Percolation

By Theo Bains (Hatfield), Tabitha Bridgman (Collingwood), Will Groves (St Mary's), Lucy Witkowski (Van Mildert)

Supervisor O. Hryniv

Computational Mathematics II Group Project Report

August 2, 2026

Plagiarism declaration

> This piece of work is a result of my own work and I have complied with the Department's guidance on multiple submission and on the use of AI tools. Material from the work of others not involved in the project has been acknowledged, quotations and paraphrases suitably indicated, and all uses of AI tools have been declared.

We, **Will Groves, Tabitha Bridgman, Lucy Witkowski, and Theo Bains** confirm that we have read, understood, and have adhered to the plagiarism declaration above.

AI Usage Declaration

- ChatGPT was used to debug some sections of code and figure presentation, e.g the 'wrapfigure' environment.
- Grammarly was used to enhance the writing style, grammar, and spelling of the report throughout.

Author Contribution Statement

Will Groves, Tabitha Bridgman, Lucy Witkowski, and Theo Bains all jointly wrote and edited the final version of the report. All authors contributed to creating section 1 (Task 1). Both Tabitha and Lucy created figures and wrote section 2 (Task 2). Tabitha created section 3 (Task 3). Lucy created section 4 (Task 4). Will created section 5 (Task 5). Theo created section 6 (Task 6 & 7). All authors contributed to the GitHub repository, produced figures, performed calculations, wrote, reviewed, and edited text. Will and Tabitha completed the meeting notes each week.

Link to GitHub Codebase

All codes used to produce the analysis and results in this report can be found in the GitHub repository: https://github.com/TheoBains/Percolation-Computational-Mathematics-II.git.

1. Definitions

Percolation is the study of connectivity within a network. On a square grid, this is visualised by *sites* occupied with probability p, where connectivity is defined by *nearest neighbours*. The following definitions will be useful throughout this paper:

Spanning Cluster A cluster is a set of connected sites, and is a spanning cluster if it connects one side to the other. If a spanning cluster exists within a system, the system is *percolating*.

Percolation Probability A system of size L spans with *percolation probability* $\Pi(p, L)$. The specific probability where this transition occurs is the *percolation threshold*, p_c.

Correlation Length The probability that 2 occupied sites separated by a distance r are connected is represented by the correlation function $g(r) = p^r = e^{-r/\xi}$ with *correlation length* ξ.

Universality *Universality* is the idea that properties of large systems are independent of dynamical details. Consequently, we can study the most simplified model, and scale independent properties after calculations.

2. The One-Dimensional Case

In one dimension, there is a spanning cluster **if and only if** all the sites are occupied, with percolation probability $\Pi(p, L) = p^L$. As $L \to \infty$, $\Pi(p, L) = 0$ if $p < 1$, and $\Pi(p, L) = 1$ if $p = 1$. From this we trivially find that the critical value is $p_c = 1$.

A cluster of size s in 1-D is a sequence of s consecutively occupied sites. The probability that a site is in a specific cluster is known as the **cluster number density**: $n(s, p; L) = p^s(1 - p)^2$.
In 1-D, this is equal to the probability that there is a cluster of size s. This highlights the rarity of large clusters unless p is very close to 1. We can estimate $n(s, p; L)$ by:

$$\hat{n}(s, p; L) = \frac{N_s}{ML^d} = \frac{N_s}{ML} \tag{1}$$

Where N_s is the total number of clusters of size s observed across M independent realisations of an $L \times L$ lattice, and d is the dimension (here, $d = 1$). We can also compute the **average cluster size:** $S(p) = \sum_s s\left(\frac{sn(s,p)}{\sum_s sn(s,p)}\right) = \frac{1+p}{1-p}$

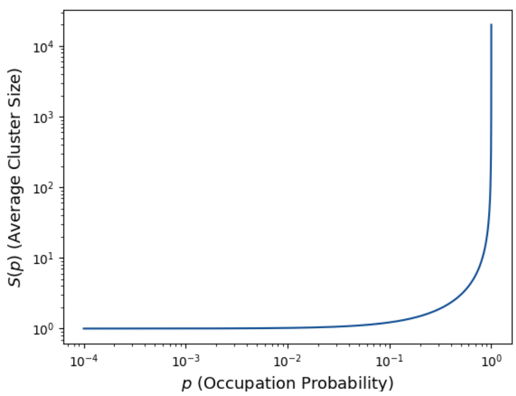
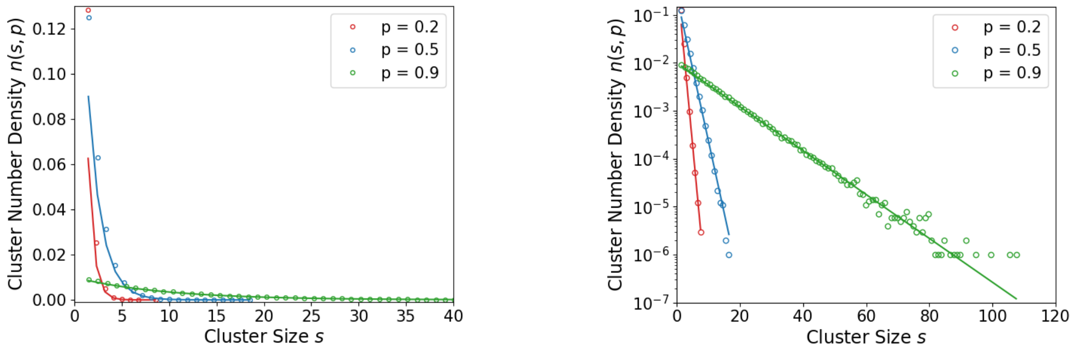


Figure 1: Average cluster size $S(p)$ as a function of p in 1-D

Figure 1 shows that for small $p \in [0.1, 0.8]$ our average cluster size $S(p)$ increases gradually, while clusters remain small. However, once p gets close to the percolation threshold p_c, we see a dramatic steepening of the curve, and much larger cluster sizes. This illustrates the divergence $S(p) \to (1 - p)^{-1}$ as $p \to p_c = 1$.

The characteristic cluster size s_ξ is the typical size of clusters in the system, and describes the scale on which the cluster number density decays. In one dimension it is given by $s_\xi = -1/\ln(p)$, which as $p \to p_c$ scales asymptotically as $s_\xi \sim (1-p)^{-1}$. Thus, s_ξ diverges as the system approaches its percolation threshold p_c.

Figure 2a is plotted on a linear axis, showing a very steep decay in the cluster number density as s increases. Figure 2b shows the same data but on a semi-logarithmic scale. This produces a linear plot and reveals some deviations on the tail end due to noise at large s.



(a) Linear axes

(b) Semi-logarithmic axes

Figure 2: A plot of the decay of the cluster number density $n(s,p) = p^s(1-p)^2$ for increasing cluster sizes s.

3. Infinite-Dimensional Case

The Bethe lattice, also called the Cayley tree, is a homogeneous tree structure in which each node has Z neighbours, where Z is called the coordination number.

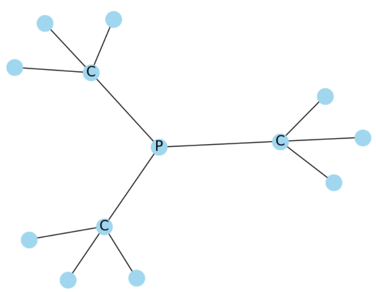


Figure 3: A Bethe Lattice with 2 generations and co-ordination number $Z = 3$

Looking at figure 3, the centre 'parent' node (labelled P) has Z neighbours, or 'children'. Moving outward to the children nodes (labelled C), now the new generation's parent nodes, each has $Z-1$ *new* child nodes, where one of those neighbours is the *old* parent node.

From the above we can see that in order to form a spanning cluster, the process of neighbouring sites must continue infinitely, and so at least one of the $Z-1$ sites must be occupied on average, i.e the occupation probability must be: $p(Z-1) \geq 1$. In which case the threshold value p_c occurs at the threshold of the above inequality:

$$p_c = \frac{1}{Z-1}.$$

In the infinite dimensional case, where each node is open with probability p, we can ask the question of whether or not an infinite cluster will form. If so, the cluster spans the infinite lattice, so is a spanning cluster, and the system percolates.

Scaling Properties

1. Crucially, the Bethe Lattice has no loops. This means nodes are connected through one route, and so an **exact** solution can be found.

2. The number of new nodes each generation grows exponentially to $(Z-1)^h$ where h is the number of generations in the lattice. In the infinite case, both the bulk and perimeter grow at this rate (as opposed to the perimeter growing faster in the finite case) and so we get the scaling relationship: Bulk $\sim$ Perimeter.

3. An infinite spanning cluster forms for $p > p_c$ and $P(p) \sim (p - p_c)^\beta$ where P is the probability for a site to be connected to the spanning cluster.

4. If, then, $p \leq p_c$, the clusters must be finite. The average cluster size, S, of these clusters is:

$$S = \frac{\Gamma}{(p_c - p)^\gamma} \sim |p - p_c|^{-\gamma}$$

where Γ is some normalisation constant and γ characterises the behaviour of S. We can see S diverges as a power law as $p \to p_c$, meaning clusters become much larger near p_c.

5. The general scaling for the cluster number density is:

$$n(s,p) = s^{-\tau} F\left(\frac{s}{s_\xi}\right) = s^{-\tau} F(s|p - p_c|^{1/\sigma}) \tag{2}$$

where the characteristic cluster size $s_\xi \sim |p - p_c|^{-1/\sigma}$.

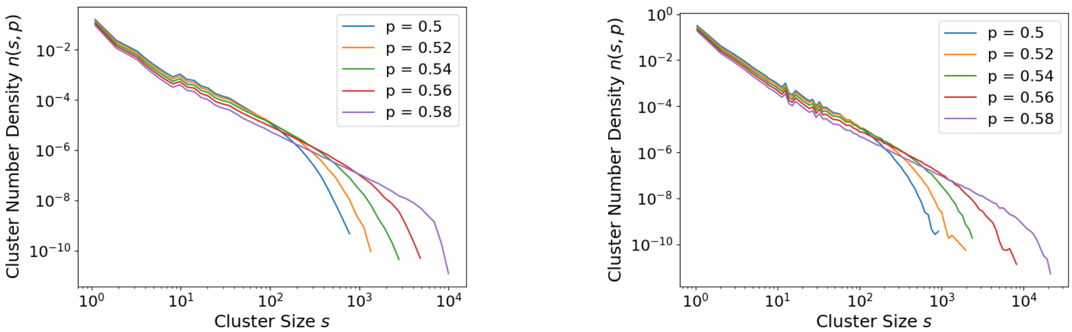
6. The correlation length is $\xi \sim |p - p_c|^{-\nu}$ which also diverges as $p \to p_c$, meaning clusters become much larger near p_c.

When a system reaches a high enough dimension, variations between sites become negligible, and so the system behaves like its own average environment, following mean-field critical behaviour. The Bethe Lattice, being infinite dimensional, therefore takes the mean-field critical exponents $\gamma = 1, \beta = 1, \sigma = \frac{1}{2}, \nu = \frac{1}{2}$, and $\tau = \frac{5}{2}$.

The above definitions are taken in part from Percolation theory using Python [2].

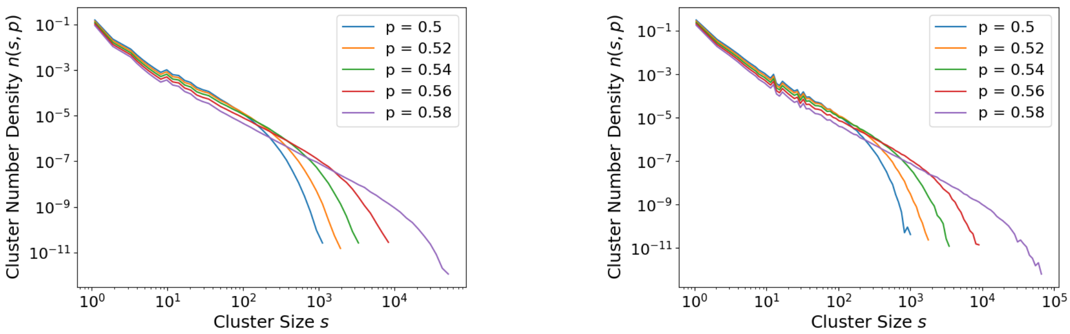
4. Estimating Cluster Number Density in 2-D

Unfortunately, we cannot examine percolation clusters analytically in 2-D. Instead, we can estimate $n(s,p)$ near to $p_c \approx 0.59275$ numerically using the same formula as in equation (1) (for $d = 2$). However, the number of clusters for each value of s decreases rapidly as cluster size increases. To reduce inaccuracy in our estimate, we can apply logarithmic binning.



(a) $a = 1.2, L = 150, [M = 2000]$

(b) $a = 1.1, L = 250, [M = 1500]$



(c) $a = 1.2, L = 500, [M = 750]$

(d) $a = 1.1, L = 1000, [M = 250]$

Figure 4: log-log plots for the logarithmically-binned $\hat{n}(s, p; L)$ over varying L and binning parameters a. Each plot shows various values of $p \in [0.50, 0.58]$. [Note that as L increases, we had to reduce M since large values for both is computationally expensive].

A larger lattice size L allows bigger clusters to form, increasing our right-tail end of s. As p approaches the critical value $p_c \approx 0.59275$, we see a power-law relationship developing; the curve is approximately linear on our log-log scale. The closer a is to 1, the narrower the bins, and hence we expect to retain a higher resolution, resulting in more noise. Increasing the log-binning parameter a makes each bin wider, decreasing the noise. Alternatively, increasing a reduces noise, producing a smoother curve. This, however, also causes a noticeable loss in resolution where we would expect a power-law to emerge.

5. Scaling Theory of Cluster Number Density

The scaling, or cut-off, function $F(x)$ dictates how the cluster number density $n(s, p)$ changes with the variable $x = s/s_\xi$. For $x < 1$, $F(x)$ is approximately constant, while it decays rapidly for $x > 1$.

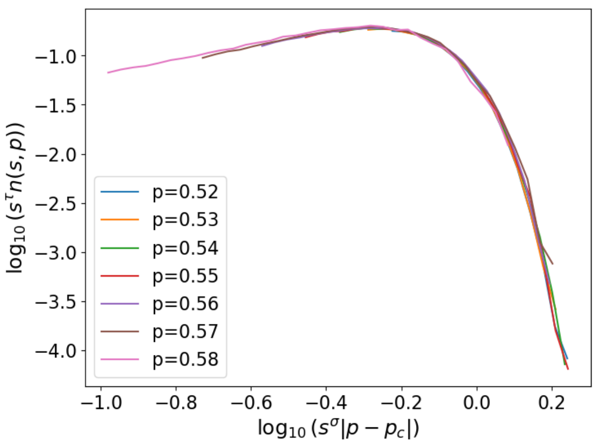


Figure 5: $d = 2, L = 850, [M = 1500]$. The curves of $\log_{10}(s^\tau n(s, p))$ against $\log_{10}(s^\sigma |p - p_c|)$, for p approaching p_c, show an approximate collapse onto a single universal curve.

Hence, as shown in equation (2), once the cluster size $s > s_\xi$, $F\left(\frac{s}{s_\xi}\right)$ causes $n(s, p)$ to decay rapidly, this behaviour is clearly visible in Figure 4, for $s_\xi \approx 2 \times 10^2$.

Plotting $s^\tau n(s, p)$ against $s|p - p_c|^{1/\sigma}$, therefore, should produce a universal curve $F(x)$, for all p near p_c. Often the transformation $F(x) = \hat{F}(x^\sigma)$ is used, as it is computationally simpler, and thus we plot $s^\tau n(s, p)$ against $s^\sigma |p - p_c|$ to produce $\hat{F}(x)$.

Hence, by trialling different combinations of τ and σ, we can plot the curves for various p values, excluding any clusters that are either too small or larger than s_ξ. Overall, the values of τ and σ that produce the best collapse onto a single curve are taken as our estimates.

Using this method of data collapse, we obtain estimates τ_{exp} and σ_{exp} which can then be compared with their current conjectured values. Moreover, using the conjectured equations [4, p. 22], we can acquire estimates for other percolation universal exponents:

$$\alpha = 2 - \frac{\tau - 1}{\sigma} \qquad \beta = \frac{\tau - 2}{\sigma} \qquad \gamma = \frac{3 - \tau}{\sigma} \qquad \nu = \frac{\tau - 1}{\sigma d}.$$

d	τ	τ_{exp}	σ	σ_{exp}	α	α_{exp}	β	β_{exp}	γ	γ_{exp}	ν	ν_{exp}
1	2	1.99	1	0.99	1	1.00	0	-0.01	1	1.02	1	1.00
2	2.05	2.04	0.40	0.43	-0.67	-0.44	0.14	0.09	2.39	2.25	1.33	1.22
3	2.19	2.22	0.45	0.45	-0.64	-0.68	0.43	0.48	1.78	1.72	0.88	0.89
4	2.32	2.40	0.48	0.49	-0.75	-0.86	0.66	0.81	1.43	1.19	0.69	0.70

Table 1: Conjectured versus experimental universal exponents values for dimensions up to $d = 4$ [2, Table 4.1], [1, Tables 7,20]. These values were calculated using code in the GitHub repository with conjectured percolation thresholds [5, Table 5], [3, Table 2].

6. Computational Analysis

The program (see GitHub) follows the steps below for a $L \times L$ lattice with occupation probability p:

1. **Lattice Generation:** A random grid is initialized where each site is occupied with probability p and empty with probability $1 - p$.
2. **Cluster Identification:** We utilize the Von Neumann neighbourhood (4-connectivity) to identify connected clusters using the `scipy.ndimage.label` function.
3. **Spanning Check:** Check for any vertically or horizontally spanning clusters.
4. **Density Calculation:** $P(p, L) = \frac{\text{Number of sites in spanning clusters}}{L^2}$
   If no unique spanning cluster exists, $P(p, L) = 0$.

Using the program, $\Pi(p, L)$ and $P(p, L)$ for $L \in \{2, 4, 8, 16, 32, 64, 128\}$ was computed.

To compare the exact values of a lattice size $L = 2$ ($N = 4$ sites), we derive the exact polynomials by enumerating all $2^4 = 16$ configurations [2, Section 1.4]:

$$\Pi(p, 2) = 4p^2 - 4p^3 + p^4, \qquad P(p, 2) = 2p^2 - p^3$$

Our simulated results match these exact polynomials well (Figure 6).

As L increases, the percolation probability $\Pi(p, L)$ approaches a step function at the critical threshold $p_c \approx 0.59275$. The curves for different L intersect near this critical point, a clear sign of continuous phase transitions.

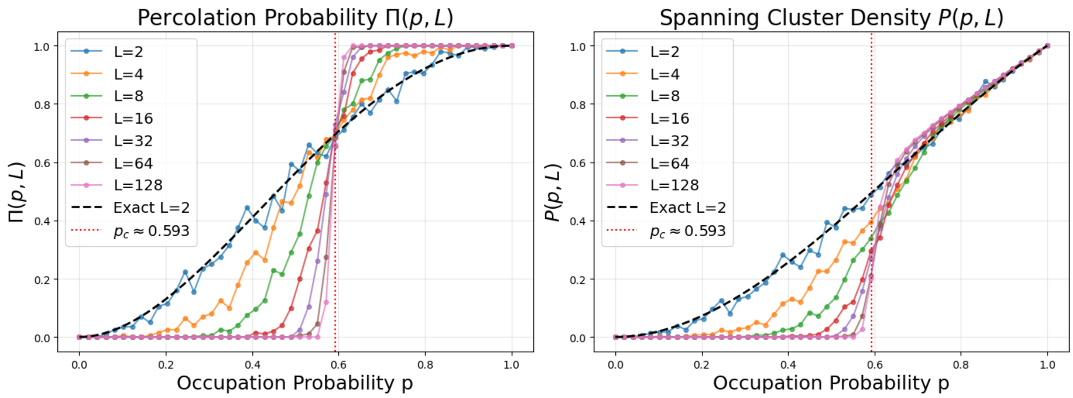


Figure 6: Simulation results for $\Pi(p, L)$ and $P(p, L)$ generated by our program

To estimate β, we perform a logarithmic regression on the simulation data for $L = 128$. The data is filtered to the scaling region $0.64 < p < 0.85$ to avoid finite size effects near p_c and saturation effects near $p = 1$.

The linear fit is shown in Figure 7. The slope of the line corresponds to β as calculated in Table 1 for 2-dimensions.

Theoretical Value (2D): $\beta \approx 0.14$.

Estimated Value: The simulations yielded $\beta \approx 0.223$ The estimated value is fairly similar to our theoretical value.

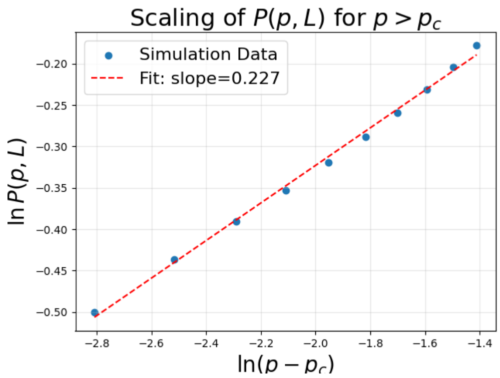


Figure 7: Log-log plot of spanning density vs. distance from threshold $(p - p_c)$.

References

[1] M. K. M. Borinsky, J.A. Gracey and O. Schnetz. Five-loop renormalization of ϕ^3 theory with applications to the lee-yang edge singularity and percolation theory. Physical Review D, 103(11):116024, 2021.

[2] A. Malthe-Sørenssen. Percolation theory using Python, volume 1029 of Lecture Notes in Physics. Springer, Cham, [2024] ©2024.

[3] S. Mertens and C. Moore. Percolation thresholds and fisher exponents in hypercubic lattices. Physical Review E, 98(2):022120, 2018.

[4] D. Stauffer. Scaling theory of percolation clusters. Physics Reports, 54(1):1–74, 1979.

[5] J.-P. L. Xiao Xu, Junfeng Wang and Y. Deng. Simultaneous analysis of three-dimensional percolation models. Frontiers of Physics, 9(1):113–119, 2014.