Complexity Project: The Oslo Model

CID: 01347129

10th February 2020

Abstract: This report aimed to investigate the Oslo model. The height of the 1st pile was recorded as well as the avalanche size every time a grain was added. The simulation work was done and analysed in python. There was found to be a scaling function $\mathcal{F}(x)$ relating height to time and system size, where $h(t;L) = L\mathcal{F}(t/L^2)$ and $\mathcal{F}(x) = K\sqrt{x}$ for $x \ll 1$ amd $\mathcal{F}(x) = \langle z \rangle$ for $x \gg 1$. There was found to be corrections to scaling for small system sizes in both the height and avalanche data. The avalanche size probability $\tilde{P}(s;L)$ was found to be $\propto s^{-\tau_s}\mathcal{G}(s/L^D)$. By showing that the height distribution was not Gaussian, the assumption that slopes z_i are not independent and identically distributed variables was proven false.

Word count: Approximately 2100 words in the report (excluding front page, figure captions, table captions, acknowledgement and bibliography).

1 Introduction

The experiment aimed to built and test the Oslo model, which demonstrates self-organised criticality. This system simulates a 1D rice pile experiment where grains are confined between two plates with the length that the grains can lie on being finite. There is a wall on one side and the other side is left open. Grains are then placed repeatedly at the same location, which eventually leads to avalanches. These avalanches are found to be scale-invariant. Initially, a set of transient configurations are present until the system reaches the steady-state, when the first grain drops off the open end. After this point the configurations of the system are recurrent.

2 Implementation

2.1 Oslo model algorithm

The Algorithm of the Oslo model is as follows[1]:

1. The system is created with L sites. The slope values are $z_i = 0$, $\forall i$. All the threshold values are set to $z_i^{th} = [1, 2]$ with equal probability.

2. A grain is added at the site most to the left. Consequently $z_1 \to z_1 + 1$

3. Site are relaxed if $z_i > z_i^{th}$ according to:
   For $i = 1$:
   $$z_1 \to z_1 - 2$$
   $$z_2 \to z_2 + 1$$
   For $i = 2, ..., L - 1$:
   $$z_i \to z_i - 2$$
   $$z_{i+1} \to z_{i+1} + 1$$
   $$z_{i-1} \to z_{i-1} + 1$$
   For $i = L$:
   $$z_L \to z_L - 1$$
   $$z_{L-1} \to z_{L-1} + 1$$

4. Now reset $z_{th} = [1, 2]$, with equal probability.

5. Return to step 2.

s, the avalanche size is defined as the number of relaxations resulting from the addition of one grain. This definition includes avalanche sizes of $s = 0$.

The height of the pile is defined in equation 1, whereas the crossover time t_c is the number of grains needed for the first grain to topple off the end of the system, shown in equation 2. Here the slope $z_i = h_i - h_{i+1}$.

$$h = \sum_{i=1}^{L} z_i \tag{1}$$

$$t_c = \sum_{i=1}^{L} z_i i \tag{2}$$

2.2 Testing the simulation

This algorithm was turned into a computational simulation, which was tested using several methods.

The first test was to run the model for $L = 16, 32$ with $4t_c$ grains added once the system was in the steady-state. The average height values after each grain were averaged then compared to the standard values. The results can be seen from the table, along with the standard errors.

The second test was to check $\langle s \rangle$ after the first grain toppled off the edge of the system. If the system was initialised correctly, $\langle s \rangle = L$. The avalanche sizes for $L = 16, 32$ were calculated as shown in the table below. Note that the actual error is bigger than the error stated due to corrections to scaling. For this reason, I determined that the system passed the test.

System size	$\langle h \rangle$	$\frac{\sigma_h}{\sqrt{N}}$	$\langle s \rangle$	$\frac{s}{\sqrt{N}}$
16	26.47	0.003	15.7	1.4
32	53.79	0.02	31.5	2.1

To save time, the code was optimized in python. This was done by removing for loops and replacing them with list comprehension where possible. The data was stored in dictionaries and saved so that it could be accessed later on, so that time could be saved by not re-running the simulations.

3 Height of the Pile

Systems of sizes $L = 4, 8, 16, 32, 64, 128, 256$ were created. The number of grains added was $5t_c$ for each system. For display purposes figure 1 shows a logarithmic scale plot. It can be seen in the transient configurations, all systems follow a power-law where the height at pile 0 grows. At t_c there is a discontinuity, after which the average height seems to be constant. The transient configurations never occur more than once, as at this point the system is growing and has yet to have a grain topple off its edge.

The fit in figure 2 shows that a power-law is present for $t_c(L)$ with a power of 2.0. Looking at equation 1, if we assume that the z_i are independent and identically distributed (iid), they will have a mean value $\langle z \rangle$. Consequently, as seen in equation 3a, we find an expression for average height $\langle h \rangle$. This then turns into equation 3b, where $\langle h \rangle \propto L$.

$$\langle h \rangle = \sum_{i=1}^{L} \langle z \rangle \tag{3a}$$

$$\langle h \rangle = \langle z \rangle \, L \tag{3b}$$

From equation 3b it is possible to find the relationship between $\langle t_c \rangle$ and L. Using equation 2, once again assuming z_i are iid variables, equation 4 is formed.

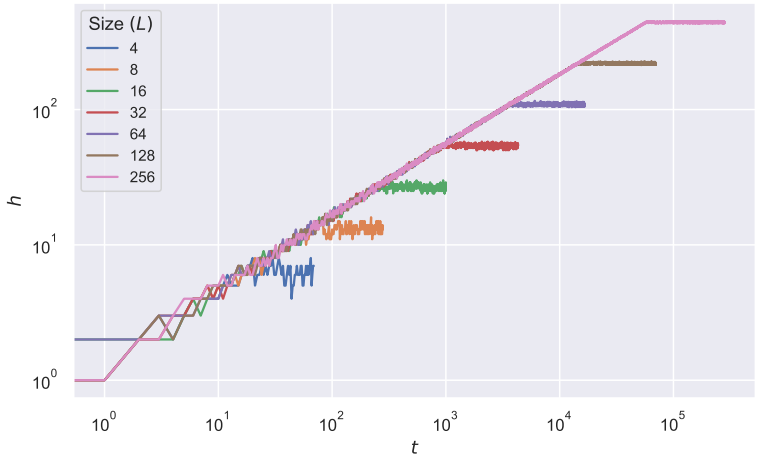

Figure 1: Logarithmic scale plot of height h vs grain added t for system sizes $L \in \{4, 8, 16, 32, 64, 128, 256\}$. This shows the heights for $5t_c$ grains added for each L.

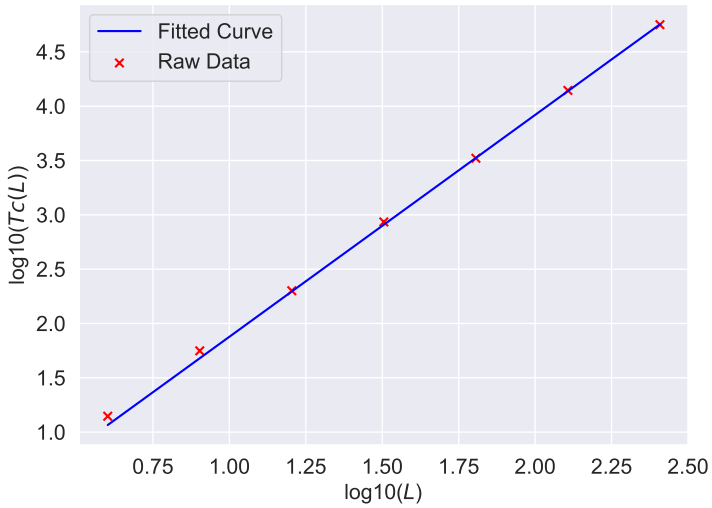

Figure 2: Logarithmic scale plot of $t_c(L)$. A straight line fit has been applied to the systems were $L \gg 1$ as complexity is displayed here. The gradient of the line is 2.01 ± 0.03, showing a power-law of approximately $t_c \propto L^2$.

$$\langle t_c \rangle = \sum_{i=1}^{L} \langle z \rangle \, i \tag{4}$$

As the contributions from larger values of i will be much bigger than for small i, as well as fact the gaps between the sum in equation 4 are small, the sum can be approximated as an integral. This integral can be solved as shown in equation 5, showing that $\langle t_c \rangle \propto L^2$. The data from figure 2 confirms this prediction, with the power-law relation as outlined previously.

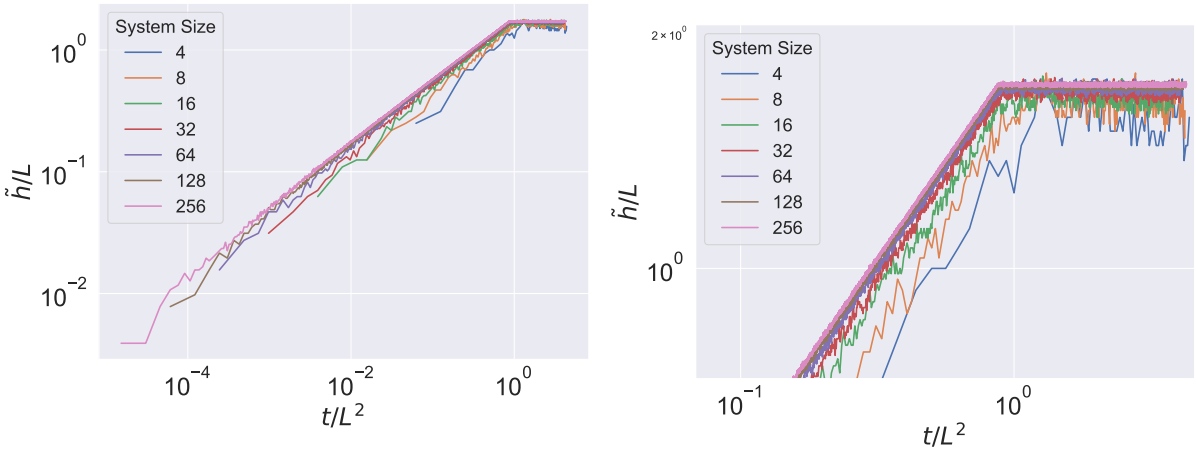

Figure 3: Data collapse for all system sizes. The data is the average of 4 independent $h(t)$ measurements for each system size, giving $\tilde{h}(t)$. On the right focuses on the area where $t = t_c$.

$$\langle t_c \rangle = \int_1^L i \langle z \rangle \, di = \frac{L^2 \langle z \rangle}{2} - \frac{\langle z \rangle}{2} \approx \frac{L^2 \langle z \rangle}{2} \propto L^2 \tag{5}$$

A data collapse can now be produced for processed height vs time. From figure 1 it is clear there is a discontinuity at $t = t_c$. From the fact that $\langle h \rangle \propto L$ and $\langle t_c \rangle \propto L^2$ for $L \gg 1$, if we plot $\frac{h}{L}$ against $\frac{t}{L^2}$ where t is the number of grains added, the data for many systems should collapse onto a single pattern. The scaling function is therefore expressed as shown in equation 6. For values of t where $t < t_c$ the scaling function follows a power-law relation as outlined previously.

$$h(t; L) = L\mathcal{F}(\frac{t}{L^2}) \tag{6}$$

For large values of the argument $x = t/L^2$, the scaling function $\mathcal{F}(x)$ should be constant with time. As $\mathcal{F}(x) = \tilde{h}(t; L)/L$, and equation 3b shows that $\frac{\langle \tilde{h} \rangle}{L} \propto \langle z \rangle$ for a system with $t > t_c$, it can be determined that for a system with $t > t_c$ we get equation 7.

For large values of x where $t < t_c$, $\tilde{h}$ is independent of L, as the system has not reached the edge with no way existing for it to know how far away the edge is. Therefore $\mathcal{F}(x) \propto 1/L$ hence $\mathcal{F}(x) \propto t/L^2$. Therefore equation 7 is valid in this region, with a constant K. Note that for $t = t_c$, $t_c/L^2 = L^2 \langle z \rangle /2L^2 = \langle z \rangle /2$.

Using the continuity condition K can be calculated. $\langle z \rangle = K\sqrt{\langle z \rangle /2}$, therefore $K = \sqrt{2 \langle z \rangle}$.

$$\mathcal{F}(x) = \begin{cases} K\sqrt{x} & \text{for } x \leq \langle z \rangle /2 \\ \langle z \rangle & \text{for } x \geq \langle z \rangle /2 \end{cases} \tag{7}$$

Figure 3 shows the result of the data collapse. Here it can be seen that for small values of L, the data collapse doesn't work as well, with the data below the expected distribution. This is evidence for corrections to scaling at these small system sizes. On figure 4 it can be seen that $\langle z \rangle$ isn't constant and is smaller for small system sizes. This is

further evidence for correction to scaling and indicates why the data for small L in figure 3 is below the expected distribution.

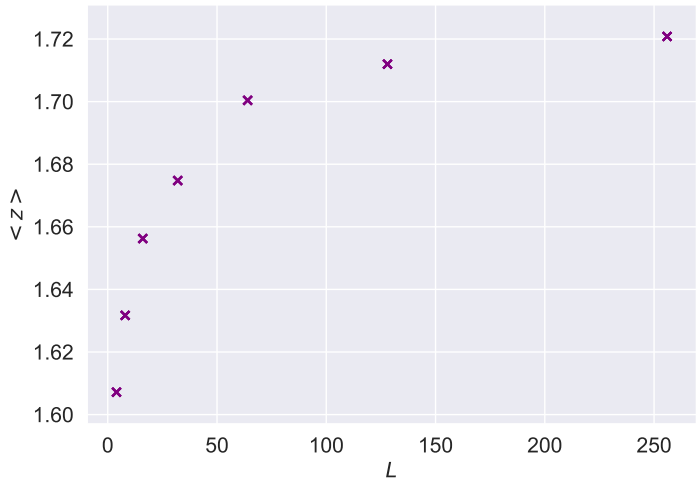

Figure 4: Plot of $\langle z \rangle (L)$. It can be seen that $\langle z \rangle$ converges as $L \gg 1$, but for small L there are corrections to scaling.

$$\langle h(t;L) \rangle = a_0 L(1 - a_1 L^{\omega_1} + a_2 L^{\omega_2} - ...) \tag{8}$$

Equation 8 shows the form of corrections to scaling that we assume. It is clear to see that for $L \to \infty$, the terms in the brackets disappear. This means that as $L \to \infty$, the value of a_0 will be the same as the value of $\langle z \rangle$, after equation 3b is applied, as is seen in figure 4.

$$1 - \frac{\langle h(t;L) \rangle}{a_0 L} = a_1 L^{\omega_1} \tag{9a}$$

$$\log_{10}(1 - \frac{\langle h(t;L) \rangle}{a_0 L}) = \log_{10}(a_1) + \omega_1 \log_{10} L \tag{9b}$$

$$h_{mod} = 1 - \frac{\langle h(t;L) \rangle}{a_0 L} \tag{9c}$$

From equations 9b and 9c it is clear that a log-log plot of h_{mod} against L should have a linear trend if a_0 is at the right value. Then the gradient of this plot would be ω_1 and the intercept would be $\log_{10} a_1$. a_0 is found by maximising the linearity of this plot. By minimising the errors on a linear fit, it was determined that $a_0 = 1.75 \pm 0.08$. The gradient of this fit is equal to ω_1, which was $\omega_1 = 0.40 \pm 0.03$. The fit is shown in figure 5.

The standard deviation of the heights was plotted on a log-log graph as shown in figure 6. It is seen that there is a power-law relationship $\sigma_h = AL^s$. The fit in figure 6 is therefore use to determine equation 10.

$$\sigma_h = AL^{0.25} \tag{10}$$

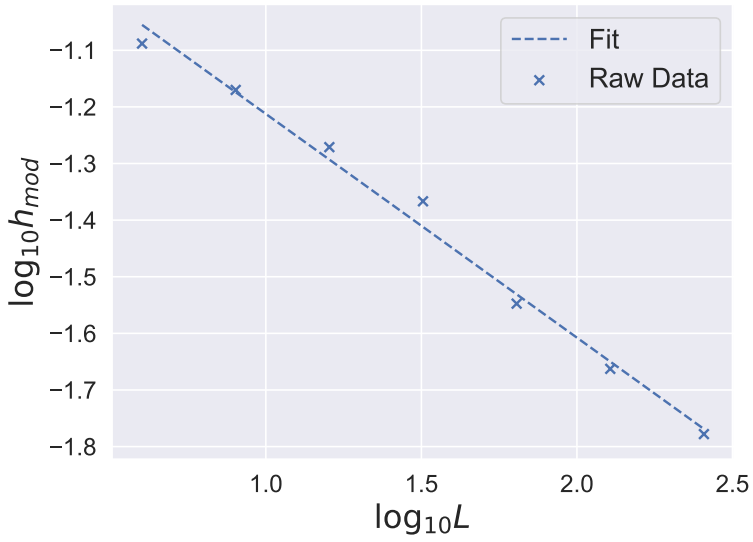

Figure 5: Logarithmic scale plot of h_{mod} against L. It was determined from this plot that $a_0 = 1.75 \pm 0.09$ and $\omega_1 = 0.40 \pm 0.07$.

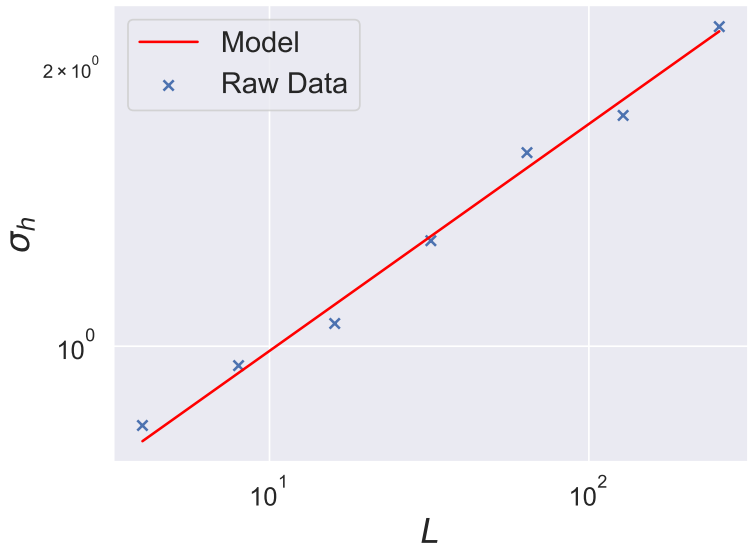

Figure 6: Log-log plot of $\sigma_h(L)$. A linear fit was done on the three highest values of L where $L \gg 1$. The power exponent was found by the gradient to be $s = 0.25 \pm 0.01$.

By using the propagation of errors and equations 3b and 10, it can be expected that the standard deviation of the slope can be determined by $\sigma_z = \sigma_h / L \propto L^{-0.75}$. This is a negative exponent, meaning that the $\sigma_z \to 0$ as $L \to \infty$. As determined previously, $\langle z \rangle \to a_0$ as $L \to \infty$.

The height probability is defined in equation 11, where $\sum_{i=1}^{L} P(h; L) = 1$. If z_i are assumed to be iid variables, when $L \gg 1$, $P(h; L)$ would be expected to have a Gaussian distribution. This is because of the Central Limit Theorem (CLT), which states that when

iid variables are summed, the normalized sum tends to a Gaussian distribution, shown in equation 12.

$$P(h;L) = \frac{\text{No. of observed configurations with height h in pile of size L}}{\text{Number of observed configurations}} \tag{11}$$

$$G(h) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(h-\langle h\rangle)^2}{2\sigma^2}\right) \tag{12}$$

To investigate if this is the case, the probability distribution was calculated for all system sizes according to equation 11. Figure 7 shows this distribution. The widening of the peaks in figure 7 is consistent with equation 10.

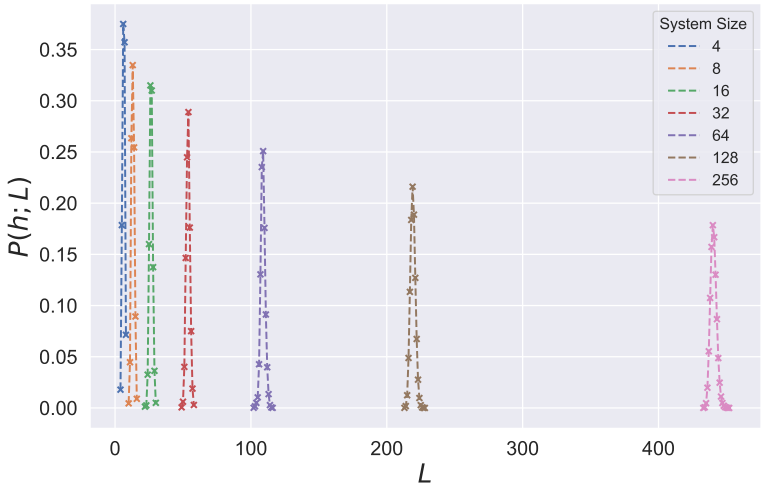

Figure 7: Plot P(h;L), here the Gaussian shape of the distribution for all system sizes can be seen. Clearly, the smaller system sizes have a higher peak and lower width than the larger systems, making sense as the area under each curve must equal one.

A data collapse can be produced for the $P(h;L)$, as σ_h and $\langle h\rangle$ are functions of L. This is done by first making the mean equal to zero, by subtracting the $\langle h\rangle$ from h. After which the standard deviation of the new plot is made equal to one, by dividing the $h - \langle h\rangle$ by σ_h. Finally, to scale the heights and ensure normalization multiply the probability by the standard deviation. This is equivalent of plotting $G(x)$ in equation 13 , where $x = \frac{h-\langle h\rangle}{\sigma_h}$The result is shown in figure 8.

$$G(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) \tag{13}$$

However, the Gaussian fit isn't perfect. The CLT predicts that the variance of the distribution is the sum of the individual variances. This means that $\sigma_h \propto L^{0.5}$. However previously, it was seen that $\sigma_h \propto L^{0.5}$ meaning that the CLT doesn't apply here. Consequently it can be concluded that the z_i variables are not independent and identically distributed.

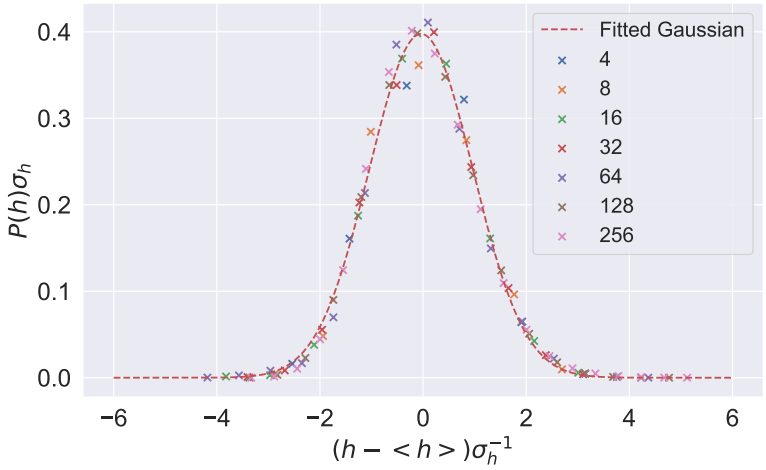

Figure 8: Data collapse of the probability function. A Gaussian distribution is fitted onto the data points and seems to work well. For larger L, the data fits more closely around the Gaussian fit.

4 Avalanche size

The original probability distribution for avalanche sizes are described in equation 14. As the total number of avalanches $N \to \infty$, the probability distribution tends to the actual avalanche size probability $P(s)$.

$$P_N(s) = \frac{\text{No. of avalanches with size } s_i = \text{s}}{N} \tag{14}$$

$$\tilde{P_N}(s^j) = \frac{\text{No. of avalanches in bin j}}{N\Delta s^j} \tag{15a}$$

$$s^j = \sqrt{s^j_{min} s^j_{max}} \tag{15b}$$

With a finite N, if there is a finite cutoff to $P(s)$ it won't be revealed. For this reason, log binning is used. Here, the bins are exponentially increasing in size and the probability is represented by $\tilde{P_N}(s^j)$ and s^j is the geometric mean of all s in bin j, as shown in equation 15b. $\tilde{P_N}(s^j)$ approximates the underlying avalanche probability however, information is lost on the exact size of s. The results for both distributions are shown in figure 9.

The curves all have a similar pattern for the log binned data. Initially, there is a negative power-law relationship, seen as a negative gradient on a log-log plot. After which there is a bump, followed by a quick drop off. The location of the bump and the drop off is a larger s for larger L. This indicates that the reason for the bump and drop off is because L is finite and for large, system-wide avalanches the finite size of the system L limits the size of those avalanches. It is also clear to see that the linear bins work far better for small L than $L \gg 1$.

The finite-size scaling ansatz is shown in equation 16 and should be valid for $L \gg 1$, $s \gg 1$. In figure 10, a data collapse is used to find the values of the avalanche dimension D and the avalanche-size exponent τ_s. In equation 16 the $s^{-\tau_s}$ characterises the negative

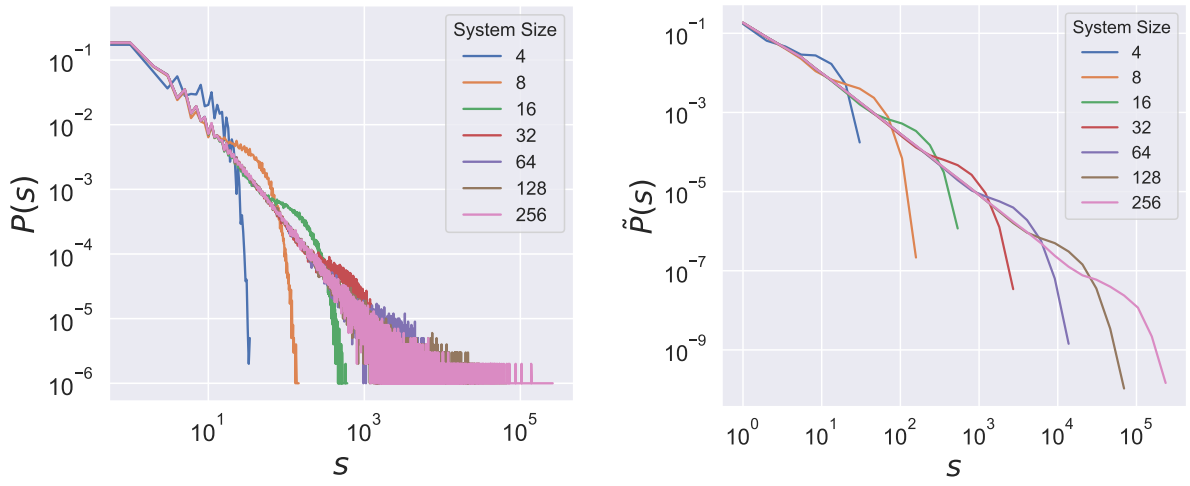

Figure 9: In the left figure the linearly binned avalanche probability distribution is shown. On the right, the log binned distribution is shown. The number of grains added was 10^6 after the system had reached the steady state. Both figures only show the avalanches when the system was in the steady state. $a = 1.5$ is used for the log binned data.

power-law, whereas the $\mathcal{G}(s/L^D)$ represents the rapid decay at the cutoff avalanche size s_c as $s_c \propto L^D$. Thus if the probability is multiplied by s^{τ_s} the gradients of the plot will align, whereas if s/L^D the cutoff will align.

$$\tilde{P_N}(s;L) \propto s^{-\tau_s}\mathcal{G}(s/L^D) \tag{16}$$

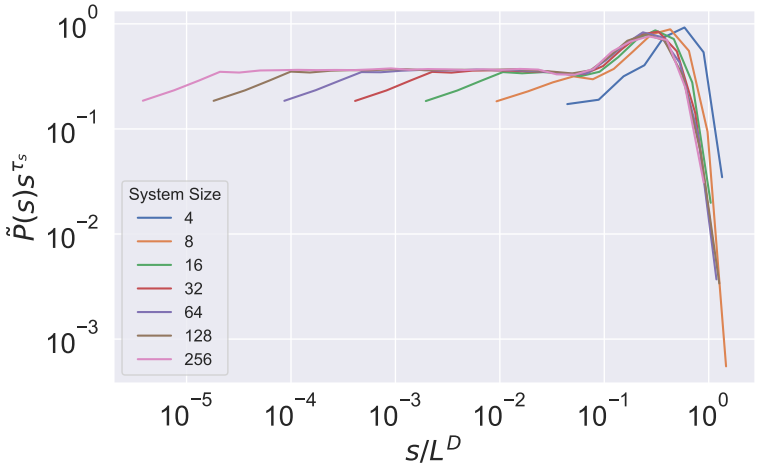

Figure 10: Data collapse of avalanche sizes for $\tau_s = 1.55$ and $D = 2.25$. Clearly, for $L \gg 1$, the collapse works significantly better than it does for smaller L.

The k'th moment of avalanche size is defined in equation 17. Here, s_t is the avalanche size at time t. The system is in the steady state as $t_0 > t_c$. Figure 11 shows the k_{th} momemnt for $k \in \{1, 2, 3, 4, 5, 6, 7, 8\}$.

$$\langle s^k \rangle = \lim_{T\to\infty} \frac{1}{T} \sum_{t=t_0+1}^{t_0+T} s_t^k = \sum_{0}^{\infty} s^k P(s;L) \tag{17}$$

Equation 18e shows that if a logarithmic scale is used of $\langle s^k \rangle$ vs L is done, the gradient of this plot will be $D(1+k-\tau_s)$. Once this is done, estimates for D and τ_s can be found by plotting $D(1+k-\tau_s)$ against k. Consequently, for $k=1$, we expect that $D(2-\tau_s)=1$, which is confirmed as the gradient for the $k=1$ linear fit is 1.00 ± 0.01.

$$\langle s^k \rangle = \sum_{0}^{\infty} s^{k-\tau_s} \mathcal{G} s/L^D \tag{18a}$$

$$\langle s^k \rangle = \int_{0}^{\infty} s^{k-\tau_s} \mathcal{G} s/L^D \text{ for } L \gg 1 \tag{18b}$$

$$\langle s^k \rangle = \int_{1/L^D}^{\infty} (uL^D)^{k-\tau_s} \mathcal{G}(u) du \text{ where } u = s/L^D \tag{18c}$$

$$\langle s^k \rangle = L^{D(1+k-\tau_s)} \int_{1/L^D}^{\infty} u^{k-\tau_s} \mathcal{G}(u) du \tag{18d}$$

$$\langle s^k \rangle \propto L^{D(1+k-\tau_s)} \tag{18e}$$

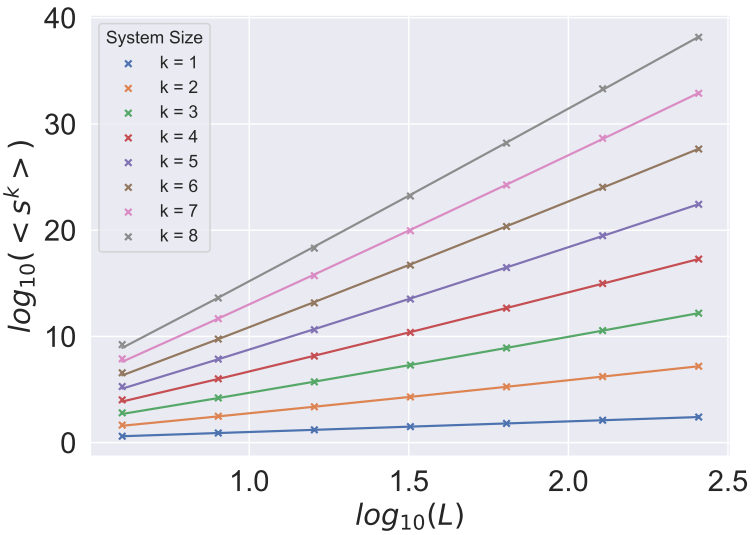

Figure 11: Moment analysis of avalanche size probability. The theory is only valid for $L \gg 1$ therefore in this figure, the fit for each moment k is done using data from systems where $L \geq 64$, where the corrections to scaling are very small. All the systems $\langle s^k \rangle$ are displayed on the figure as scatter points, however.

Figure 12 gives the result that $D = 2.21 \pm 0.07$ and $\tau_s = 1.50 \pm 0.12$. The errors on these values were obtained from the errors on the line of best fit gradient and intercept, then the errors were propagated meaning they do not include the errors for corrections

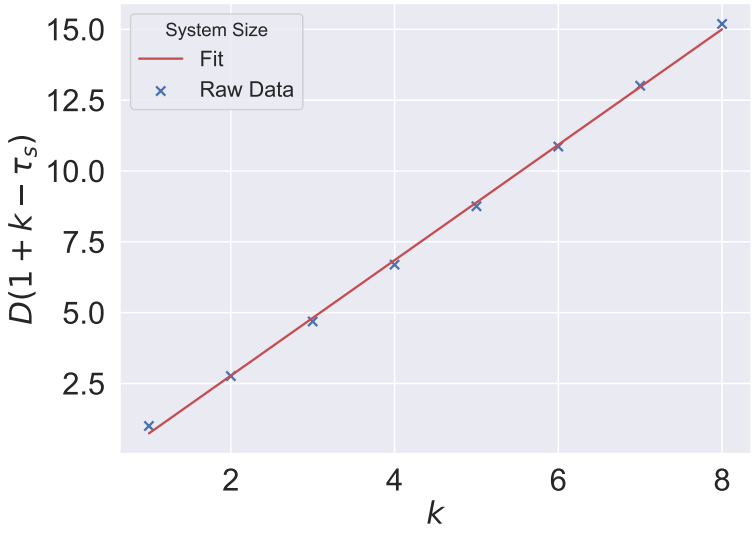

Figure 12: Plot of gradient $D(1 + k - \tau_s)$ vs k. There is a linear relationship between these two variables, as expected by the theory. The gradient of the line is 2.21 ± 0.07 and the y-intercept is -1.10 ± 0.08. This means that $D = 2.21 \pm 0.07$ and therefore $\tau_s = 1.50 \pm 0.12$.

to scaling. These values give $D(2 - \tau_s) = 1.1 \pm 0.1$. This is just outside the theoretical prediction that $D(2 - \tau_s) = 1$ but is close. Even though the corrections to scaling are small in systems $L \geq 64$, this may account for the additional error. These values are also consistent with the values used to create the data collapse in figure 10.

The corrections to scaling can be seen if the measured moment data is compared to the theoretical prediction for each system size L, using equation 18e and the values for τ_s and D that have been obtained. Note that the different moments likely have a different proportionality constant, as equation 18e only shows the proportion. If there are no corrections to scaling, the line for each moment should be flat with a gradient of zero across all system sizes.

The results for the comparison are shown in figure 13. It can be seen that the lines for each moment k flatten off as $L \to \infty$. This happens the fastest for larger moments of k. Consequently, the previous assumption of only using data from $L \geq 64$ was valid. With further time, the data would be used from only $L \geq 128$ however, as $L = 128$ system has significantly less corrections to scaling at the $k = 1$ moment than the $L = 64$ system.

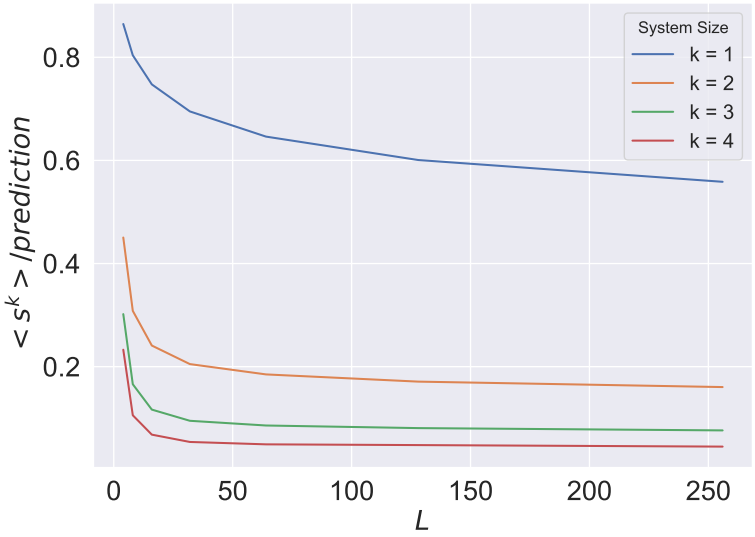

Figure 13: Plot of measured $\langle s^k \rangle$ divided by the predicted $\langle s^k \rangle$ vs L. As the lines aren't flat, for small L there are clearly corrections to scaling.

5 Conclusion

This report investigated the properties of the Oslo model of a rice pile in one dimension. The Oslo model algorithm was written in python where the system size could be modified. The height and avalanche data were collected every time a grain was added into the system. There was found to be a scaling function for the heights $h(t;L) = L\mathcal{F}(t/L^2)$. The distribution of heights was found to look like a Gaussian distribution however, is not actually a Gaussian as the standard deviation was found to scale differently than expected. This means that the assumption that z_i are independent and identically distributed is false.

The average height has corrections to scaling and can be expressed as $\langle h \rangle = a_0 L(1 - a_1 L^{-\omega_1})$, where $a_0 = 1.75 \pm 0.08$ and $\omega_1 = 0.40 \pm 0.03$. The avalanche probability $\tilde{P}_N(s;L)$ was found be $\propto s^{-\tau_s}\mathcal{G}(s/L^D)$ with $D = 2.21 \pm 0.07$ and $\tau_s = 1.50 \pm 0.12$. The Oslo model as a whole was found to exhibit criticality for $L \gg 1$.

References

[1] K.Christensen and N.Maloney, *Complexity and Criticality*, Imperial College Press, London, 2005.