entropy

MDPI

Article

Complex Networks and Interacting Particle Systems

Noam Abadi * and Franco Ruzzenenti

Integrated Research on Energy, Environment and Society (IREES), Energy and Sustainability Research Institute Groningen (ESRIG), University of Groningen, Nijenborgh 6, 9747 AG Groningen, The Netherlands; f.ruzzenenti@rug.nl

* Correspondence: n.abadi@rug.nl

Abstract: Complex networks is a growing discipline aimed at understanding large interacting systems. One of its goals is to establish a relation between the interactions of a system and the networks structure that emerges. Taking a Lennard-Jones particle system as an example, we show that when interactions are governed by a potential, the notion of structure given by the physical arrangement of the interacting particles can be interpreted as a binary approximation to the interaction potential. This approximation simplifies the calculation of the partition function of the system and allows to study the stability of the interaction structure. We compare simulated results with those from the approximated partition function and show how the network and system perspective complement each other. With this, we draw a direct connection between the interactions of a molecular system and the network structure it forms and assess the degree to which it describes the system. We conclude by discussing the advantages and limitations of this method for weighted networks, as well as how this concept might be extended to more general systems.

Keywords: complex networks; interacting systems; maximum entropy; statistical physics; Lennard-Jones

check for updates

Citation: Abadi, N.; Ruzzenenti, F. Complex Networks and Interacting Particle Systems. *Entropy* **2023**, *25*, 1490. https://doi.org/10.3390/e25111490

Academic Editor: Leonid M. Martyushev

Received: 18 September 2023
Revised: 20 October 2023
Accepted: 26 October 2023
Published: 27 October 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The field of network theory is a growing branch of science aimed at representing and understanding interacting systems. The main notion is that the system can be represented by listing its components, known as nodes, and the connections between each pair of components, called links [1,2]. These simple requirements allow for the description and common study of a wide range of phenomena, generalized by studying the structural properties defined on the nodes and links of the network representation instead of on the system itself [3]. Structural properties range from simply the number of nodes and links to fractal dimensions related to notions of space occupied by the network [4].

Interacting particle systems serve, in many senses, as a reference case for network concepts [5] due to the natural representation that they enjoy as networks (in the form of nodes that represent particles connected by edges that represent a stable equilibrium distance) [6–8]. This representation inherently shares well-defined notions of nodes (particles) and an intrinsic relation between interactions and structure (e.g., molecular forces mediated by physical distance) grounded on physical concepts and experimental evidence. It has seen applications ranging from the study of molecule recognition [9] to self-organization [10] and to the notion of complexity [11,12]. In the field of complex networks theory, it has become increasingly clear that a better understanding of the relation between network structure and the interactions of the system would be useful to both the application and the framework [13,14]. Additionally, network structure formation models often draw on the notion of structure through link creation and destruction [15,16], which is natural to molecular systems in the same sense that the network representation of the spatial configuration is, and it provides a very intuitive view of the formation process. This motivates a positive outlook for complex networks providing a general view on patterns observed between interacting systems at different scales [17].

On the other hand, within fixed scales (in particular, at the molecular scale), network analyses of structure tend to either take data-analytic approaches [18,19] or impose external "control parameters" [8] built on the notion of adaptive links [20]. In this context, a change in temperature which allows certain molecular structures to emerge would be modeled by setting an external parameter which regulates how the links are created and destroyed. The problem with this is that the emergent phenomenon (in this case, simply stable links in the network) is imposed instead of actually obtained from the interaction. A question that then naturally arises is whether it is possible to construct an "environment-independent" network model of the particle system which does not consider the links as adaptive and still learns something about structure formation.

In this work, we study the emergence of a molecular structure in a simulated Lennard-Jones particle system in thermodynamic equilibrium by representing it as a complex network. In particular, we remove the standard restriction of molecular graph representations depicting only stable links and interpret these as binary approximations of the interaction potential. The goal of this is to free the network representation of details beyond the interaction itself, as these can change according to the environment or definitions of stability. We study the structural properties of the resulting network in equilibrium at different temperatures and pressures and show how the binary approximation of the potential allows the use of network and thermodynamic properties in combination, in particular to detect where the studied network structure is stable in thermodynamic equilibrium.

Particle systems in thermodynamic equilibrium are subject to description by maximum entropy. One of the main results [21] is that such systems can be described by a partition function, which is essentially a function $\mathcal{Z}$ of a parameter β from which average values (such as the average kinetic and potential energy) in equilibrium can be calculated. For a system of particles $1 \leq i \leq N$ with kinetic energy $mv_i^2/2$ and potential energy $V_{ij}(r_i, r_j)$. The canonical partition function is calculated as

$$\mathcal{Z}(\beta) = \int \exp(-\beta(mv_i^2/2 + \sum_i \sum_{j>i} V_{ij}(r_i, r_j)))\Pi_{i=1}^N dv_i dr_i \,. \tag{1}$$

As the potential energy is independent of the velocity, the integrations over velocities can be separated from the spatial term. The former gives the partition function of an N classical particle system with no interactions, $\mathcal{Z}_v = (\int \exp(-\beta m v^2/2) dv)^N$. Assuming unit mass and two dimensions, this will result in a (kinetic) energy $K = 1/\beta$ with $\beta = 1/T$ representing the temperature of a noninteracting gas with that amount of (kinetic) energy, and the corresponding distribution of velocities will be Maxwell–Boltzmann. However, the spatial part is, in general, very difficult to calculate. We show that the network representation introduced in this work simplifies this task. Finally, we remark that while maximum entropy is well-established for thermodynamic equilibrium, the field of maximum entropy production [22,23] is a developing counterpart for out-of-equilibrium systems through which our approach might be extended.

2. Methods and Experiments

2.1. Lennard-Jones Potential

The Lennard-Jones potential is a useful, well-studied idealization of the interactions between two atoms at a molecular scale. In this work, the potential is written as

$$\begin{aligned} V(r) &= V_o \left[\left(\frac{r_o}{r}\right)^{12} - 2\left(\frac{r_o}{r}\right)^{6} \right] \\ \Rightarrow \vec{F}(r) &= -\vec{\nabla} V(r) = 12\frac{V_o}{r_o}\left[\left(\frac{r_o}{r}\right)^{13} - \left(\frac{r_o}{r}\right)^{7} \right] \hat{r} \,. \end{aligned} \tag{2}$$

For two particles, a short distance $r < r_o$ between them produces an enormous repulsive force. At long distances $r > r_o$, it is very weakly attractive, and in the middle lies the equilibrium point $r = r_o$, as shown in Figure 1. This allows for particles with low

kinetic energy (below V_o) to "bind" together. If the kinetic energy is much larger than V_o, the energy decrease close to r_o is practically unnoticed, but a minimum distance is eventually reached at which the repulsion forces win over and the particles bounce back away from each other. These two limits represent low and high temperature, respectively, and between them lie a diversity of possibilities which give rise to different network structures.

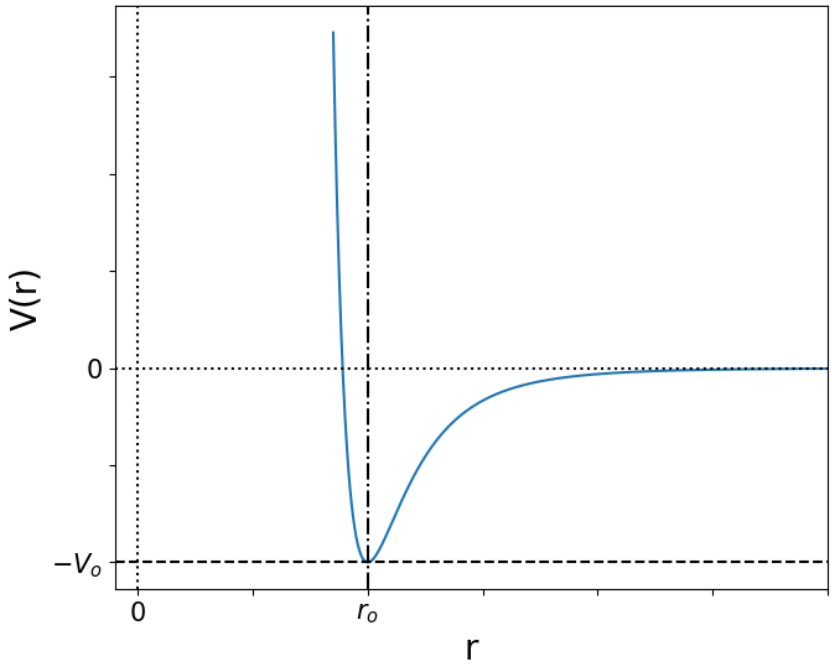

Figure 1. Potential energy as a function of distance in the Lennard-Jones potential in Equation (2). For kinetic energies below the minimum value V_o, particles will remain at a distance of approximately r_o.

For the particle simulation, we took the Lennard-Jones potential parameters to be $V_o = 1$ and $r_o = 1$. $N = 900$ particles of mass 1 were set on a uniform 2-dimensional grid with cells of side d with random velocities, uniformly distributed between $-V_M$ and V_M. The total force on each particle due to its interactions with others allows for the numerical integration of the equations of motion of each particle using a leap-frog algorithm [24]. Initial conditions varied were d (between 1 and 4), in order to control initial potential energy through particle separation, and V_M (between 0.1 and 5), in order to control initial kinetic energy. The size of time steps and total simulation time were chosen in order to ensure energy conservation and at least 3/4 of simulation time in equilibrium (constant proportions of kinetic and potential energy). Additionally, a set of "cooled" conditions were obtained, continuing the evolution of the last step of simulations started with $V = 0.1$ after multiplying their velocities at this step by 0.5 (effectively cooling the system). Once these reached equilibrium and stayed there for at least 3/4 of the simulation time, the process was repeated with the last step of the new simulation. This was performed four times for each initial separation.

2.2. Network Construction and Stability

As discussed in the introduction, we aim to construct a network representation that captures the interactions between the particles, without relying on the particular notion of equilibrium (which in this case is thermodynamic). For this, we first consider an approximation of the interaction potential as a step function of unspecified heights V_a and V_b, with the step $\Delta V = V_a - V_b$ at a separation radius r_T, that is, $V(|r_i - r_j|) \approx V_a\Theta(r_T - |r_i - r_j|) + V_b\Theta(|r_i - r_j| - r_T)$. One might protest that this is not a suitable representation, as it does not allow for repulsive potentials. However, we argue that one can imagine that a real collision is replaced by a virtual information exchange, where the

particles exchange their properties, as shown in Figure 2. This replacement is valid for the regions far away from the interaction range (i.e., it is valid where the particle trajectories are approximately straight lines, which in a two-level potential would be the outer range $|r_i - r_j| > r_T$), while for the short range, it is equivalent to projecting the particle velocities onto their initial unitized velocity vectors.

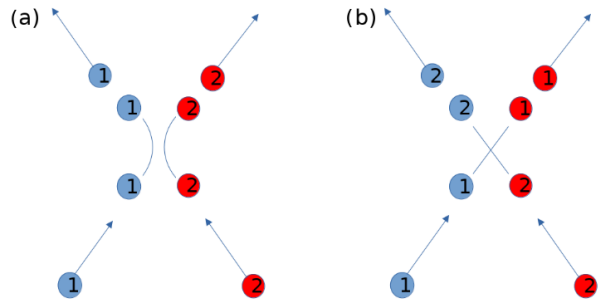

Figure 2. A repulsive interaction between two particles as shown in (**a**), can be replaced by an information exchange where each particle acquires the properties of the other, following straight trajectories, as shown in (**b**). We set a "name" (1 and 2) in order to track the particles in both scenarios, but the name is not a property as it is not exchanged.

With the approximation, the inner radius is then considered as the "active" region of the interaction between different particles i and j, defining a binary undirected network

$$A_{ij} = (1 - \delta_{ij})\Theta(r_T - |r_i - r_j|)\,. \tag{3}$$

From it, we study the degree k_i, which (in the two-level approximation of the potential) represents the number of particles with which i is interacting with

$$k_i := \sum_j A_{ij}\,, \tag{4}$$

assortativity a, the correlation between the number of neighbors of a randomly selected pair of connected nodes,

$$\begin{aligned} L &= \sum_i k_i \\ \alpha &= \sum_i k_i^2 / L \\ a &:= \frac{\sum_i k_i \sum_j A_{ij} k_j / L - \alpha^2}{\sum_i k_i^3 / L - \alpha^2} \end{aligned} \tag{5}$$

and clustering coefficient c_i, for the tendency of pairs of links to form triangles,

$$c_i = A_{ii}^3 / (k_i(k_i - 1)) \tag{6}$$

as they represent the first three orders of network structure description (i.e., first, second and third neighbors).

Note that the links are not only independent of the particular notion of equilibrium but also of V_a and V_b. Instead of fixing these values into the network model, they are inferred from a combination of network and system data (in our case, the simulations). This will allow the same binary representation of the interactions to replicate physical system properties as temperature changes and, fundamentally, the differentiation of stable ($V_a < V_b$) and unstable ($V_a > V_b$) regimes of the network structure. The value of r_T, on the other hand, must be chosen somewhat arbitrarily. We take the value $r_T = r_o = 1$ as it is the minimum of the potential. Consider that, otherwise, a particle starting with arbitrarily low energy will produce a "discontinuity" in the representation if the energy of the system is

increased enough for it to cross the value of the potential energy at some other r_T which does not correspond to the minimum. This may be an interesting property in some contexts but not for the purpose of this study.

We show two different ways we can calculate V_a and V_b. In the first, we study the high-temperature limit by assuming that "free" pairs of particles (those at a distance greater than r_T) have interaction potential $V_b = 0$ and calculating V_a from the average potential energy measured from the simulations. In the second, we do not impose either V_a or V_b, obtaining both from the potential and number of links measured in the network. The goal of these two approaches is to show the flexibility and limitations of the method to physical assumptions we may make. In both cases, the relation between the calculated properties and the ones measured from the system are obtained by using the partition function of the approximated potential, which is obtained in Appendix A. In Appendix B, we discuss how this method can be extended to other potentials and weighted networks.

3. Results and Discussion

As the gas is not initially in equilibrium (its velocities are distributed uniformly), the simulations show a dynamic phase before reaching it. This means that initial kinetic and potential energies do not correspond to equilibrium ones. In Figure 3, we show the (absolute value of) equilibrium proportions of potential to kinetic energy as a function of the initial velocity for the different initial cell sides used. The ratios range from around 10 times as much potential to kinetic energy to the inverse of that, showing that a wide variety of conditions are covered.

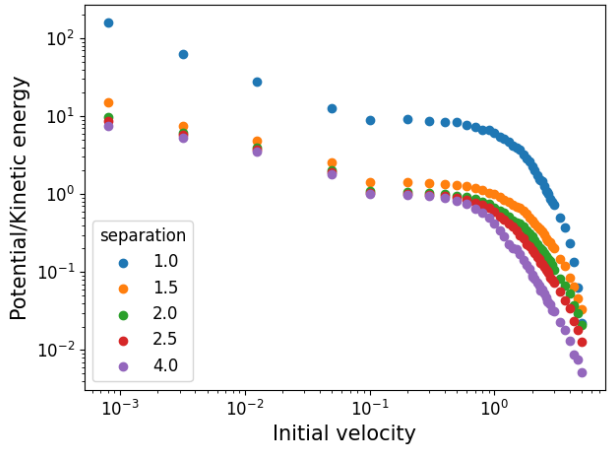

Figure 3. Ratio of equilibrium kinetic to potential energy as a function of initial velocity distribution boundaries.

All the results presented in this work are in equilibrium conditions. As simulation parameters were chosen in order to obtain at least 3/4 of the simulation time in equilibrium, we use the final half of the simulation (instead of 3/4, as the transition from dynamics to equilibrium is not a sudden event) when we refer to this regime.

3.1. Temperature and Density

As discussed in the introduction (and in much more detail by Kardar [21]), the distribution of particle speeds $v = |\vec{v}|$ in equilibrium is the same as that of a noninteracting particle system with the same kinetic energy; so, it can be used to determine the temperature T by fitting the two-dimensional Maxwell–Boltzmann distribution. In two dimensions, this temperature is also equal to the average kinetic energy, which gives us a way to verify the fitted temperature and whether the simulation has reached equilibrium.

$$\rho(v) = \frac{v}{T} \exp(-\frac{v^2}{2T}) \tag{7}$$

On the left hand side of Figure 4, we show the temperatures as a function of the initial velocities for the different initial particle separations, obtained from adjusting the temperature T in Equation (7) to the velocity distribution measured from the simulations. On the right hand side, we show the agreement between the inverse of the temperature with the kinetic energy.

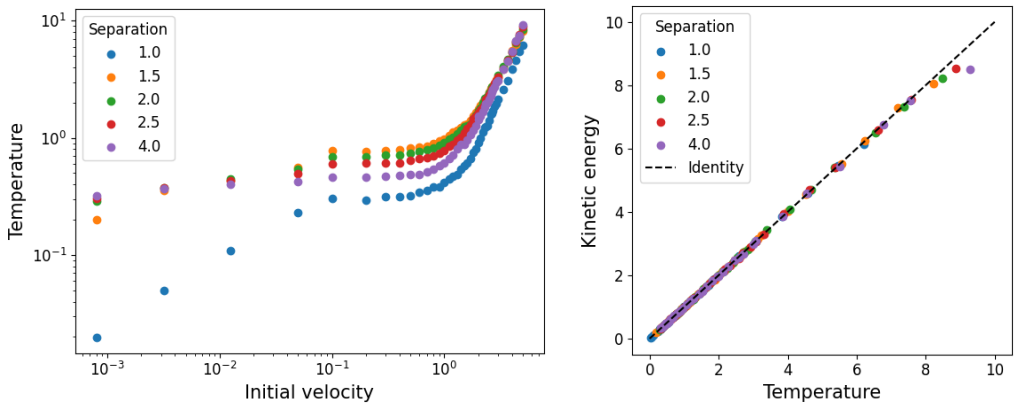

Figure 4. On the left: fitted temperature in equilibrium as a function of the initial velocity range V_M for different initial separations. On the right: agreement between fitted temperature and the average kinetic energy.

As for the spatial distribution, the average separation of two particles is given by the average of the average distance from each particle to others,

$$D^2 := \frac{1}{N}\sum_i \frac{1}{2(N-1)}\sum_{j>i}\vec{r}_{ij}^{\,2} = \frac{N}{N-1}\left(\langle\vec{r}^{\,2}\rangle - \langle\vec{r}\rangle^2\right). \tag{8}$$

This average interparticle distance (squared) defines a cell of equal sides s whose area is the inverse of the particle density. Knowing that $D^2 = 2s^2$, the particle density is $\rho = 2/D^2$. Its equilibrium value is shown as a function of the initial velocity in Figure 5. The error bars represent density fluctuations in the average distances of individual particles, $\Delta\rho = 2/(D^2)^2\Delta D^2 = 2/(D^2)^2\mathrm{var}(D^2)^{1/2}$. The values of D^2 present a mostly constant density for high temperatures but an increasing trend below a certain temperature. The size of density fluctuations increases with density, suggesting that the spatial distribution is composed of dispersed clusters of tightly packed particles, and increasingly so the lower the temperature.

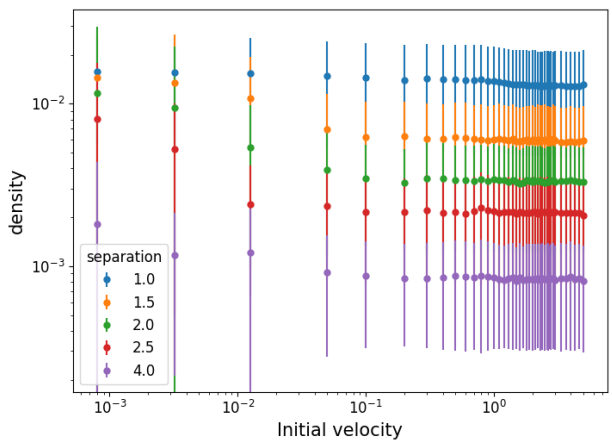

Figure 5. Density $\rho = 2/D^2$ calculated from the particle separation D in equilibrium and fluctuations as a function of the initial velocity range V_M for different initial separations.

3.2. Network Properties

In Figure 6, we show the average degree of the network: on the left as a function of the temperature, while on the right as the potential energy with inverted sign. The degree of each node (particle) is calculated for each time step in equilibrium, and the average is taken over all time steps and nodes. All densities seem to present a two-phase behavior of average degree first decreasing with temperature (as the stable structure dissolves up to the phase transition) and then increasing once again (through random fluctuations of unstable interactions). As we see in Figure 5, particle separations are roughly constant at high temperatures, meaning that the increase in degree can only be attributed to the change in kinetic energy. This implies that slowly moving particles will reach others faster (increasing link formation), but due to the relatively low velocity, these will not be able to bounce off as easily (meaning that the corresponding increase in broken links is not as large). It is also interesting to note that in the high-potential (or low-temperature) regime, the degree becomes independent of the separation.

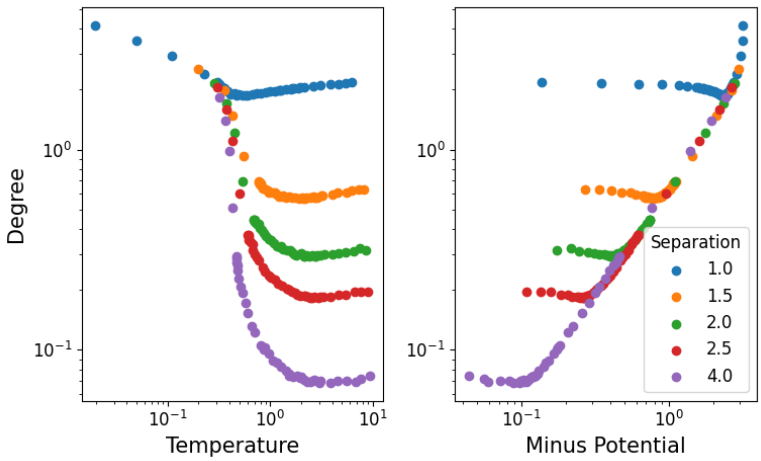

Figure 6. Average degree of nodes in equilibrium as a function of temperature for different densities, with error bars representing the spread of the degree values among nodes at any given time.

The distribution of degrees of different nodes is shown on the left-hand side of Figure 7. We see that it resembles a Gaussian distribution, suggesting that the degree is simply the result of identically distributed independent random variables. However, as the distribution is cut off at only six links (this is because the minimum potential energy is achieved in a uniform hexagonal grid), many concave functions will produce reasonable approximations. On the right-hand side, the average spectral density of a node's degree over time is shown. The spectral density represents the "intensity" of fluctuations in different frequencies of a signal. This means that a slowly changing signal will have high spectral density at low frequencies (and low spectral density at high frequencies), while a rapidly changing one will have high spectral density at high frequencies and low spectral density at low frequencies. In this case, the signal is taken to be the degree of a single node as it varies over time in the simulation (for the equilibrium range), and the result shown in Figure 7 is the average of the resulting spectral densities over all nodes in the network.

In Figure 8, we show the assortativity of the network as a function of temperature. Assortativity represents the tendency of nodes to connect to those with a different or similar degree to themselves (between -1 and 1, respectively). The fact that assortativity is a global measure that we cannot average over nodes produces an image with considerably more noise than the other measures. However, we can still make out that assortativity is positive. This represents the fact that structure (as in connected particles) is always relatively similar throughout the system (it is unlikely to find dense agglomerations of particles with long chains extending from them, for example). Starting at low temperature, the particles form a hexagonal grid, so every node is identical (except for the boundaries).

As temperature increases, some of the particles are freed from the grid and travel around, adding asymmetry to the grid and thus decreasing the assortativity. After a certain point, the free particles become the dominant property of the system (structure becomes unstable) and whatever is left of the grid is now the exception; so, as it is dissolved, nodes become more and more similar and assortativity increases again. This suggests a connection between assortativity minimum and the stability of a structure when there is a transition from repulsive to attractive behavior. It is then also interesting that the minimum of the degree happens simultaneously with assortativity, as can be seen in the degree–assortativity curves on the right of Figure 8.

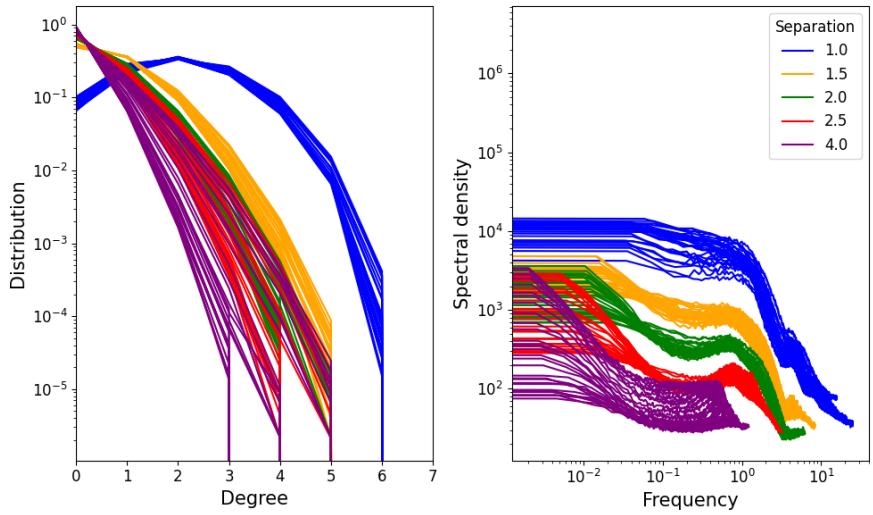

Figure 7. (**Left**) Degree distribution for different densities and temperatures. (**Right**) Average over nodes of degree spectral density.

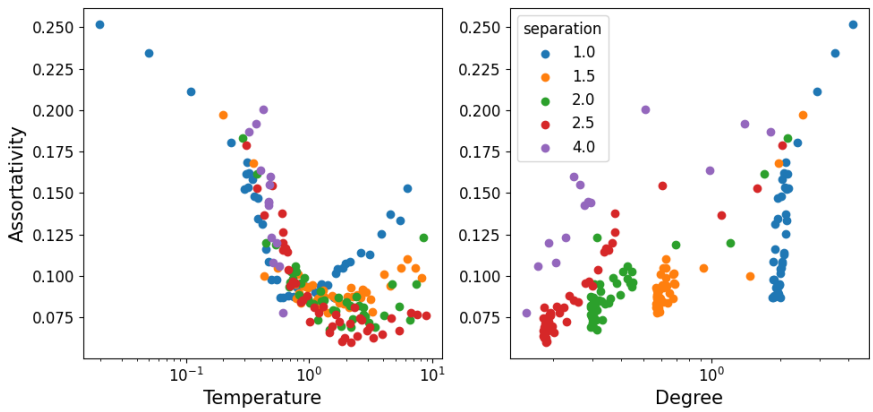

Figure 8. Assortativity (tendency of nodes to connect to others of similar degree) as a function of temperature (**left**) and degree (**right**). The figure has considerably more noise than the others, as assortativity is a global property, as opposed to the degree which we can average over nodes for each instant.

In Figure 9, we show the clustering of the network as a function of the temperature and degree. The clustering coefficient represents the tendency of a link with two neighbors to form a closed cycle of three links. This is a useful structural property, as it measures how often pairs of connections "close"; so, the fact that it scales to the average degree independently of the density makes it interesting for models of structure formation.

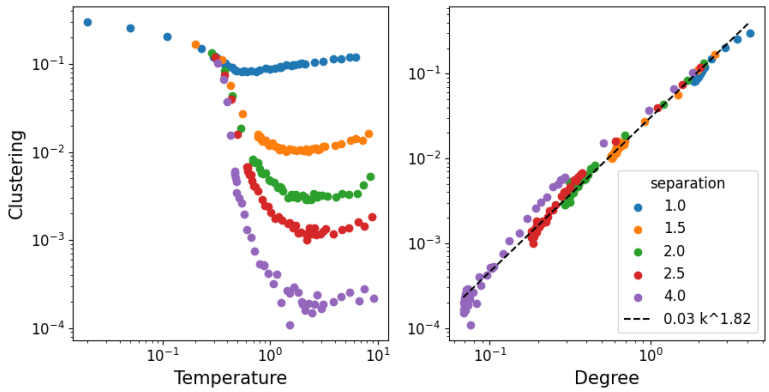

Figure 9. Clustering coefficient (tendency of pairs of connections to form triangles).

3.3. Structural Stability

We now turn to calculating the parameters V_a and V_b. Following Appendix A, we find that we can write the spatial term of the partition function

$$\mathcal{Z}_r = \left(L^2 \exp^{-\beta V_b}\right)^N \exp\left(-\frac{(N-1)\pi r_T^2}{2L^2}(1-\exp^{-\beta\Delta V})\right)^N \tag{9}$$

where the first term corresponds to the contribution of "free" particle states with energy V_b (which are beyond a distance r_T to any other), while the second corresponds to the contributions from active interactions where the pair has energy $V_a = V_b + \Delta V$. The central assumption for this result to hold is that the interaction radii r_T are small enough (with respect to the box size L) to assume that overlapping disks do so in pairs. The bad news is that this approach will likely fail for long-range potentials (nevertheless, long-range potentials are a problem in statistical physics), but the good news is that this is independent of the particular shape of the potential, as discussed in Appendix B. It is not clear how to modify the calculation in order to impose the cutoff at six links observed in Figure 7. In this form, there is a clear relation to the individual particle contribution z_1, just as for the partition function of the velocity. Since the order of integration in both the spatial and momentum space is subject to the same particle-renaming symmetry, which produces the Gibbs paradox if not accounted for [21], Equation (9) is technically missing a $1/N!$ factor. However, as we consider only fixed particle number cases, this is not expected to give us any problems. The interesting thing, however, is that any network representation will be subject to the same symmetry, which amounts to the row/column renaming of a network's adjacency matrix. Studying entropy during the mixing of two initially independent networks in order to determine whether the Gibbs paradox also occurs is an interesting future line of research.

According to Equation (9), the average potential energy of the system is

$$\langle V \rangle = -\partial_\beta \ln(\mathcal{Z}) = NV_b + \frac{N(N-1)\pi r_T^2}{2L^2} \exp^{-\beta\Delta V} \Delta V \tag{10}$$

which must match the potential energy measured from the simulations to relate our network and physical model. The value of β is known, as it is given by the temperature (kinetic energy), but V_b and ΔV are both free parameters. One of them will be determined by Equation (10), but the other is still unknown.

As a first approach, we simply assume the high-temperature limit in which the interactions beyond the strong repulsive force at short distances is negligible. In this case, the hard-shell approximation should hold, so we can set $V_b = 0$. This leaves only $V_a = \Delta V$ to be determined numerically from Equation (10). Then, considering that all the potential energy is stored in links of energy ΔV, the number of links should be $\langle V \rangle / \Delta V$. In Figure 10, we

show, on the left-hand side, the height $V_a = \Delta V$ of the high-temperature limit ($V_b = 0$) of the two-level potential. In the middle, the agreement between the measured and resulting potential energy shows that V_a correctly recovers this value. On the right-hand side, we compare the number of links measured in the simulations with that expected from $\langle V \rangle$, using only data from simulations with $T > 1$ to show the validity of the approximation beyond imposed values. It should be noted that the number of links varies much more in the low-temperature regime, making this value seem practically constant. Also, as the potential energy of the system is negative, so is ΔV, which seems to suggest that these links are stable. We interpret this as a failure of this approximation to provide results on the stability of structure due to the imposed value $V_b = 0$.

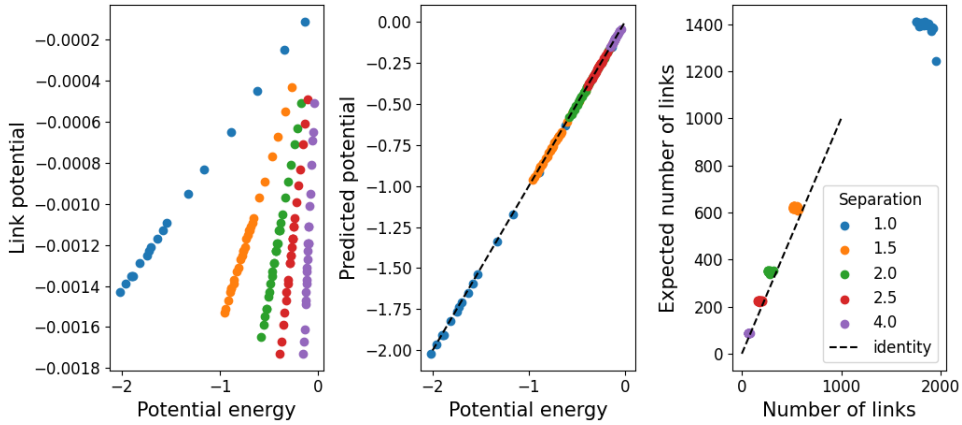

Figure 10. (**Left**) Potential $V_a = \Delta V$ determined from Equation (10) in the high-temperature limit $V_b = 0$. (**Middle**) Agreement between measured and calculated potential. (**Right**) Number of links measured from the simulations in comparison with the number of links expected from $\langle V \rangle / \Delta V$ for temperatures $T > 1$.

For the second case, we do not impose either V_a or V_b. We require an additional equation to obtain one of them. For this, we recall that a macrocanonical partition function Z_M (where the average energy and number of components are fixed by the temperature β and the chemical potential μ) obeys

$$-\partial_\beta \ln(Z_M(\beta, \mu)) = \langle E \rangle - \mu \langle N \rangle \tag{11}$$

where $\langle N \rangle$ is the average value of a fluctuating number of components. We then propose that Equation (10) takes this form so that $\Delta V = -\mu$ represents the chemical potential and $\langle N \rangle = N\langle k \rangle / 2 = N(N-1)\pi r_T^2 \exp^{\beta\mu} / 2L^2$ is the average number of links observed in the simulation. Note that μ tells us about the stability of the structure ($\mu > 0$ is stable while $\mu < 0$ is unstable) because of its definition from ΔV, which aligns with the notion of chemical potentials in statistical physics. In Figure 11, we show the values obtained for V_a and V_b in this approximation. We see that the structure becomes stable ($\mu < 0 \to \mu > 0$) below $T = 1$. In Figure 12, we show the agreement between the number of links and the measured and calculated potential.

Finally, we must verify that this way of analyzing the stability of the structure agrees with the simulations. For this, consider that when the structure is unstable, links will form randomly throughout the nodes of the network. When the structure is stable, the links are persistently associated with a pair of nodes. In order to capture this, we first measure the average over time (in equilibrium) of the adjacency matrix of the network $A_{ij} = \sum_{t=0}^{\tau} A_{ij}(t)/\tau$. If the links are unstable, the whole matrix will have roughly the same value, while if they are stable, some pairs will have a high value and others a low one. The relative fluctuations around the mean value $\mathrm{var}(A_{ij}) / \langle A_{ij} \rangle^2$ (where both $\mathrm{var}(\cdot)$ and $\langle \cdot \rangle$ are taken over links i, j) capture this effect, simultaneously correcting for the specific link density. In Figure 13, we compare this quantity (rescaled by a factor of 20 to be visible)

to the chemical potential shown in Figure 11, showing that the latter predicts a transition from unstable to stable slightly before (when the temperature is being lowered) what is shown by the relative fluctuations.

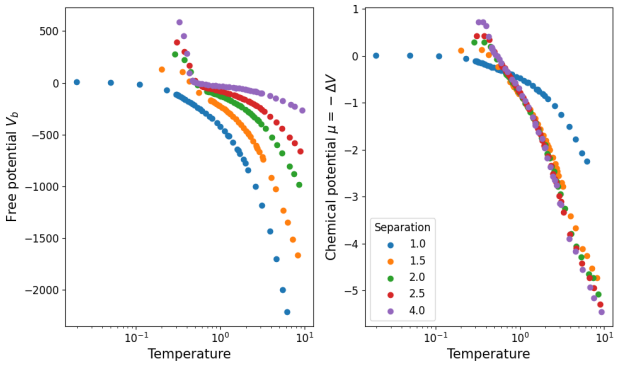

Figure 11. Partition function parameters calculated from the measured average potential energy and average number of links in the system.

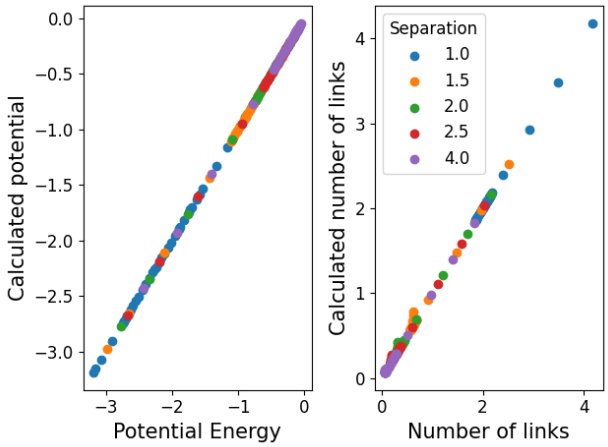

Figure 12. Agreement between simulated and calculated values of the potential energy and the number of links.

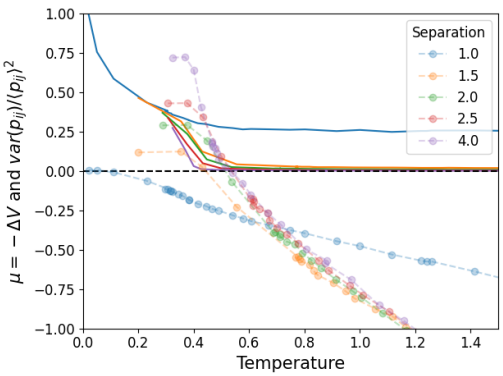

Figure 13. Relative fluctuations measured from simulations (in full lines) and chemical potential calculated from the corresponding data. The chemical potential predicts a transition from stable to unstable structure at a slightly higher temperature than observed.

4. Conclusions and Further Work

A Lennard-Jones particle system in thermodynamic equilibrium was mapped to a binary undirected network by using a two-level approximation for the interaction potential according to the distance between the particles. This was motivated by constructing a

"environment-blind" (in reference only to the interactions, without resorting to emergent properties like the temperature) network model for particle interactions, which allowed us to simplify the calculation of the system partition function. Said partition function was shown to be able to reproduce the system energy and number of connections exactly, and its parameters were used to determine the stability of the interaction structure. The partition function can also be interpreted as a macrocanonical partition function with a varying number of links, defining a chemical potential which captures the same notion of structural stability. This supports the point of view of treating links in the network as the effective fluctuating "particles" presented in statistical mechanics of networks [25,26].

The specific network structural properties studied were the degree, assortativity and clustering coefficient, taken as representative of the first three orders of interaction in a network. All three properties exhibit a minimum as a function of the temperature and, in particular, clustering was found to have the same relation to degree across different densities. These minima are aligned with the change from a stable to unstable structure (middle- to high-temperature), suggesting the two are related. Indeed, as the binary network is essentially a partition into physically distinct regions of the interaction potential, with the physical properties of the system determined through the partition function of the approximation, then phase transitions that render certain structures stable correspond to changes in the "average behavior" of the system in each region. In our case, this is the transition from weak, short-range attractive forces in low temperature to strong repulsive (yet also short-range) interactions at high temperature. This suggests that under such changes in the structural properties, it may be better to increase the "resolution" of the network approximation, using a weighted network, which has not been tested and is interesting for future development in order to see if the method can improve accuracy. Another suggested way to perform this is to relate interaction terms beyond pairwise ones in the partition function to network properties, such as the degree distribution to estimate what order the interactions should be accounted for.

If we were to add a further approximation to the potential, for example, three ranges of distance instead of two, we would attain a better model of the potential but also a network with three possible connection values. It would require further conditions on the relation between these three values (beyond the potential energy and the number of links). Assuming that this is possible, increasing the number of values to a countable infinity would produce a weighted network with integer-valued levels, and finally, the continuous limit continuous weights that match the potential. This suggests that the "natural" network representation of a system that interacts through a potential is the potential itself. However, it comes at high analytical and computational difficulty, suggesting that other representations might be more convenient.

Independently of the most convenient representation, this draws a relation between a network link and the interaction energy contained in the region of space it represents. By checking whether these links are stable, we can tell if the energy is physically contained in a structure. In this case, the network (on average) resembles the traditional representation of particles at their stable distance. This method then allows to tell when energy is contained in structure, which could be a useful tool for a more systematic study of energy density across scales [27]. For this (and other applications beyond potential-mediated interacting particle systems), the reliance on maximum entropy is particularly convenient, as the conservation of energy, which puts kinetic and potential energy in the integrand of the partition function, can be replaced by a constraint on average values of arbitrary pairwise system states, which does not require the physical notion of an interaction potential. Any specific average value that is known to reproduce properties of a system through maximum entropy (for example, $\sum_{i,j}(\ln(s_i) - \ln(s_j))^2/N^2$) then takes the place of the interaction potential, and the link represents the measure of this function (replacing the energy) contained in the states (replacing positions) corresponding to a link.

Author Contributions: Conceptualization, N.A. and F.R.; Methodology, N.A. and F.R.; Software, N.A.; Validation, N.A. and F.R.; Formal analysis, N.A.; Investigation, N.A.; Writing—original draft, N.A. and F.R.; Writing—review & editing, N.A. and F.R.; Supervision, F.R.; Funding acquisition, F.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Faculty of Science and Engineering of the University of Groningen.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Python code for generating the simulation data and results, as well as the data themselves, is available on request from the corresponding author in order to comply with the data management policy of the Integrated Research on Energy, Environment and Society.

Acknowledgments: The authors would like to thank Kaan Hidiroglu, Danyang Zhang and Jin Yan for their participation in discussions on complex networks.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. System Partition Function

In order to calculate the partition function of the system in the two-level potential approximation, we start by recalling that an ideal gas has uniformly distributed particle positions. This implies that if we set links between particles when there distance is smaller than some threshold, the probability p_n that a test node (particle) has n links is the probability that there are exactly n (randomly placed) particles within a radius of r_T . This is given by Equation (A1), and the expected degree of a single node is Equation (A2). The expected total number of links is $N\langle k\rangle/2$, ultimately leading to a link density equal to $\pi r_T^2/L^2$ in the infinite temperature limit.

$$p_n = \binom{N-1}{n}\left(\frac{\pi r_T^2}{L^2}\right)^n\left(1-\frac{\pi r_T^2}{L^2}\right)^{N-1-n} \tag{A1}$$

$$\langle k\rangle = \sum_n n p_n = n\binom{N-1}{n}\left(\frac{\pi r_T^2}{L^2}\right)^n\left(1-\frac{\pi r_T^2}{L^2}\right)^{N-1-n} = (N-1)\frac{\pi r_T^2}{L^2} \tag{A2}$$

We now improve this estimation to the first order in the interactions with the partition function

$$Z(\beta) = \int\int ... \int \Pi_{i=1}^N \exp(-\beta\sum_{j>i} V(r_{ij}))dr_i \tag{A3}$$

and, introducing the two-step potential approximation, we can visualize the first integration of the partition function $z_1 = \int \exp(-\beta\sum_{j>1} V(r_{1j}))dr_1$ as the particle 1 "sweeping" over the whole area L^2 and measuring the height $\exp^{-\beta\sum_{j>1} V(r_{1j})}$ at each point. This height is given by a surface of height $\exp -\beta V_b$, on top of which a series of $N-1$ columns of radius r_T reach a total height $\exp^{-\beta V_a}$, where the columns' bases do not overlap with each other and have a height of $\exp -\beta n V_a$ wherever n columns overlap. Note that only the regions that overlap have a height of $\exp^{-\beta n V_a}$, as shown in Figure A1.

A first approximation of the integral would be $L^2\exp^{-\beta V_b}$ (disregarding the areas with at least one column). A second, better approximation would be $L^2\exp^{-\beta V_b} + n\pi r_T^2$ $(\exp^{-\beta V_a} - \exp^{-\beta V_b})$ (disregarding the areas of two or more intersecting columns). In general, an approximation to n intersections is improved by introducing the area where $n+1$ or more disks overlap, correcting the overcounted area and undercounted height, which is the difference between the previous height approximation and the current estimate. For the second order, we can write

$$
\begin{aligned}
z_1(\beta) := &\int \exp(-\beta \sum_{j>1} V(r_{1j}))dr_1 \\
\approx &L^2 \exp^{-\beta V_b} + \Big[(N-1)\pi r^2 - A_2(r_{j>2})\Big](\exp(-\beta V_a) - \exp(-\beta V_b)) \\
&+ A_2(r_{j>1})[\exp(-\beta 2V_a) - \exp(-\beta V_b)] \\
= &L^2 \exp^{-\beta V_b} + (N-1)\pi r^2(\exp(-\beta V_a) - \exp(-\beta V_b)) \\
&+ A_2(r_{j>1})[\exp(-\beta 2V_a) + \exp(-\beta V_b) - 2\exp(-\beta V_a)]
\end{aligned} \tag{A4}
$$

where $A_2(r_{j>1})$ represents the area where at least two circles $j \neq j'$ (both indices greater than 1) intersect for the fixed configuration $(r_j)_{j>1}$. This is also the only term dependent on r_2, which is the next in line to integrate. This means that if the approximation of $A_2 \approx 0$ stays small throughout the integration over different particles i to obtain $\mathcal{Z}$, then the integration over the previous particle $i-1$ will appear constant to each term. In order for the approximation to hold as we integrate the different particles, we need to make sure that the perturbation A_2 stays small. Decomposing it into the contributions from intersections with $j = 2$ and $j > 2$, we write $A_2(r_{j>2}) = A_2(r_2, r_{j>2}) + A_2(r_{j>2})$. This way, $z_1(\beta) = a_{N-1} + b(r_2) + c(r_{j>2})$, which means that the integration over r_2 will only interact with $b(r_2)$, while a and $c(r_{j>2})$ are constants. For these, integration over r_2 will go just as for z_1, but since the first particle has already been integrated, there are $N-2$ columns instead of $N-1$ (also note that the sum in the exponent is $\sum_{j>2}$ for this term).

$$
\begin{aligned}
z_2(\beta) = &\int z_1(\beta) \exp(-\beta \sum_{j>2} V(r_{2j}))dr_2 \\
= &(a_{N-1} + c(r_{j>2})) \int \exp(-\beta \sum_{j>2} V(r_{2j}))dr_2 \\
&+ \int b(r_2) \exp(-\beta \sum_{j>2} V(r_{2j})dr_2 \\
= &(a_{N-1} + c(r_{j>2}))(a_{N-2} + b(r_3) + c(r_{j>3})) \\
&+ \int b(r_2) \exp(-\beta \sum_{j>2} V(r_{2j})dr_2
\end{aligned} \tag{A5}
$$

In order to integrate the second term, consider that $A_2(r_2, r_{j>2}) = b(r_2)/(e^{-\beta 2V_a} + e^{-\beta V_b} - 2e^{-\beta V_a}) > 0$ only if r_2 is within a distance r_T of some $r_{j>2}$. As r_2 is moved around, the area varies between 0 (where there are no particles $j > 2$ within a distance r_T of r_2) and πr_T^2 (which corresponds to r_2 coinciding with the center of another disk $j > 2$). The value of the energy exponential will take on discrete values $\exp^{-\beta n V_b}$, just as shown in Figure A1. So, a first approximation makes the height that r_2 sees constant $\exp^{-\beta V_b}$ (note that the two or more intersections must include r_2; so, there is actually 1 or more connection for the particles $j > 2$, which is the height that r_2 sees). The first correction occurs where there are intersections of 2 or more $r_{j>2}$; so, we can set a bound to the first order of $\int A_2(r_2, r_{j>2}) \exp^{-\beta \sum_{j>2} V(r_{2j})} dr_2$ to $(N-2)\pi r_T^2(\exp^{-\beta V_a} - \exp^{-\beta V_b})$. A better estimation would be to replace πr_T^2 by the average area of the intersection of two circles, keeping one fixed and moving the other over the interior of the first, which is smaller than πr_T^2, as the area (the random variable) is bounded by πr_T^2. One can already see that this term is much smaller than a_n so long as $r_T \ll L$ (the interaction regions cover a small area of the box). The second-order term is the average area of intersection of one circle with two others fixed inside the region where they both intersect. This is even smaller than the first-order term because, if the intersection of the circles is small, the intersection of the three will be given by the intersection of the fixed ones. If the intersection is large, it will still be less than the case where the fixed circles overlap perfectly, which gives the same result as the first-order

term. All of this means that we can place a weak upper bound (in the sense that it is likely far from the true value) on the term to integrate in Equation (A5), and this upper bound remains below the leading order, thus conserving the order of the approximation. Moreover, when this term is integrated to obtain the partition function of the third particle (note that it depends on r_3, taking it to be fixed when r_2 is integrated), it will only vary depending on whether r_2 and r_3 intersect, which happens in an area πr_T^2 (which we discussed should be small compared with L^2). Thus, it is constant for almost all the integration range and will practically remain the same when r_3 is integrated. The remaining term proportional to $b(r_3)$ in Equation (9) will become analogous to the term just discussed; so, the same arguments hold. Therefore, the perturbation stays bounded for successive integrations, and we can calculate the partition function to the first order keeping only the independent terms

$$z_n = L^2 \exp^{-\beta V_b} + (N-n)\pi r_T^2 \left(\exp^{-\beta V_a} - \exp^{-\beta V_b} \right) \tag{A6}$$

to leading order in $\pi r_T^2 / L^2$. Therefore, the partition function is

$$Z(\beta) = \Pi_{n=1}^{N} z_n = \left(L^2 \exp^{-\beta V_b} \right)^N \Pi_{n=0}^{N-1} \left(1 - n \frac{\pi r^2}{L^2} (1 - \exp(-\beta(V_a - V_b)) \right). \tag{A7}$$

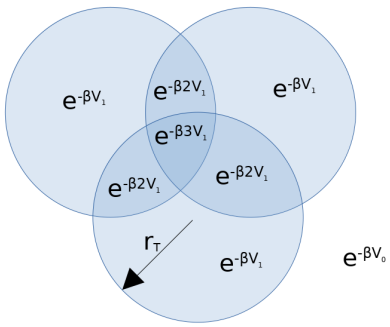

Figure A1. Representation of the total height of the surface integrated in the first integrand of Equation (A3). This is also the two-level approximation of the potential "seen" by particle 1 with a fixed configuration of the other particles.

Finally, we discussed that $r_T \ll L$; so, we can write

$$\begin{aligned} &\Pi_{n=0}^{N-1} \left(1 - n \frac{\pi r^2}{L^2} (1 - \exp(-\beta(V_a - V_b)) \right) = \\ &\exp\left(\sum \ln \left(1 - n \frac{\pi r^2}{L^2} (1 - \exp(-\beta(V_a - V_b)) \right) \right) \approx \\ &\exp\left(- \sum n \frac{\pi r^2}{L^2} (1 - \exp(-\beta(V_a - V_b)) \right) = \\ &\exp\left(\frac{N(N-1)}{2} \frac{\pi r^2}{L^2} (1 - \exp(-\beta(V_a - V_b)) \right) \end{aligned} \tag{A8}$$

which is equivalent to Equation (9) if we write $\Delta V = V_a - V_b$ and $\exp^{xN} = \exp x^N$.

Appendix B. Beyond the Lennard-Jones Potential and Binary Networks

Two things should be highlighted about this method. First, the central assumption is that overlaps of more than two particle disks can be disregarded, which means that the interaction radii must be small. The second is that, since the partition function counts states with the same energy uniformly, our two-level approximation of the real potential between two particles can have as many steps as we want, while taking only values V_a and V_b (i.e., assuming the network is binary). These are all "accumulated" by the partition function into a single radius r_T spanning the total lengths of the steps, as shown in Figure A2. Of

course, the steps should be used close to the particle in order for the short-range hypothesis to hold, but this depends on the potential, not the approximation.

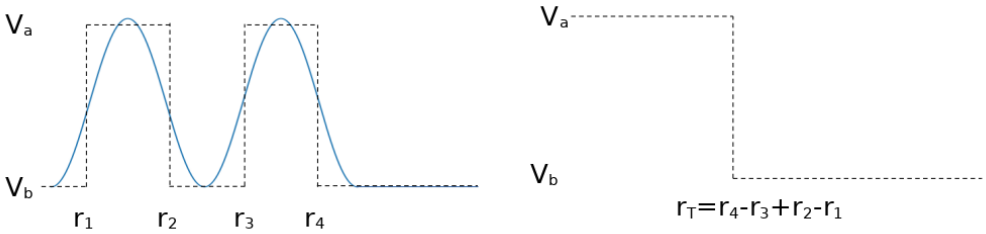

Figure A2. Arbitrary two-level approximations of a potential are "measured" by the partition function as a single radius of the total length of the steps, meaning that we can place steps however we like as long as they obey the short-range approximation, and this will always amount to an effective radius r_T.

One can also choose (just as the high-temperature limit is chosen in Section 3.3) to relax the approximation of both steps on the left side of Figure A2, which have the same potential, and then obtain a weighted network with three different values, mapping to V_b, V_a and a new $V_{a'}$. This adds a variable that can be calculated by measuring thermodynamic variables beyond the potential energy, such as the entropy of the distribution (which can also be calculated from the partition function). It is not clear how other network properties might be involved in this process, but it may be that the restriction of first-order interactions may be too strong for this purpose. On the other hand, this approximation could perhaps be corrected by using network properties. In particular, the degree, which represents the number of connections a particle has, could improve the approximation that interactions occur mostly in pairs.

References

1. Van Steen, M. Graph theory and complex networks. *Introduction* **2010**, *144*, 1–287.
2. Reichardt, J. Introduction to complex networks. In *Structure in Complex Networks. Lecture Notes in Physics*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 1–11.
3. Estrada, E. Introduction to complex networks: Structure and dynamics. In *Evolutionary Equations with Applications in Natural Sciences*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 93–131.
4. Wen, T.; Cheong, K.H. The fractal dimension of complex networks: A review. *Inf. Fusion* **2021**, *73*, 87–102. [CrossRef]
5. Juszczyszyn, K.; Musial, A.; Musial, K.; Bródka, P. Molecular dynamics modelling of the temporal changes in complex networks. In Proceedings of the 2009 IEEE Congress on Evolutionary Computation, Trondheim, Norway, 18–21 May 2009; pp. 553–559.
6. Ozkanlar, A.; Clark, A.E. ChemNetworks: A complex network analysis tool for chemical systems. *J. Comput. Chem.* **2014**, *35*, 495–505. [CrossRef] [PubMed]
7. Amamoto, Y.; Kojio, K.; Takahara, A.; Masubuchi, Y.; Ohnishi, T. Complex network representation of the structure-mechanical property relationships in elastomers with heterogeneous connectivity. *Patterns* **2020**, *1*, 100135. [CrossRef]
8. García-Sánchez, M.; Jiménez-Serra, I.; Puente-Sánchez, F.; Aguirre, J. The emergence of interstellar molecular complexity explained by interacting networks. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2119734119. [CrossRef] [PubMed]
9. Hann, M.M.; Leach, A.R.; Harper, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856–864. [CrossRef] [PubMed]
10. Goodby, J.W.; Saez, I.M.; Cowling, S.J.; Gasowska, J.S.; MacDonald, R.A.; Sia, S.; Watson, P.; Toyne, K.J.; Hird, M.; Lewis, R.A.; et al. Molecular complexity and the control of self-organising processes. *Liq. Cryst.* **2009**, *36*, 567–605. [CrossRef]
11. Bertz, S.H. The first general index of molecular complexity. *J. Am. Chem. Soc.* **1981**, *103*, 3599–3601. [CrossRef]
12. Böttcher, T. An additive definition of molecular complexity. *J. Chem. Inf. Model.* **2016**, *56*, 462–470. [CrossRef]
13. Milano, M.; Agapito, G.; Cannataro, M. Challenges and limitations of biological network analysis. *BioTech* **2022**, *11*, 24. [CrossRef]
14. Rodríguez, F.Y.; Muñuzuri, A.P. A Goodwin Model Modification and Its Interactions in Complex Networks. *Entropy* **2023**, *25*, 894. [CrossRef] [PubMed]
15. Watts, A. A dynamic model of network formation. *Games Econ. Behav.* **2001**, *34*, 331–341. [CrossRef]
16. Chandrasekhar, A. Econometrics of network formation. In *The Oxford Handbook of the Economics of Networks*; Oxford University Press: Oxford, UK, 2016; pp. 303–357.
17. Barzel, B.; Barabási, A.L. Universality in network dynamics. *Nat. Phys.* **2013**, *9*, 673–681. [CrossRef]
18. Cameron, M.; Vanden-Eijnden, E. Flows in complex networks: Theory, algorithms, and application to Lennard–Jones cluster rearrangement. *J. Stat. Phys.* **2014**, *156*, 427–454. [CrossRef]

19. Forman, Y.; Cameron, M. Modeling aggregation processes of Lennard-Jones particles via stochastic networks. *J. Stat. Phys.* **2017**, *168*, 408–433. [CrossRef]
20. Kasatkin, D.V.; Nekorkin, V.I. Transient Phase Clusters in a Two-Population Network of Kuramoto Oscillators with Heterogeneous Adaptive Interaction. *Entropy* **2023**, *25*, 913. [CrossRef] [PubMed]
21. Kardar, M. *Statistical Physics of Particles*; Cambridge University Press: Cambridge, UK, 2007.
22. Martyushev, L.M. Maximum entropy production principle: History and current status. *Phys.-Uspekhi* **2021**, *64*, 558. [CrossRef]
23. Dewar, R.C. Maximum entropy production as an inference algorithm that translates physical assumptions into macroscopic predictions: Don't shoot the messenger. *Entropy* **2009**, *11*, 931–944. [CrossRef]
24. Hut, P.; Makino, J.; McMillan, S. Building a better leapfrog. *Astrophys. J.* **1995**, *443*, L93–L96. [CrossRef]
25. Park, J.; Newman, M.E. Statistical mechanics of networks. *Phys. Rev. E* **2004**, *70*, 066117. [CrossRef]
26. Albert, R.; Barabási, A.L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **2002**, *74*, 47. [CrossRef]
27. Smil, V. *Energy in Nature and Society: General Energetics of Complex Systems*; MIT press: Cambridge, MA, USA, 2007.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.