, the mean values of interaction
energy and bending energy were computed for eleven thousand three-dimensional chains. The
average energy was plotted as a function of $N$ (Fig. 11). The plot shows a linear and uniform
behavior for the two energy (which is expected [18, 32]). When $N$ is large, the behavior of the
interaction energy changes from decreasing to increasing, having less negative values, which can be
interpreted as representative of more stretched chains, with fewer contacts but with small lengths
of persistence still prevailing.

Figure 10: Total Energy histogram \((E + H)\) for three-dimensional chains with mean \(\mu = -1499.14\) and standard deviation \(\sigma = 53.24\).

Figure 11: Interaction energy \(E\) and bending energy \(H\) as a function of the number of steps \(N\). The horizontal dotted lines represent the values between which the energies are distributed (see Fig. 10). On the other hand, the vertical ones, represent the values between which distribute all sizes of the three-dimensional chains generated.

The great majority of the three-dimensional chains generated are large \((N > 1700)\). These chains, with energies between \(-1350\) and \(-1650\), belong to the second regime of the graph of the total energy and follow a Gaussian distribution (see Fig. 10). The rest of the chains, with energy between 0 and \(-1350\), belong to the first regime of the energy graph and follow a uniform distribution.
In the “discrete” bending energy approach, the second regime does not appear (see Fig. 11). This term has the same linear behavior, independent of the presence of neighbors interacting with other monomers. The folding of the chains in the large \( N \) regime adds negative energy to the system. The graph of Fig. 11 shows a lower density of points for energies between \(-450\) and \(0\) (corresponding to \(0 \lesssim N \lesssim 500\)) confirming the decrease in the number of configurations for small chains, as shown in the dot density of Fig. 7 and Fig. 8 for \(N \lesssim 500\).

The Fig. 12 a typical linear polymeric chain of \(N = 1772\) generated by the simulation is illustrated. The generated chain forms two clusters separated by a tail. The interaction energy decreases because the chain “escapes” and does not have neighbors that contribute to this energy. On the other hand, the bending energy remains constant, since the flexibility of the chain is maintained and causes the chain to continue folding even in the “bridge” that separates the clusters.

![Figure 12: Typical structure of \(N = 1772\) generated by the simulation which shows the formation of two clusters. The spheres represent the monomers of a homopolymeric chain.](image)

In the total energy graph of Fig. 11, two regimes can also be seen due to the contribution of the interaction term. For chains of size below \(N = 1775\) (mean value of \(N\)), the energy decreases with \(N\) at a rate of 0.86, practically twice the rate of decrease of each term separately, because the two energy contributes in this regime in the same proportion. For larger chains, the energy increases with \(N\) and the characteristic rate of this increase is 0.29. Here the different contributions of the two types of energy make the increase less significant than in the case of the interaction energy.
4. CONCLUSIONS

In this work, an algorithm based on natural self-avoiding random walk in a cubic network with no boundary was used to generate linear chains. The implementation of this algorithm is simpler and its computational cost is significantly lower than other approaches generally used to generate self-repelling chains. Although our method does not consider all possible configurations for $N$ size of the chain we obtain values of the characteristic critical exponent’s of the biopolymers, similar to those reported in literature. The distribution of the characteristic distances for the linear chains obtained from our simulation showed a reasonable correspondence with that reported in the literature. Specifically, we prove that although it is possible to fit the two distributions ($R_g$ and $R_{ee}$) using both Lhuillier’s theory and the law proposed by Fisher-McKenzie-Moore-des Cloiseaux, the obtained values of critical exponents such as $\delta$, $\alpha$, $\nu$ show us that the distribution of $R_g$ responds better to the theory of Lhuillier, while in the case of the distribution of $R_{ee}$, the best fit is obtained with the Fisher-McKenzie-Moore-des Cloizeaux function. These results are in full agreement with the literature and reinforce the validity of the algorithm used when characterizing this type of system.

Both $R_{ee}$ and $R_g$ resulted in power functions of $N$ and the values of the Flory exponent, in both cases, are quite close to the theoretical value, especially the value corresponding to $R_g$, showing that this characteristic distance is more appropriate when characterizing structurally this type of chains. By comparing the behavior with $N$ of these two distances, we can validate the theoretical linear dependence between these two parameters in the case of real chains.

Our study revealed interesting behaviors related to the flexibility of the chains. Small chains are more correlated and consequently, less flexible. Medium-sized chains with $N$ values below 1650 behave uniformly, being more flexible as $N$ increases.

Comparing the behavior of the energy $E$, related to the monomer-monomer interactions of the chains, with the energy $H$, related to its flexibility, we found that chains start with clusters and doubles over distances of the order of $R_g$, increasing its energy. When the chains reach a large number of steps, the interaction energy revealed that the chain stretches and escapes from the cluster, which results in a loss of interaction energy. The behavior of the bending energy reveals that in this escape regime, the chains keep the folding behavior they showed before the escape. This process lasts until the clusters reach a size comparable with the mean radius of gyration.

The analysis presented in this work, including both the bending and interaction energies is important because it allows differentiating between chains of equal interaction energy but with different structures and hence different bending energies.

Acknowledgments

This work has received financial support from CAPES and CNPq (Brazilian Federal Grant Agencies) and from FACEPE (Pernambuco State Grant Agency). We thank Pedro Hugo Figueirêdo, Raúl Cruz Hidalgo and Juan Miguel Parra Robles for useful comments of this work.
References

[1] P. J. Flory, Principles Of Polymer Chemistry, Cornell University Press, Ithaca, United States, 1953.

[2] N. Madras, A. D. Sokal, The pivot algorithm: A highly efficient monte carlo method for the self-avoiding walk, Journal of Statistical Physics 50 (1988) 109. doi:10.1007/BF01022990.

URL https://doi.org/10.1007/BF01022990

[3] S. Gordon, Self-avoiding walks, The Mathematical Intelligencer 16 N1 (1994) 29.

URL https://www.imsc.res.in/~sitabhra/teaching/asm17/Slade_MathIntelligencer_1994_Self_avoiding_walks.pdf

[4] H. E. Stanley, Dependence of critical properties on dimensionality of spins, Phys. Rev. Lett. 20 (1968) 589. doi:10.1103/PhysRevLett.20.589.

URL https://link.aps.org/doi/10.1103/PhysRevLett.20.589

[5] P. de Gennes, Exponents for the excluded volume problem as derived by the wilson method, Physics Letters A 38 (5) (1972) 339. doi:https://doi.org/10.1016/0375-9601(72)90149-1.

URL http://www.sciencedirect.com/science/article/pii/0375960172901491

[6] N. Madras, G. Slade, The Self-Avoiding Walk, Birkhauser , Boston-Basel-Berlin, 1993.

[7] H. Yamakawa, Modern Theory of Polymer Solutions, New York: Harper and Row, 1971.

[8] K. G. Wilson, J. Kogut, The renormalization group and the $\varepsilon$ expansion, Physics Reports 12 (2) (1974) 75. doi:https://doi.org/10.1016/0370-1573(74)90023-4.

URL http://www.sciencedirect.com/science/article/pii/0370157374900234

[9] A. Sokal, Molecular Dynamics Simulations in Polymer Science, Oxford University Press, 1995.

[10] A. Guttmann, A. Conway, Square lattice self-avoiding walks and polygons, Annals of Combinatorics 5 (3) (2001) 319. doi:10.1007/PL00013842.

URL https://doi.org/10.1007/PL00013842

[11] I. Jensen, Enumeration of self-avoiding walks on the square lattice, Journal of Physics A: Mathematical and General 37 (21) (2004) 5503.

URL http://stacks.iop.org/0305-4470/37/i=21/a=002

[12] B. Li, M. Neal, A. D. Sokal, Critical exponents, hyperscaling, and universal amplitude ratios for two and three-dimensional self-avoiding walks, Journal of Statistical Physics 80 (3) (1995) 661. doi:10.1007/BF02178552.

URL https://link.aps.org/doi/10.1007/BF02178552

[13] D. J. Amit, G. Parisi, L. Peliti, Asymptotic behavior of the "true" self-avoiding walk, Phys. Rev. B 27 (1983) 1635–1645. doi:10.1103/PhysRevB.27.1635.

URL https://link.aps.org/doi/10.1103/PhysRevB.27.1635

[14] M. Rubinstein, R. H. Colby, Polymer Physics, New York: Oxford University Press, 2003.

[15] I. Teraoka, Polymer Solutions, An Introduction to Physical Properties, New York: Wiley-Interscience, 2002.

[16] S. M. Bhattacharjee, A. Giacometti, A. Maritan, Flory theory for polymers, Journal of Physics: Condensed Matter 25 (50) (2013) 503101.

URL http://stacks.iop.org/0953-8984/25/i=50/a=503101

[17] P. H. Figueiredo, M. A. Moret, S. Coutinho, J. Nogueira, The role of stochasticity on compactness of the native state of protein peptide backbone, The Journal of chemical physics 133 (2010) 1. doi:10.1063/1.3481485.

URL https://doi.org/10.1063/1.3481485

[18] R. A. Broglia, G. Tiana, D. Provasi, Simple models of the protein folding and of non-conventional drug design, Journal of Physics: Condensed Matter Vol 16 (6) (2004) 111.

URL http://stacks.iop.org/0953-8984/16/i=6/a=R02

[19] C. Tang, Simple models of the protein folding problem, Physica A: Statistical Mechanics and its Applications 288 (1) (2000) 31 – 48. doi:https://doi.org/10.1016/S0378-4371(00)00413-1.

URL http://www.sciencedirect.com/science/article/pii/S0378437100004131

[20] S. P. Vijay, Y. G. Alexander, T. Toyoiuchi, Statistical mechanics of simple models of protein folding and design, Biophysical Journal 73 (1997) 3192. doi:https://doi.org/10.1016/S0006-3495(97)73435-0.

URL http://www.sciencedirect.com/science/article/pii/S0006349597783450

[21] A. Y. Grosberg, A. R. Khokhlov, Statistical Physics of Macromolecules, New York: AIP Press, 1994.

[22] H. Hsiao-Ping, P. Wolfgang, K. B., Standard definitions of persistence length do not describe the local “intrinsic” stiffness of real polymer chains, Macromolecules 43 (2010) 3094. doi:10.1021/ma902715e.

URL https://doi.org/10.1021/ma902715e

[23] L. L. L. E., Theory of Elasticity, Oxford Press: Pergaman, 1986.
[24] S. Schöbl, S. Sturm, W. Janke, K. Kroy, Persistence-length renormalization of polymers in a crowded environment of hard disks, Physical Review Letters 113 (2014) 238302–1. doi:10.1103/PhysRevLett.113.238302.

[25] D. Lhuillier, A simple model for polymeric fractals in a good solvent and an improved version of the flory approximation, J. Phys. France 49 (5) (1988) 705–710. doi:10.1051/jphys:01988004905070500. URL https://doi.org/10.1051/jphys:01988004905070500

[26] J. M. Victor, D. Lhuillier, The gyration radius distribution of twodimensional polymer chains in a good solvent, The Journal of Chemical Physics 92 (2) (1990) 1362–1364. arXiv:https://doi.org/10.1063/1.458147, doi:10.1063/1.458147. URL https://doi.org/10.1063/1.458147

[27] D. S. McKenzie, M. A. Moore, Shape of self-avoiding walk or polymer chain, Journal of Physics A: General Physics 4 (5) (1971) L82–L85. doi:10.1088/0305-4470/4/5/017. URL https://doi.org/10.1088/0305-4470/4/5/017

[28] J. Des Cloizeaux, Lagrangian theory for a self-avoiding random chain, Phys. Rev. A 10 (1974) 1665–1669. doi:10.1103/PhysRevA.10.1665. URL https://link.aps.org/doi/10.1103/PhysRevA.10.1665

[29] J. Cloizeaux, G. Jannink, Polymers in solution: their modelling and structure, Oxford science publications, Clarendon Press, 1990.

[30] S. Caracciolo, M. S. Causo, A. Pelissetto, End-to-end distribution function for dilute polymers, The Journal of Chemical Physics 112 (17) (2000) 7693. doi:10.1063/1.481362. URL https://doi.org/10.1063/1.481362

[31] T. Vettorel, G. Besold, K. Kremer, Fluctuating soft-sphere approach to coarse-graining of polymer models, Soft Matter 6 (2010) 2282–2292. doi:10.1039/B921159D. URL http://dx.doi.org/10.1039/B921159D

[32] N. V. Dokholyan, S. V. Buldyrev, H. E. Stanley, E. I. Shakhnovich, Discrete molecular dynamics studies of the folding of a protein-like model, Folding and Design 3 (6) (1998) 577. doi:https://doi.org/10.1016/S1359-0278(98)00072-8. URL http://www.sciencedirect.com/science/article/pii/S1359027898000728