Physics 344 Assignment 3

Jean du Plessis 23787295

August, 2022

1 Introduction

We would like to investigate how geometric graphs can be understood from a statistical mechanics perspective. Previous approaches have studied models where there are transitions between geometric and random graphs [2], between geometric and complete graphs [1] and between geometric and disconnected graphs [3]. No (known to the author) major progress has been made linking transitions between lattices and geometric graphs, or even more ambitiously, all of the above. The model described here has two clear lattice phases, two clear disconnected phases, and one phase which appears geometric (but has not yet been shown to be such). The investigated model also has no interactions yet, whose addition might also be vital to more complicated/intricate behaviour.

2 Preliminaries

All graphs we will be dealing with are simple undirected graphs.

Remark. *Underlying graph* here will always refer to a square lattice graph in 2D, unless explicitly stated.

Definition 1. In a graph a *short-cycle* refers to some sequence of vertices of the graph, such that two of the vertices in the sequence are adjacent if and only if they are consecutive in the sequence (regarding the initial and final vertices of the sequence as consecutive).

Since only consecutive vertices are adjacent in a short cycle, we can define the following:

Definition 2. In a short cycle a *shortcut* is an edge (which doesn't necessarily exist in the graph) connecting any pair of vertices in the short cycle.

Not all shortcuts are created equal, so we can quantify this with:

Definition 3. The *strain* of a shortcut is the distance in the underlying graph between its two vertices.

For example we can see then that every edge in a graph that is part of a short-cycle will be a shortcut of strain 1.

3 Defining the model

For any graph G, if we consider the set S of shortcuts in G, such that the strains can be seen as a function $s : S \to \mathbb{Z}^+$. We will then be considering the ensemble defined as the powerset $\mathcal{P}(S)$, with a 'free Hamiltonian' $H : \mathcal{P}(S) \to \mathbb{R}^+$ defined by

$$H(X) = \sum_{x \in X} (s(x) - 1) \tag{1}$$

in analogy with the 'energy distance' used in [2]. Since we consider the entire powerset as our ensemble (and can see edges as analogous to Fermionic particles) we are in the grand canonical ensemble, giving us a distribution over our ensemble parameterized by β, μ, the inverse temperature and chemical potential respectively. The probability of a state $\in \mathcal{P}(S)$ is given by

$$P(X) = \frac{e^{\beta(\mu|X| - H(X))}}{\mathcal{Z}}, \tag{2}$$

where the grand partition function $\mathcal{Z}$ is given by

$$\mathcal{Z} = \sum_{X \in \mathcal{P}(S)} e^{\beta(\mu|X| - H(X))}. \tag{3}$$

Here we are interpreting edges as indistinguishable Fermionic particles with chemical potential μ, and energy given by what type of shortcut they are.

4 Special Case to be investigated computationally

The above described model is quite general, and we will only be investigating a special case in this report. We will be concerned with the case where the underlying graph is a $n \times n$ square lattice graph. If we then call V the amount of edges in the underlying graph ($V \approx 2n^2$), we can (in the limit V large, or with periodic boundary conditions) approximate the grand partition function quite easily.

4.1 Grand Partition function and expectation values

We first notice that since there are no interactions, each edge can safely be treated as independent. Then for the square lattice there are only two types of shortcuts (of V each), the first are edges in the underlying graph, which will contribute nothing to the Hamiltonian, and the second are shortcuts of strain 2, contributing 1 to the Hamiltonian each. Then for a fixed number of edges N, we easily get the grand partition function

$$\mathcal{Z} = \left(1 + e^{\beta\mu}\right)^V \left(1 + e^{\beta(\mu-1)}\right)^V, \tag{4}$$

which in turn gives us the grand potential

$$\Omega(V, \beta, \mu) = -\frac{V}{\beta}\left[\ln\left(1 + e^{\beta\mu}\right) + \ln\left(1 + e^{\beta(\mu-1)}\right)\right]. \tag{5}$$

We can use this to calculate the expected number of edges

$$\langle N \rangle = -\frac{\partial \Omega}{\partial \mu} \tag{6}$$

$$\frac{\langle N \rangle}{V} = 2 - \frac{1}{1 + e^{\beta\mu}} - \frac{1}{1 + e^{\beta(\mu-1)}} \tag{7}$$

which one can very is the same as calculating it using probability theory. Similarly for the Gibbs entropy (with $k_B = 1$)

$$S = -\frac{\partial \Omega}{\partial T} \tag{8}$$

$$\frac{S}{V} = \beta^2 \frac{\partial \Omega}{\partial \beta} \tag{9}$$

$$= \beta - \beta \frac{1}{1 + e^{\beta(\mu-1)}} - \beta\mu\left(2 - \frac{1}{1 + e^{\beta\mu}} - \frac{1}{1 + e^{\beta(\mu-1)}}\right) + \ln\left(1 + e^{\beta\mu}\right) + \ln\left(1 + e^{\beta(\mu-1)}\right) \tag{10}$$

$$S = \ln(\mathcal{Z}) + \beta V\left(1 - \frac{1}{1 + e^{\beta(\mu-1)}}\right) - \beta\mu\langle N \rangle. \tag{11}$$

This last form allows us to identify $\langle E \rangle = V\left(1 - \frac{1}{1+e^{\beta(\mu-1)}}\right)$, and because of how our Hamiltonian is set up, this will just tell us the amount of 'excited' edges there are. This can again be checked to match with basic probability theory calculations.

4.2 Percolation theory

Using percolation theory we can define 5 phases in this special case. We can do this by using percolation theory order parameters as follows. The percolation conditions considered will be

I The bond percolation of the resulting graph.

II The site percolation of vertices of degree 8 in the resulting graph.

III The site percolation of vertices of degree 4 in the resulting graph.

IV The site percolation of vertices of degree 0 in the underlying graph.

We can then name the phases in the following way

- A completely disconnected phase where percolation IV happens.
- A minimally disconnected phase where neither percolation I nor IV happens.
- A amorphous graph phase when only percolation I happens.
- A lattice graph phase when percolation III happens.
- A degenerate lattice graph phase when percolation II happens.

We can then naively (ignoring differences between edge probabilities, hoping some universality class magic saves us) estimate what our phase diagram will look like. Looking it up on Wikipedia we find percolation threshold probabilities (noting that the full resultant graph is a sq-1,2 lattice):

I $0.250368\ldots$

II $0.407\ldots$

III $0.407\ldots$

IV $0.59274\ldots$

which gives us the following approximate conditions (using some very straightforward but tedious probability theory calculations):

I $\frac{1}{2}\left(2-\frac{1}{1+e^{\beta\mu}}-\frac{1}{1+e^{\beta(\mu-1)}}\right)>0.250368$

II $\left(\frac{e^{\beta\mu}}{1+e^{\beta\mu}}\frac{e^{\beta(\mu-1)}}{1+e^{\beta(\mu-1)}}\right)^4>0.407$

III $\frac{1}{128}\left[18+16\cosh(\beta)+\cosh(2\beta)\right]\operatorname{sech}^4\left(\frac{\beta\mu}{2}\right)\operatorname{sech}^4\left(\frac{\beta(\mu-1)}{2}\right)>0.407$

IV $\left(\frac{1}{1+e^{\beta\mu}}\frac{1}{1+e^{\beta(\mu-1)}}\right)^4>0.59274$

This gives us the regions seen on the left of Figure 1. We can then use these regions together with the definitions of the phases to get a naive view of what we can expect the phase diagram to look like, as can be seen on the right of Figure 1.

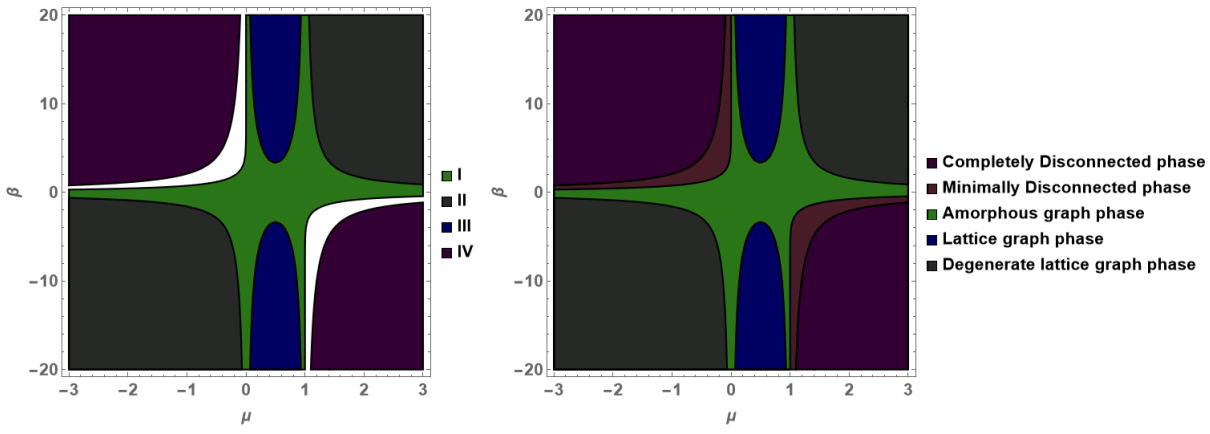

Figure 1: (Left) Regions in $\beta-\mu$ space where the different forms of percolation happens. (Right) Naive phase diagram.

5 Methods

5.1 In general

Throughout this report, typical graphs are generated by computing the probability of a strain 1 and 2 edge existing respectively. Then we can sample a binomial distribution of each to get a typical amount of each edge type, which we can then use to sample that amount from all possible edges of the corresponding type (without repetition) to construct a graph from. Possible edges are generated by finding all short-cycles (in this case cycles of length 4) in the underlying graph, and then from each enumerating all possible pairs of vertices that are in the same short cycle. We can discard duplicates, and split the remaining edges by what the distance in the underlying graph is between the two vertices of the edge. Since this will often be called repeatedly for the same underlying lattice, memoisation is used to offer significant speedup. This gives us a fast way to generate graphs that are typical for a given β, μ in a statistically unbiased way, such that ensemble averages can be calculated by a simple average of quantities calculated from these graphs.

5.2 Percolation

In order to find the percolation transitions, we will in all the cases use the probability of a random vertex being in the largest connected component as the order parameter. Below the transition we should find the probability limiting to 0, while above the transition it should limit to some finite value. Since we have a rough idea of where the transition should occur, we can do a binary search, where we look for the transition point between limiting down and up. We use 50×50 and 100×100 underlying lattices as our small and large lattices in each case. For several values of μ we can then use binary search to find at which β value the transition occurs, allowing us to observe how close to the naive theoretical behaviour described above the simulations are. At each value of μ we generate 1024 typical graphs for the given value of μ and β to estimate the mean size of the largest component (which gets rescaled by dividing by the total number of vertices to get a percentage), and do 8 steps of binary search (each step generating 1024 50×50 and 1024 100×100 typical graphs).

5.3 Entropy

We would like to investigate the ensemble average entropy of the degree distribution of the graphs (we will refer to it as degree entropy S_v from here on). To estimate the ensemble average degree entropy at a given μ and β we will generate 1024 typical graphs with a 50×50 square lattice with periodic boundary conditions as underlying graph. The periodic boundary conditions here avoid boundary effects which are relatively severe when investigating vertex degree. We will investigate $-3 \leq \mu \leq 3, -20 \leq \beta \leq 20$ by sampling on a 50×50 grid of points, in order to get an estimate of how it changes in the chosen region. This can then be compared to the Gibbs entropy S calculated above, which has more information, since it can distinguish between different edgetypes, while the degree distribution cannot.

5.4 Geometry

Since it has been shown [2] that in a certain sense geometric graphs are maximum entropy systems in the ensemble of graphs with a fixed mean vertex degree and clustering coefficient, it seems important to get a numerical handle on these quantities in the investigated model, to hopefully show that the amorphous phase yields graphs that are typical for some class of random geometric graphs. These are also standard quantities under consideration in graph and complex network theory. Mean local clustering coefficient was found to be well-behaved over the entire investigated region, while global clustering coefficient gives issues when computing it on graphs without (or with very few) edges. Since they agree in the well-behaved regions (including the amorphous phase we are interristed in), mean local clustering coefficient is used as a proxy for global clustering coefficient. The same region as the entropy is investigated, and using the same method as described above.

5.5 Mean Graph Distance

The ideal quantity to investigate here would be the graph diameter of the largest component (normalised by the diameter of the inderlying lattice), as one might intuitively expect it to diverge when approaching the percolation I transition from the one side, and to be 0 on the other. This however is extremely computationally expensive to compute (as both finding the largest component, and calculating the graph diameter are very expensive for large graphs). What is instead tractable with the currently available computing resouces is to estimate the mean graph distance between vertices in the same component. To estimate this we choose 100 random vertices (excluding singlets to speed up computation and introduce neglible bias in the considered region) and calculate the mean distance to every other vertex in it's component. We then return the mean of these 100 means. This is expected to have comparable qualitative results to the (normalised) graph diameter of the largest

component. Even with these more tractable problems, the rate at which it limits is extremely slow, so it is neccessary to go to very large lattice sizes, making it challenging to get completely stable statistics. The described statistics is calculated on the line $\mu = -\beta$ for $0.5 < \beta < 1$, which crosses the percolation I threshold line. This interval is sampled by dividing it into 50 equally sized pieces. In this interval we also then divide the interval $0.8 < \beta < 0.9$ into 50 equally sized pieces to get more finely grained results around the point where the theoretical transition happens (easily calculate to be approximately $\beta \approx 0.847$ on the investigated line).

6 Results

6.1 Percolation

Using the methods described in the previous section we have found estimates of where the phase transitions happen, which can be seen in Figure 2.

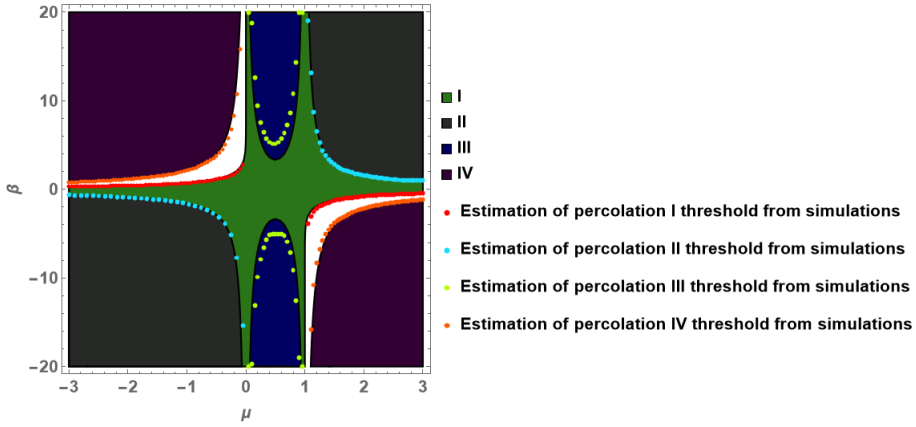

Figure 2: Estimates of phase transition lines found in simulations plotted on top of theoretical percolation regions.

We can see that for all but percolation III we have excellent agreement. If we rather conjecture that since the percolation III transition occurs in regions where one of the two edgetypes are of probability almost 0, that it 'looks' like a percolation on a square lattice, such that the percolation threshold is actually around $0.59274\ldots$. If we then alter our conditions above, we find excellent agreements for all percolation types, as can be seen in Figure 3.

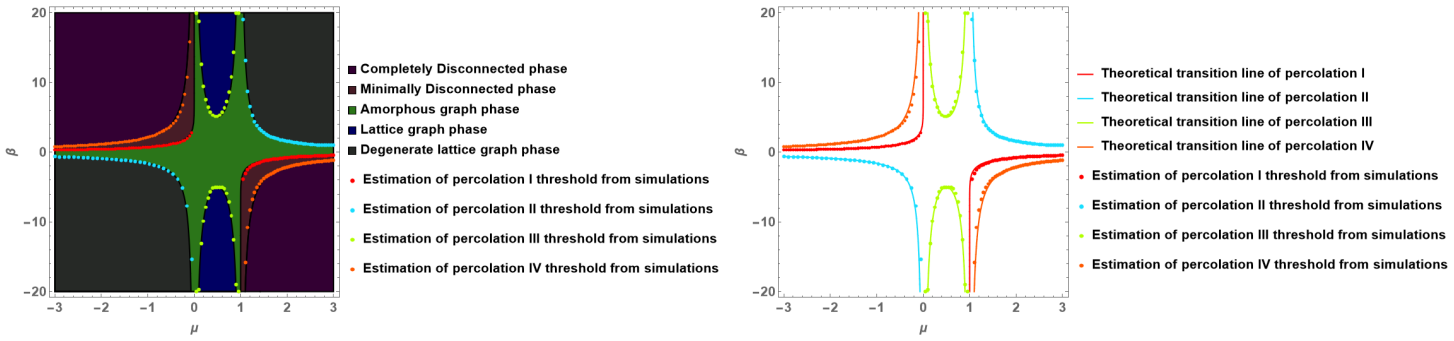

Figure 3: (Left) Corrected theoretical phase diagram, matching numerical data from simulations. (Right) The theoretical transition lines are plotted against the data from simulations.

6.2 Entropy

Again implementing methods described above we find the results plotted on the left in Figure 4. On the right we can see the Gibbs entropy calculated from the grand partition function.

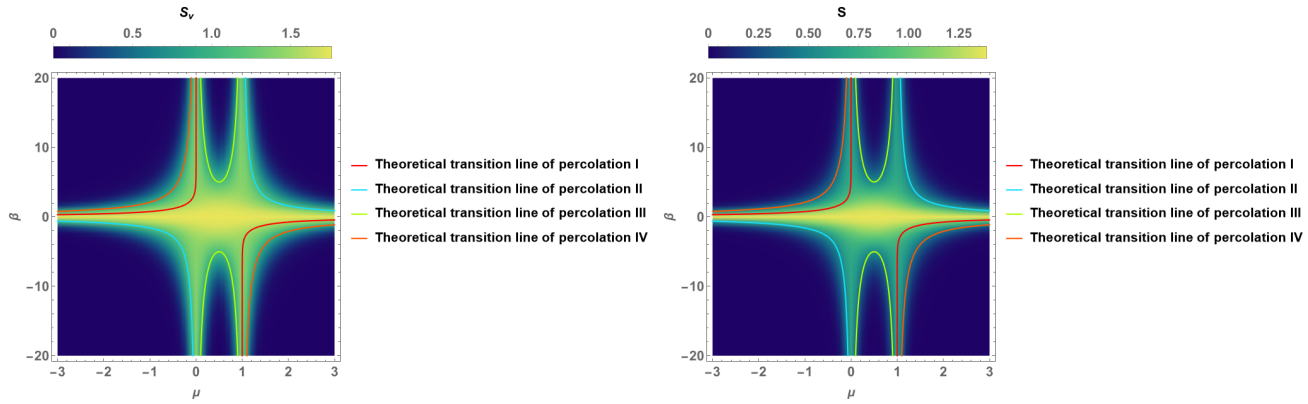

Figure 4: (Left) Estimates of degree entropy found in simulations plotted with theoretical percolation transition lines. (Right) Gibbs Entropy calculated from grand partition function.

The nature of the agreement between the two is surprising, and warrants further investigation.

6.3 Geometry

The mean vertex degree $\overline{\mathrm{D}}$ and mean local clustering coefficient $\overline{\mathrm{C}}_\ell$ in the investigated region can be seen in Figure 5.

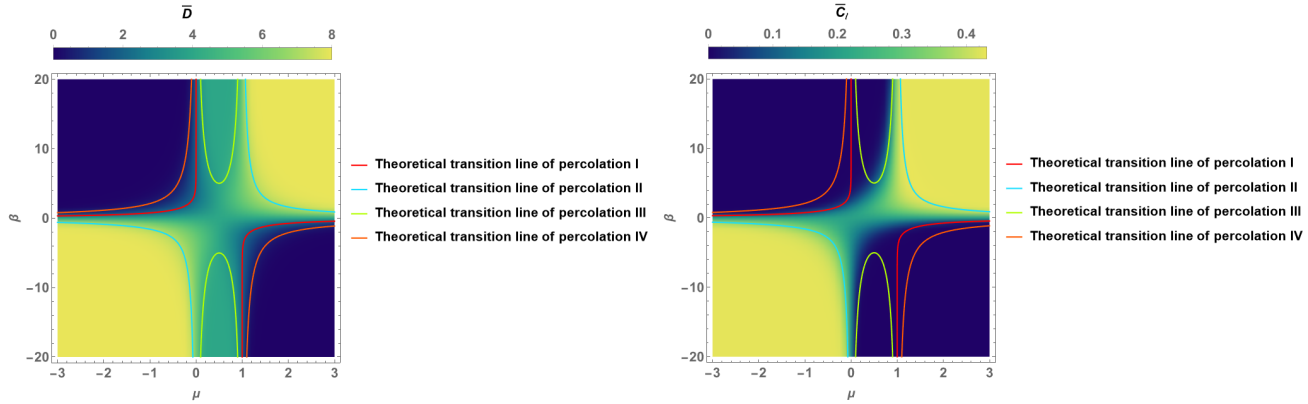

Figure 5: (Left) Estimates of mean vertex degree found in simulations plotted with theoretical percolation transition lines. (Right) Estimates of mean local clustering coefficient found in simulations plotted with theoretical percolation transition lines.

Unsurprisingly we can see that the mean degree is approximately 8 across the region where percolation II happens, 4 across the region where percolation III happens and 0 across the region where percolation IV happens. We can also see the amorphous phase split into three regimes with high, medium and low clustering respectively. It is doubtful that the low clustering (and possibly also high clustering) regime would represent a geometric graph regime.

6.4 Mean Graph Distance

As explained above, this is quite a computationally expensive problem, even with the simplifications considered. We can see that satisfactory statistical stability has not been achieved in Figure 6. Nevertheless we get a good qualitative idea of how distances act near the transition line for percolation I. Similar behaviour is expected for the diameters of the largest components. The plot inset shows behaviour close to the theoretical transition point.

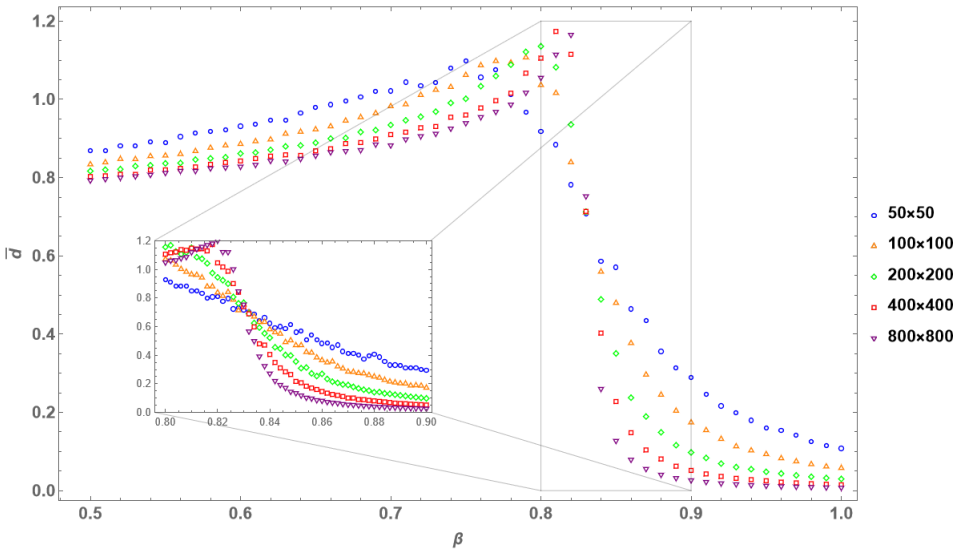

Figure 6: A measure of mean graph distance along the line $\mu = -\beta$ around the transition point. Inset plot is a less coarse sampling around the transition point.

7 Conclusion and Future Work

This simple model that flows naturally from the definition of shortcuts and strains gives 5 distinct phases seperated by different percolation phase transitions, with the amorphous phase having 3 regimes, one of which is suspected to be geometric in some sense. This is essentially a free Hamiltonian, since no interactions have been added. The natural interaction to consider is to have a term in the hamiltonian proporional to the absolute differences of vertex degrees of vertices connected by a single edge. For a positive coupling constant this should 'encourage' phases with homogonous vertex degree. Another possible direction to take this is to further investigate different values, and on different lattices or in different dimensions. On the theoretical side it would be interristing to formalise how temeprature and chemical potential can be derived from the vertex degree distribution, which is intuitively analagous to the speed distribution of an ideal gas.

References

[1] Tomasz Konopka, Fotini Markopoulou, and Simone Severini. Quantum graphity: A model of emergent locality. *Phys. Rev. D*, 77:104029, May 2008.

[2] Dmitri Krioukov. Clustering implies geometry in networks. *Phys. Rev. Lett.*, 116:208302, May 2016.

[3] Massimo Ostilli and Ginestra Bianconi. Statistical mechanics of random geometric graphs: Geometry-induced first-order phase transition. *Phys. Rev. E*, 91:042136, Apr 2015.