# Complexity Project:
# The Oslo Model

CID: 01708553

19th February 2023

***Abstract:*** ***A simulation of the Oslo Model was built and investigated for system sizes of 4, 8, 16, 32, 64, 128 and 256. It was found that the height, height probability and avalanche size probability all demonstrate scale-free behaviour and their characteristic scaling functions were found. The difference between theory and data that corrections to scaling at small system sizes was also explored, with a possible discrepancy of up to 8%.***

**Word count: 2483**

# 1 Introduction

A system self-organises if it can reach an equilibrium state with no fine-tuning of external parameters. Criticality in a system is where measured parameters follow the same statistical model regardless of the size of the system. This is known as scale-free behaviour. There are also characteristic cutoffs that increase with system size. The Oslo Model, a version of the $d = 1$ BTW model displays self-organised criticality. The system has a defined length, divided into individual sites. Rice grains are dropped at the $i = 1$ site one by one. When the slope, the difference between adjacent heights, exceeds a threshold, an avalanche is induced. This threshold, $z^{th}$, is determined by a random probability distribution. The simplest of which is where $z^{th}$ equal to 1 for all sites. The distribution used in the majority of the analysis here is $z^{th} = [1, 2]$, with $P(z^{th} = 1) = P(z^{th} = 2) = 1/2$. It is worth noting that this does not mean that $P(z = 1) = P(z = 2)$. This is because a $z = 1$ site will never relax but a $z = 2$ site will have a 50% possibility of relaxing.
A simulation of this model has been built and investigated to further explore the characteristics of self-organised criticality.

# 2 Building and Testing

The first test of the simulation was to see if the expected behaviour for the $p = 1$ case could be recovered. This system only has one recurrent configuration and once the steady state is reached, all avalanche sizes are equal to the system size. In Figure 1 this recurrent configuration can be seen, produced by adding grains to system sizes of 16 and 32 with $p = 1$ until they reached their steady state.

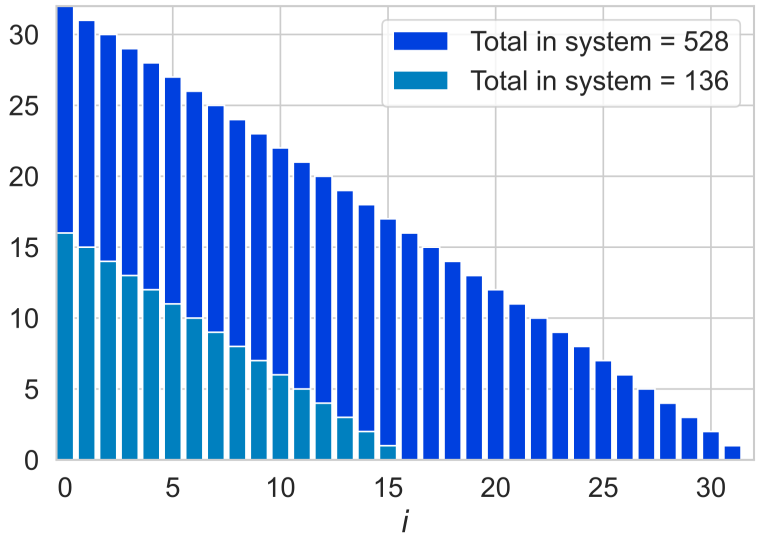



Figure 1: *Visualisation of the steady state configuration for $p = 1$ systems of $L = 16$ and $L = 32$. Grains are arranged like a stair-case and an added grain will cause each site to relax successively and then fall out at $i = L$*

The next tests were performed on $p = 1/2$ systems which are used in the rest of the analysis. It was verified that the piles reached a steady state when driven for a sufficient period of time. Figure 2 demonstrates how this was done by monitoring the total number of grains in piles of $L = 16$ and $L = 32$ over time. It can be seen that after an initial transient period, the piles

remain at a constant average total size.

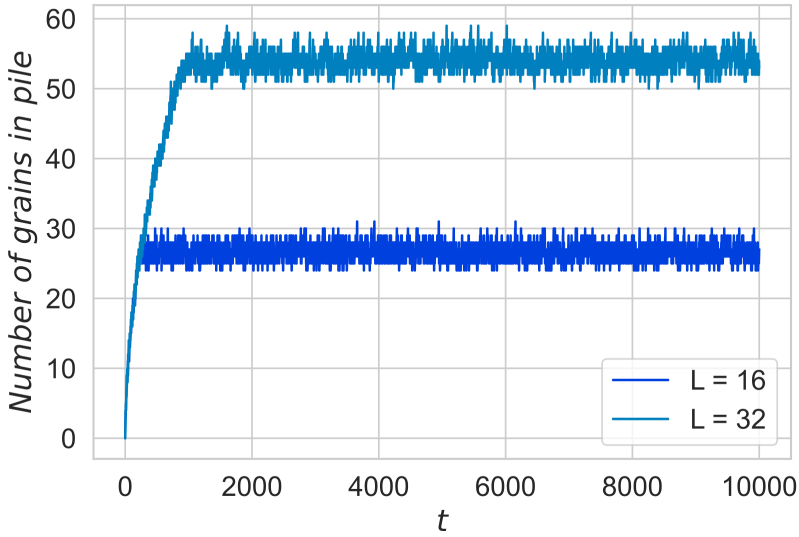



Figure 2: *Total number of grains in piles of length 16 and 32 against time. Transient and steady state stages are clearly shown.*

Finally, the average height of the $i = 1$ site was measured for the same system sizes as before, shown in Table 1. Since these agreed with the expected values, the simulation was deemed to be working correctly.

Table 1: *Average heights of the $i = 1$ site*

| System Size | Average Height | Expected Value |
|---|---|---|
| 16 | 26.54313 | 26.5 |
| 32 | 53.88213 | 53.9 |

# 3 Results and Discussion

## 3.1 Height Analysis

### 3.1.1 Initial Height Analysis

The height of a pile, $h(t; L)$ is the height of the $i = 1$ site. Since $z_i = h_i - h_{i+1}$, the height can be found by,

$$h(t; L) = \sum_{i=1}^{L} z_i, \tag{1}$$

where $z_i$ are the slopes at sites $i$. The heights of system sizes L = 4, 8, 16, 32, 64, 128 and 256 were all measured against time, shown in Figure 3. Each of the systems were driven a total of 60000 times. While the height is consistently increasing, it can be assumed the systems are in transient configurations as the total number of grains in the pile is constantly growing. Once the curves becomes horizontal, the systems are in the set of recurrent configurations. This is

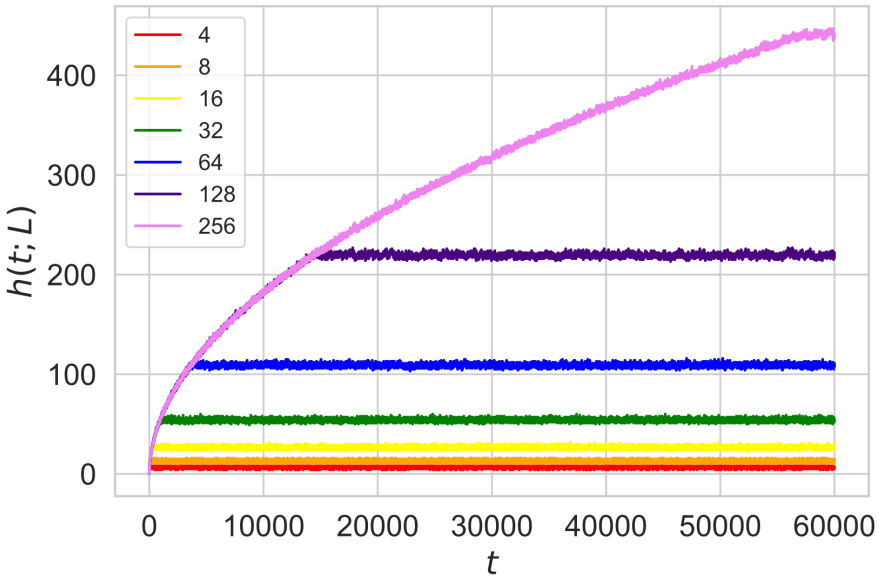



Figure 3: *Height of the piles against time. The horizontal sections show when the systems have reached a steady state as the height is at an approximately value. The transient stage can be seen in the curved sections. Each system size follows the same curve, with the time spent in transient configurations increasing with system size.*

a statistically stationary state where the number of grains added to the system is equal to the number leaving. The $L = 256$ system spends the majority of that time in transient configurations, only reaching its steady state towards the end. Figure 3 also shows that the time taken for the system to reach a steady state clearly increases with system size.

The cross-over time for a system is the number of grains in the system before an added grain causes an avalanche. Alternatively but equivalently, it is the number of times the system is driven before the first grain falls. It can be calculated using,

$$t_c(L) = \sum_{i=i}^{L} z_i \cdot i. \tag{2}$$

The average crossover time for the systems mentioned previously was estimated by recording $t_c$ a total of 10 times and taking the mean . The values found from this are shown in Table 2.

Table 2: *Average cross-over time for the different system sizes*

| **System Size** | **Average Cross-over Time** |
|---|---|
| 4 | 15.3 |
| 8 | 53.9 |
| 16 | 218 |
| 32 | 862 |
| 64 | 3480 |
| 128 | 14000 |
| 256 | 56100 |

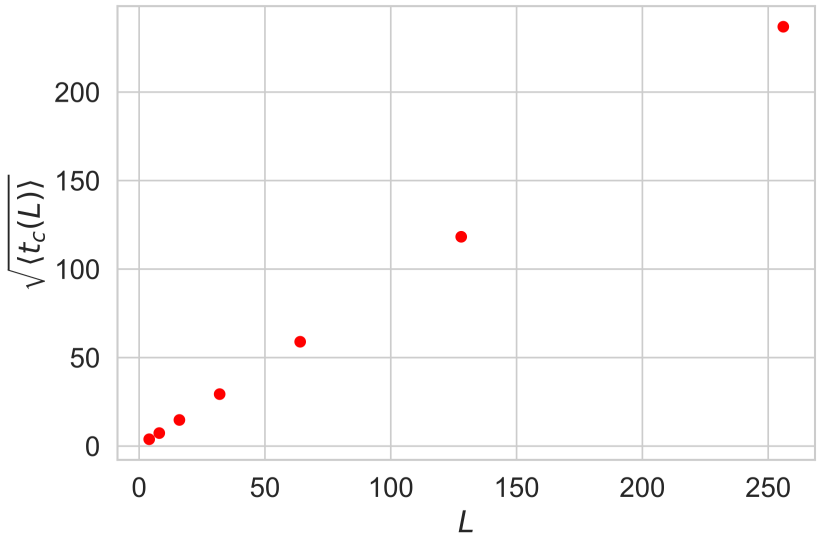



Figure 4: *Square root of the average cross-over times against system size. The distribution of the points, particularly for the larger four system sizes implies an approximately quadratic relationship.*

The square root of $\langle t_c(L)\rangle$ against $L$ can be seen in Figure 4. For L = 32, 64, 128 and 256, the data points clearly follow an approximately linear relationship, implying $\langle t_c(L)\rangle$ scales as $L^2$.

In order to explain this relationship, the scaling of the average height in the steady state with system size must first be deduced. This is given by,

$$\langle h(L)\rangle_t = \sum_{i=1}^{L} \langle z_i\rangle_t, \tag{3}$$

where $\langle z_i\rangle_t$ is the average slope at site $i$. However, since the slope of a site is independent of system size, a mean of $\langle z_i\rangle_t$ can also be taken over all of the sites, which will be a number between 1 and 2, denoted by $Z$. This can be written as $\langle h(L)\rangle_t = Z\sum_{i=1}^{L} 1$, which means that $\langle h(L)\rangle_t \propto L$. The average over repetitions of the crossover time can be written as,

$$\langle t_c(L)\rangle = \sum_{i=i}^{L} \langle z_i\rangle \cdot i. \tag{4}$$

Following similar logic as the time-average of the heights, $\langle z_i\rangle$ can be removed from the sum. This leaves $\langle t_c(L)\rangle \propto L(L+1)/2$ which will scale as $L^2$ for large $L$, as expected from the data.

### 3.1.2 Data Collapse of Smoothed Heights

Each system was driven 5000 times past its average cross-over time, ten times. In order to reduce noise in the data, averages were taken over these repetitions according to,

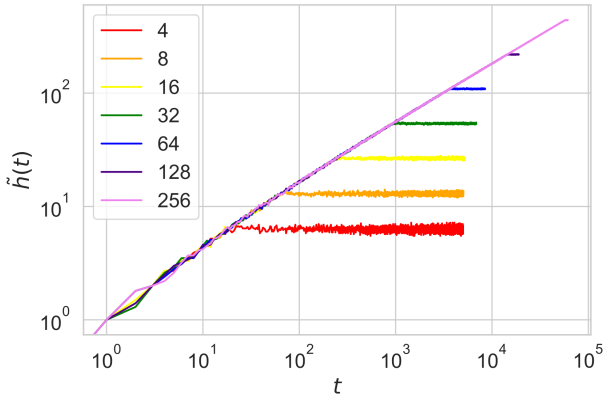
$$\tilde{h}(t;L) = \frac{1}{10}\sum_{j=1}^{10} h^j(t;L), \tag{5}$$

where $h^j$ denotes the $j^{th}$ repetition.

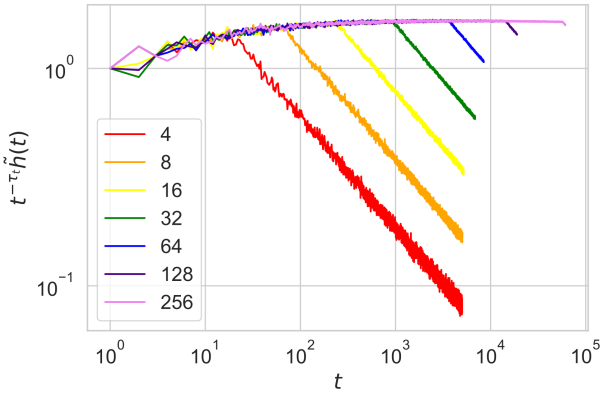
A logarithmic plot of the smoothed heights is shown in Figure 5a. There is a the linear relationship during the transient stages, implying an exponentially increasing relationship between the height of the pile and time, i.e. $\tilde{h}(t;L) \propto t^{\tau_t}$. In the steady state, the horizontal line implies that $\tilde{h}(t;L)$ is no longer changing in time. These features prompt a data collapse attempt using the ansatz,

$$\tilde{h}(t;L) = t^{-\tau_t}\mathcal{F}(t/t_c). \tag{6}$$

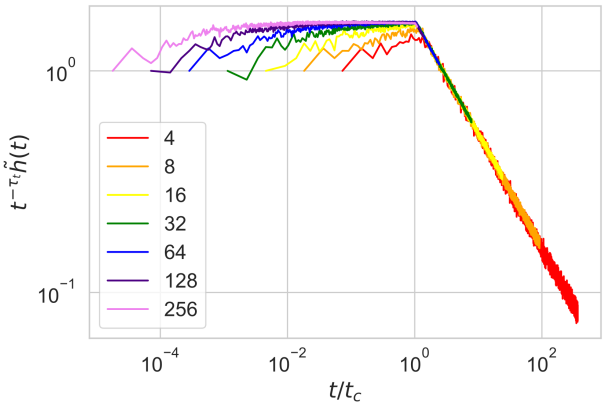
Using this prediction, $\tau_t$ is the gradient of the straight line in the logarithmic plot, found to be $0.51072 \pm 0.00002$ by performing a linear fit. The validity of the prediction was testing by multiplying the smoothed height by $t^{-\tau_t}$. If correct, this product should have no time dependence during the transient. In Figure 5b, the plot of this against time shows that the transient stage is now horizontal, implying that this rescaling has indeed removed the time dependence.




(a) *Logarithmic plot of smoothed height against time. The transient stages appear linear with the distributions becoming horizontal once the system reaches its steady state.*




(b) *Rescaled height against time. The transient stage is now horizontal, implying it is unchanging with time .*




(c) *Data collapse of height and time. Steady state regions now all fall on the same linearly decreasing function.*

Figure 5: *Process for performing a data collapse on the smoothed height data. Plots were made at each stage to inspect the result of the rescaling. The final result, shown in (c) shows that the same behaviour is observed in the steady state, regardless of system size.*

The crossover time have previously been found to scale quadratically with $L$, therefore time was divided by $L^2$ to remove the dependence on system size and collapse into one scaling function. Figure 5c shows the result of this, which is the final data collapse produced from this

method.

From processing the data, it has been possible to align the steady state regions of each system, demonstrating that the behaviour observed is scale-free. The transient stages, however, while following a similar pattern, do not align as well. This is likely due to how the smaller system sizes effect the scaling relations used. To verify this, the same procedure could be carried out, only using much larger system sizes to see if better alignment can be achieved.
The scaling function itself can be divided into two regions, transient and steady state with a transition at $x = 1$, where $x = t/t_c$. The transient stage, $x \ll 1$, is approximately horizontal, implying it is independent of the rescaled time. This corresponds to an exponentially increasing relationship between height and time during the transient. The steady state region, $x \gg 1$ decays very quickly with $t/t_c$. This means the height will reach a finite value and higher moments of the height can exist.

### 3.1.3 Corrections to Scaling

In the limit that $L \gg 1$ and $t \gg 1$, the scaling relations with system size tend to those found in Section 3.1.1. However, when these limits are not met, correction to scaling effects can be observed. More generally then, the average height of the piles once they have reached a steady state have the distribution,

$$\langle h(t; L)\rangle = a_0 L(1 - a_1 L^{-\omega_1} + a_2 L^{-\omega_2} + ...), \tag{7}$$

where the $\omega_i$ are positive. This recovers the linear relationship for large values of L as the higher order terms will tend to 0, however, for smaller system sizes, like some used here, there are noticeable effects. Fitting the average heights against the system sizes linearly gives $a_0 = 1.714 \pm 0.003$. Figure 7 shows how the the percentage error from this approximation scales with system size and it is apparent that for $L < 100$ correction to scaling is significant. For a system size of 4, the error could be up to 8% if only a linear distribution is assumed, while is is close to 0% for $L = 128$.

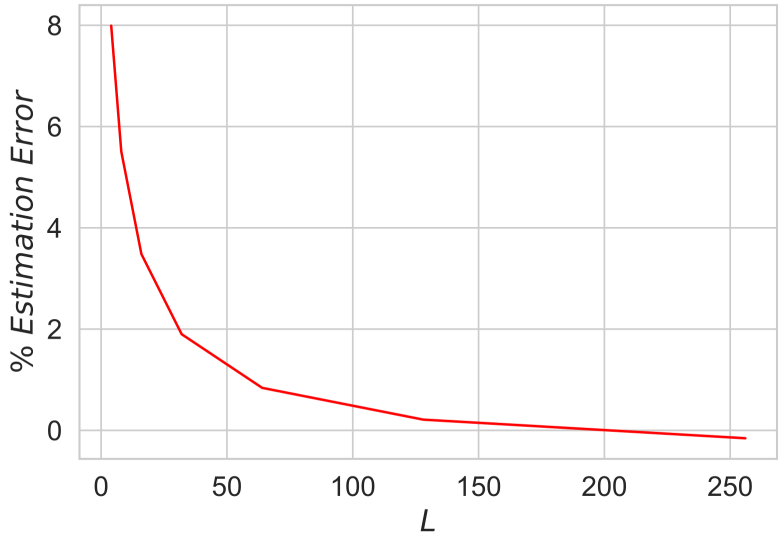



Figure 6: *Percentage difference between data and first order, linear fit. The error increases steeply as system size decreases.*

To find a more accurate representation for the data, the next order in Equation 7 was accounted for. This was done by dividing the data by the linear approximation and then fitting $1 - a_1 L^{-\omega_1}$ to the result. From this method, values of $a_1 = 0.22 \pm 0.08$ and $\omega_1 = 0.71 \pm 0.03$ were found. To test if this procedure had reduced the error, chi-squared tests were performed, shown in Table 3 along with their p-values. Although the linear approximation was reasonable, it is apparent from the reduction in $\chi^2$ and the increase in the p-value that accounting for the higher order scaling with system size has lead to a much better representation of the data.

Table 3: $\chi^2$ *values for the corrections to scaling*

| **Distribution** | $\chi^2$ | **p-value** |
|---|---|---|
| $a_0 L$ | 0.1484569 | 0.9999355 |
| $a_0 L(1 - a_1 L^{-\omega_1})$ | 0.02243885 | 0.9999998 |

The improvement can also be seen in Figure 7. From looking at the full range, the difference is difficult to see, but a close-up on the smaller system sizes shows the linear fit passing over the data but the higher order estimation going through it.

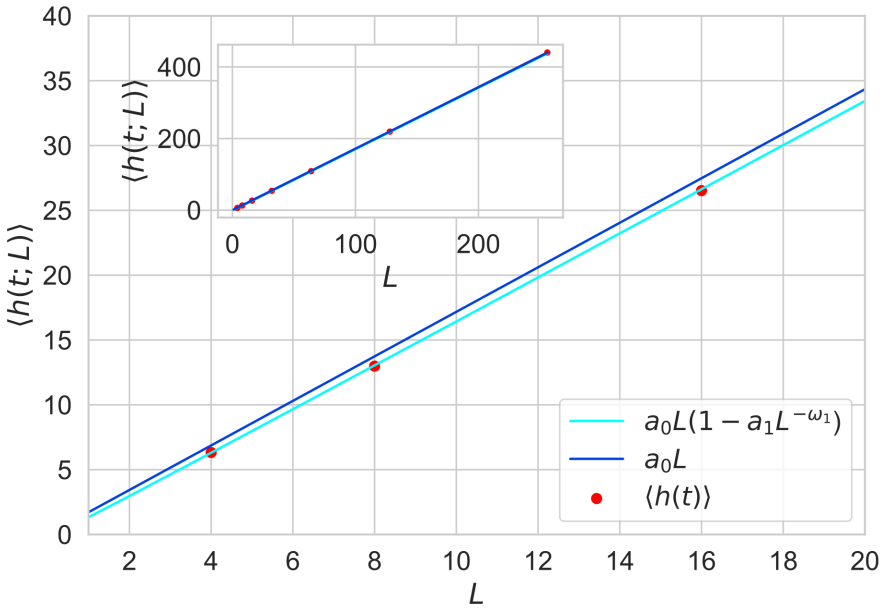



Figure 7: *Comparison of the fits from the linear and next order distributions. Overlay is the full range while the main plot is a closeup on the smaller system sizes where the difference is most apparent.*

### 3.1.4 Height Probability

The standard deviation of the height with time once in the steady state is given by,

$$\sigma_h = \sqrt{\langle h^2 \rangle - \langle h \rangle^2}. \tag{8}$$

Since $\langle h \rangle^2$ scales as $L^2$ and $\langle h^2 \rangle$ scales as $L(L+1)$, $\sigma_h$ is expected to be proportional to $\sqrt{L}$. Upon fitting the data, it was found $\sigma_h \propto L^{0.55 \pm 0.02}$. The difference between the prediction and the experimental data can be attributed to smaller system sizes causing corrections to scaling. The average slope of the system can be found by dividing the average height by the system

size. Since in the limit $L \longrightarrow \infty$, the average height is proportional to $L$, the average slope will tend to a constant. The standard deviation of the slope, $\sigma_z$, will be proportional to $1/\sqrt{L}$.

The height probability is defined as the number of observed configurations with a given height divided by the total number of configurations. Each height is a sample of the average slope multiplied by the system size. The slopes are are independent, identically distributed random variables, so the expected probability distribution is Gaussian, given by,

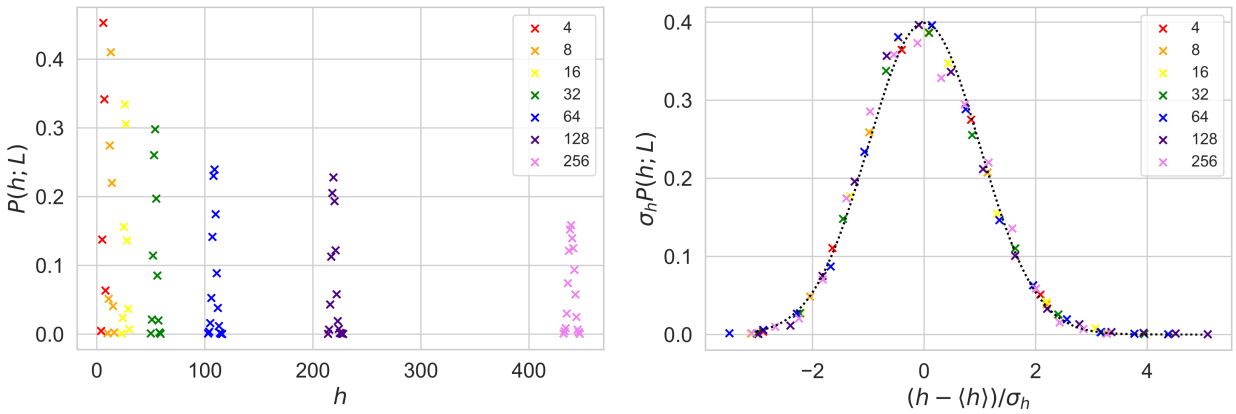
$$P(h; L) = \frac{1}{\sigma_h \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{h-\langle h \rangle}{\sigma_h}\right)^2} \tag{9}$$

due to the central limit theorem. The mean of the distribution will be the average height. The standard deviation is found by,

$$\sigma_h = L\frac{\sigma_z}{\sqrt{N}}, \tag{10}$$

where $N$ is the number of iterations in the steady state. This means, as found previously, $\sigma_h$ will be proportional to $\sqrt{L}$, but it also dependent on the time spent in the steady state. When $t \gg 1$, $\sigma_h$ will tend to 0.

In Figure 8a the raw probabilities have been plotted. As expected, they appear Gaussian, peaking around $\langle h \rangle$ for each distribution.




(a) *Height probability distributions of all system sizes. Curves appear to get lower and wider as system size increases, agreeing with the assumed $\sigma_h$ relationship with $L$.*

(b) *Data collapse of the height probabilities. Overlayed is a Gaussian distribution with mean of 0 and standard deviation of 1. Points follow this closely other than a slight left skew that can be seen at the peak.*

Figure 8: *Height probability data, both raw and collapsed.*

In order to determine if each system size follows the same probability distribution, a data collapse was performed. The probability was multiplied by $\sigma_h$ and the height was rescaled according to,

$$h' = \frac{h - \langle h \rangle}{\sigma_h}. \tag{11}$$

Theoretically, the probabilities would all collapse onto a Gaussian of mean 0 and standard deviation 1. This collapse is shown in Figure 8b. All the points appear to fall into the same distribution, implying this process was a success.

The overlayed plot in Figure 8b is the expected Gaussian distribution. The collapsed data appears to mostly follow its shape however there are some discrepancies. The peak of the data

Table 4: *Probabilities sums for the upper and lower halves of the distributions*

| **System Size** | $\mathbf{P(h' < 0)}$ | $\mathbf{P(h' > 0}$ |
|---|---|---|
| 4 | 0.595 | 0.405 |
| 8 | 0.323 | 0.673 |
| 16 | 0.514 | 0.486 |
| 32 | 0.387 | 0.603 |
| 64 | 0.444 | 0.555 |
| 128 | 0.596 | 0.403 |
| 256 | 0.549 | 0.450 |

is slightly negatively shifted and tail is longer to the right than the left. A true Gaussian should be symmetric.

In Table 4 are the probabilities of $h'$ being less than or greater than 0. From a Gaussian distribution, it is expected that both these probabilities should be equal to 1/2 but this is not the case in the data. A Gaussian distribution is continuous while the probabilities here are discrete as the height can only take integer values. This could be the cause of the difference in observed and predicted distributions, or it could be that the assumed distribution is incorrect. The number of different heights observed in the steady state, and therefore the number of data points in the distribution increases with system size, with only 5 for $L = 4$ but 16 for $L = 256$. While an actual continuous distribution could not be achieved, a sufficiently large number of points would given a reasonable approximation. The same process of calculating and scaling the height probabilities could be done for a larger system, at least 1 order of magnitude greater than those here and the resultant probability distribution analysed to see if the same discrepancies are observed. If that were the case it would imply that the predicted Gaussian actually is not the best representation of the data.

## 3.2 Avalanche size probability

The avalanche size probability is the number of avalanches of a given size, $s$, divided by the total number of avalanches observed during the steady state. This includes $s = 0$ avalanches, where a grain is added but does not induce any sites to relax. A data collapse according to the finite-size scaling ansatz,

$$\tilde{P}_N(s; L) = s^{-\tau_s}\mathcal{G}(s/L^D), \tag{12}$$

is sought with the scaling parameters $\tau_s$ and $D$ found using two different methods. For both methods, the systems have been run for 20000 iterations past their cross-over time.

### 3.2.1 Log-Binning

Due to limitations on the number of iterations possible, the avalanche size probabilities have been log-binned in order to reveal the underlying probability density function. This is done by binning the data in exponentially increasing bin widths, such that the $j^{th}$ bin is given by $[a^j, a^{j+1}]$. The value of $a$ used, 1.2, has been taken from lecture notes.

A comparison of the raw and binned data can be seen in Figure 9 for $L = 32$. Taking a larger $N$ would result in the same distribution seen in the binned data, however this is a more efficient method to attain similar results, particularly given that the $N$ required scales with system size.

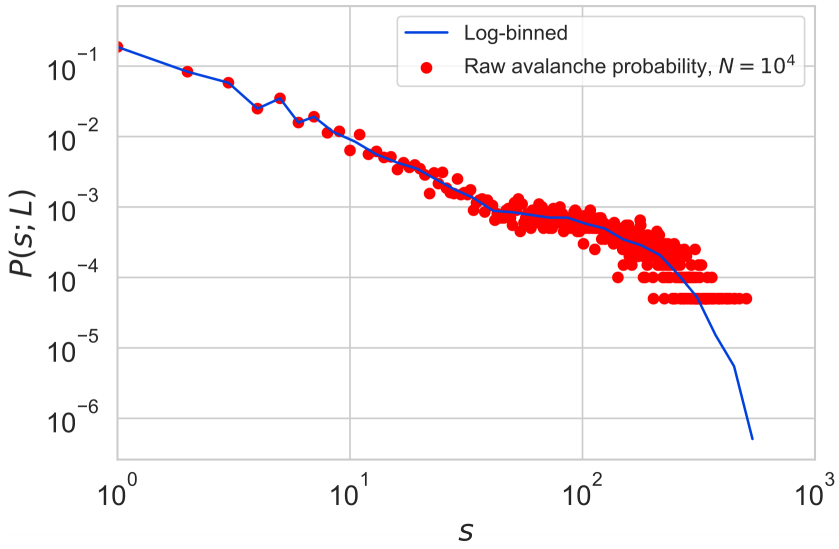



Figure 9: *Log-binned versus raw avalanche size probability for* $L = 32$ *for* $10^5$ *iterations. The binning procedure has revealed the characteristic probability distribution.*

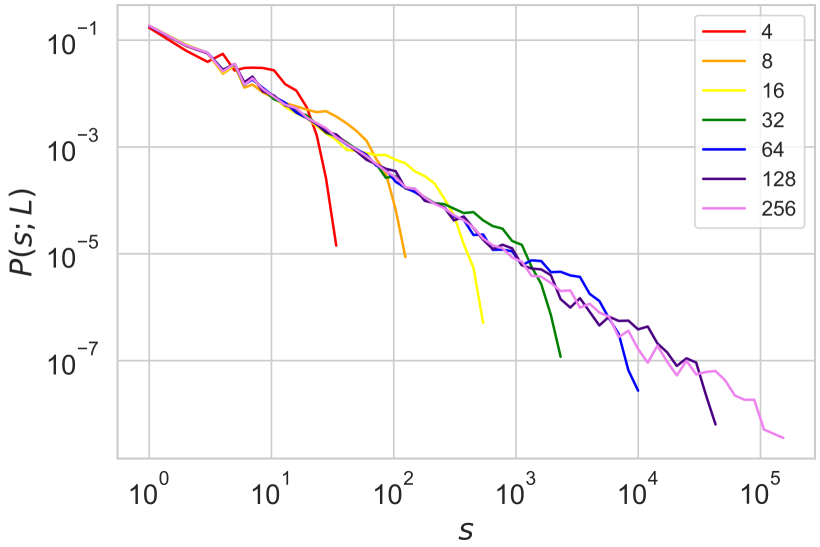



Figure 10: *Log-binned probabilities for all system sizes. The same linearly decreasing relationship can be seen throughout. Each line also has a distinctive cutoff size where the probability starts rapidly decreasing.*

In Figure 10 the binned data is shown. The probabilities all follow the same linear decay until reaching a cutoff size which depends on $L$. For $L = 256$ the cutoff is less pronounced than it is for the other system sizes.

The value of $\tau_s$ from Equation 12 was found by fitting a straight line to the linear section of the binned probabilities. The fit was performed on all the system sizes simultaneously in order to maximise the number of data points being fitted to reduce uncertainty. We know that $\langle s \rangle \propto L^{D(2-\tau_s)}$, for full derivation see [1]. Since for the Oslo Model, $\langle s \rangle \propto L$, we have,

$$D(2 - \tau_s) = 1. \tag{13}$$

Using Equation 13, the value of $D$ was found which can be seen, along with $\tau_s$ in Table 5.

Table 5: *Scaling values from log-binned data*

| **Scaling Value** | **Result from Log-Binning** |
|---|---|
| $\tau_s$ | $0.150 \pm 0.007$ |
| $D$ | $2.042 \pm 0.009$ |

These values were then used to produce the data collapse seen in Figure 11. The key features of the distributions have been aligned, verifying the ansatz in Equation 12. There is some noise present, particularly on and around the bump. This would likely be improved by increasing number of iterations in the steady state.

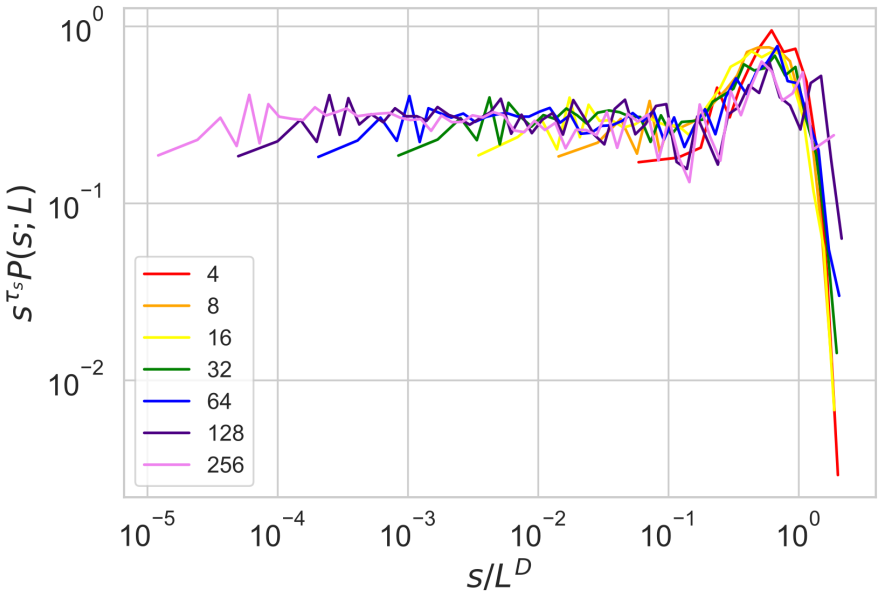



Figure 11: *Data collapse of avalanche size probabilities using log-binned data. The characteristic cutoffs are now aligned for all the system sizes. The scaling function is approximately constant until a bump around the cutoff followed by a steep decay.*

### 3.2.2 Moment Analysis

The $k^{th}$ moment is defined as,

$$\langle s^k \rangle = \lim_{T \to \infty} \frac{1}{T} \sum_{t=t_0+1}^{t_0+T} s_t^k. \tag{14}$$

Finite-size scaling means that $\langle s^k \rangle \propto L^{D(1+k-\tau_s)}$. Therefore a log-plot of the $k^{th}$ moment and $L$ should give a linear relationship with a gradient of $D(1+k-\tau_s)$. Figure 12 confirms that this prediction is reflected in the data. A plot of $D(1+k-\tau_s)$ against $k$ then gives a linear relationship with a gradient of $D$ intersects with the $k$-axis at $1-\tau_s$.

Table 6 contains the values found from performing this analysis and the subsequent data collapse is shown in Figure 13. Visually comparing this and Figure 11, it appears the moment analysis has resulted in a slightly better collapse. The steeply decaying regions are more closely aligned as is the bumped section. This was expected since moment analysis uses averages to

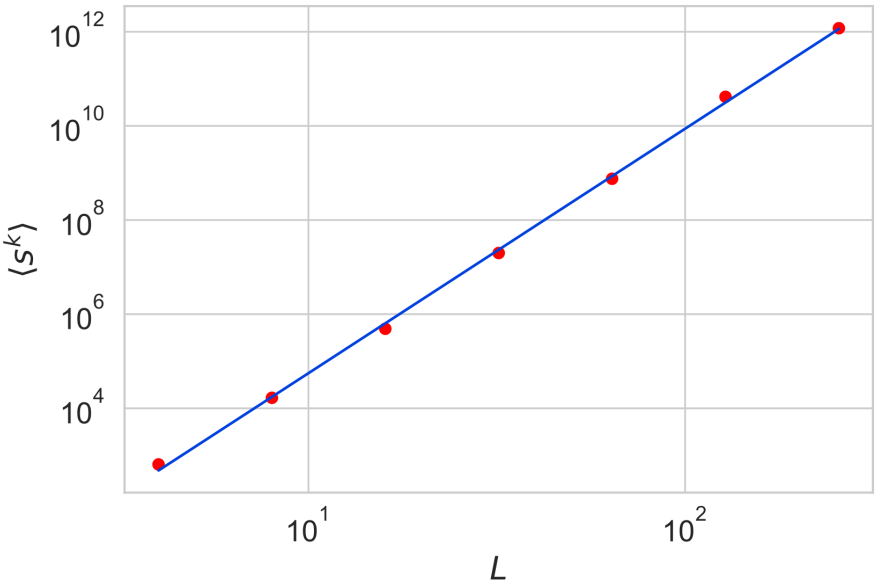



Figure 12: $\langle s^3 \rangle$ *against* $L$. *There is a clear linear relationship shown, as predicted.*

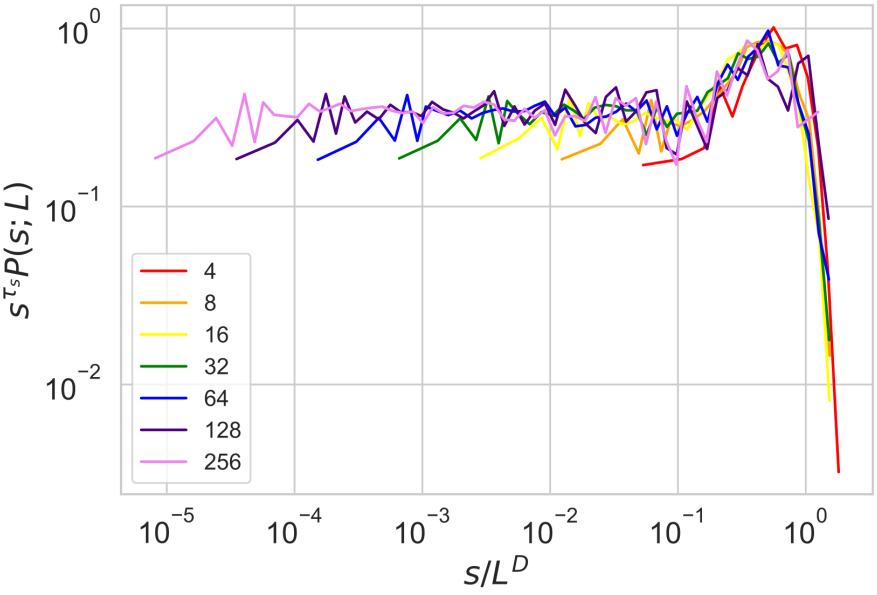



Figure 13: *Data collapse of avalanche size probabilities using moment analysis. The same characteristic shape as before has been achieved.*

Table 6: *Scaling values from moment analysis*

| **Scaling Value** | **Result from Moment Analysis** |
|---|---|
| $\tau_s$ | $1.538 \pm 0.007$ |
| $D$ | $2.11368 \pm 0.00003$ |

find the scaling values while the $\tau_s$ fitting for the log-binning required an arbitrary selection of the linear region.

# 4 Summary

The finite-size scaling properties of the Oslo Model have been explored. It was found that the height, height probability and avalanche size probability for different system sizes can all be collapsed onto single scaling functions. Correction to scaling effects were also investigated and it was shown that the relationship of average height with system size can only be approximated as linear in the limit $L \gg 1$. To further improve the investigations presented, the same analyses could be performed on larger system sizes as well as increasing the number of iterations.

# References

[1] Christensen. K. *L4 Lecture Slides, Finite-size scaling and moment analysis.*