Complexity Project Report

Suhail Shoaib Mall, 01392009

February 17, 2020

Word Count: 2549

Abstract

The Oslo Model was implemented using Python to analyse the numerous scaling relations displayed. The height of hte pile over time was found to follow the scaling form $h(t;L) = L\mathcal{F}\left(\frac{t}{L^2}\right)$ with $\mathcal{F}(x) = \text{const} \times x^{0.5}$ for $t < t_c$ and a_0 for $t > t_c$. The average height $\langle h \rangle = a_0 L\left(1 - a_1 L^{-w_1}\right)$ was investigated and the values of $a_0 = 1.733 \pm 0.001$, $a_1 = 0.168 \pm 0.01$, and $\omega_1 = 0.551 \pm 0.08$ were found. The distribution of steady-state height was then considered and assumed to be Gaussian. Using the measured values of $\langle h \rangle(L)$ and $\sigma_h(L)$, the data was collapsed onto a Gaussian function of mean 0 and width 1 confirming the Gaussian form of $P(h;L)$.

The second set of scaling relations regarded the avalanche size distribution and the scaling ansatz for the avalanche probability $P(s;L) = s^{-\tau_s}\mathcal{G}\left(\frac{s}{L^D}\right)$ was proposed. A curve was fit to the data to predict $\tau_s = 1.561$ and $D = 2.25$. Estimates for the parameters were then made by analysing the moments o the distributions to yield $\tau_s = 1.563 \pm 0.004$ and $D = 2.175 \pm 0.006$.

1 The Oslo Model

1.1 Implementation

The Oslo Model is defined as a one dimensional system of size L with each site i described by a height h_i. Each site is also associated with a slope $z_i = h_i - h_{i+1}$. The threshold slope of each site, z_i^{th} is initialised as 1 with user-inputted probability p, or 2 with probability $1 - p$.

The model is driven by adding a grain to the leftmost site, i.e. the height of the $i = 1$ site is increased by 1. After each drive, all sites must be checked to ensure that each site is stable: $z_i < z_i^{th}$ for all sites i. If a site i is unstable, the site is "relaxed" by decreasing the height at that site by 1 and increasing the $(i+1)^{th}$ height by 1 to simulate the grain toppling from an unstable site to the next site:

$$h_i \to h_i - 1$$
$$h_{i+1} \to h_{i+1} + 1$$

The slope at the final site $i = L$ is defined to be the height $h_{i=L}$ and relaxing this site causes a grain to leave the system by decreasing only $h_{i=L}$ by 1.

Furthermore, when site i is relaxed, the threshold slope of that site is randomly chosen again to be 1 with a probability p or 2 with $(1 - p)$. This represents the grain interacting with the system e.g. by changing the angle of the topmost grain through the physical process of toppling.

Checking that the system is stable can be done by iterating through the entire system and checking the stability of each site individually, and repeating this until an iteration yields no relaxations. This is a brute-force method and can be optimised.

This can be done by considering the ways by which a site can become unstable: firstly a relaxation only affects the slopes of the relaxed site (decreasing by 2) and the neighbouring sites (increasing by 1), hence the neighbouring sites must be checked after a relaxation.

The second way is best understood by "following the grain" as shown in Fig. 1. Consider a site with both a slope and threshold of 2. If the neighbouring site before it topples – and following that grain – it will pass over the considered site and possibly change its threshold from 2 to 1 and decrease its slope to 1. If the followed grain topples further, then the slope will return 2 and the considered site is now unstable. In this way, relaxing a site and following the grain provides a range over which the grain has interacted with sites and may have caused them to become unstable.

Following the grain was implemented by iterating forward and relaxing unstable sites until a stable site is reached. If the position of this site is higher than the stored maximum site then this becomes the new site that defines the range to be checked over.

Rather than brute-force iterating through the system multiple times, instead when an unstable site is encountered, the site is relaxed and the grain is followed forward. The previous site

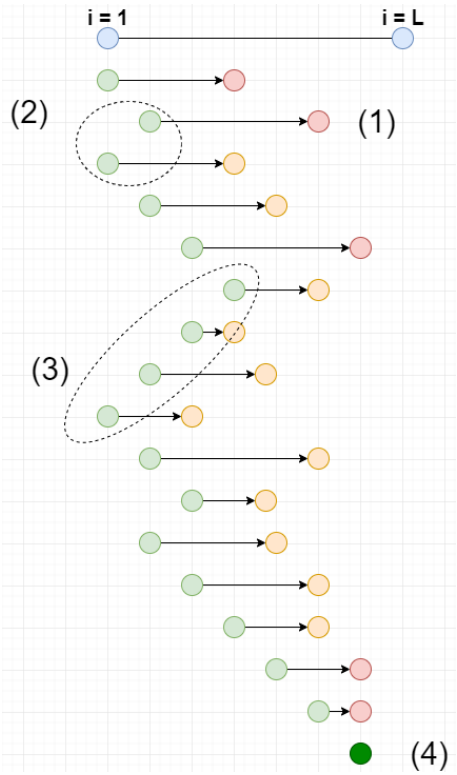

Figure 1: An example of the checks and relaxations in the Oslo Model. The green sites are relaxed and the grains followed to the orange sites. If this site defines the new range then it is red. The blue sites define the system length and green is when checks over all relevant sites have been made.
1) this is a case where the maximum site must be updated.
2) After the site has been relaxed, an iteration backwards is started and the site to its left is checked. Because this site has $i = 1$ it cannot be iterated backwards further.
3) An example of an iteration backwards where relaxing a site leads the site before it to need to become unstable. After a stable site is reached, the site immediately to its right is considered.
4) The considered site is now the stored maximum site so all sites that have been affected by the initial drive have been checked.

is then checked and an iteration backward is started where unstable grains are relaxed and followed. This continues until a stable site is reached, where the next position to the right is checked and an iteration backwards is started again if it is unstable. This is continued until the checked site is equal to the stored maximum site.

This is significantly faster at the start of the simulation when only a small portion of the system is in use, but slows down considerably later on as the extra checks against the maximum range become cumbersome.

1.1.1 Testing

The above was tested first by ensuring that all sites are indeed stable before the next grain is added. Secondly, the values of $p = 0$ and $p = 1$ were tested as these would result in a staircase configuration with slope 1 and 2 respectively. Hence the steady-state height must be L and $2L$ and this was checked for different L and found to be true. The steady-state height for $L = 16$ and 32 were given [1] and checked against.

2 Height of the Pile Over Time

Q 2.a

The height of the pile $h(t; L)$ is defined as the height of the first site and $\tilde{h}_N(t; L)$ the height of the pile averaged over N independent realisations. All investigations into the height of the pile were carried out on system sizes $L = 2^n$ where n is an integer between $n = 2$ ($L = 4$) and $n = 9$ ($L = 512$).

The system time t is defined by the number of grains added to the system and so can also be used to index the stable configuration reached after each drive of the system.

The total average height $\tilde{h}_N(t; L)$ of the pile was measured at each time for each system size and plotted, though due to the difference in scale, system sizes $L = 4$ to 64 and 128 to 512 were plotted on separate subfigures of Fig. 2.

The height array was averaged over 100 realisations for 10,000 grain additions for the former, and 10 realisations up to 300,000 ($L = 512$) for the latter.

Qualitatively, the above figures show that for each L, $\tilde{h}_N(t; L)$ scales with time up to a crossover time $t_c(L)$ after which it becomes roughly constant. The crossover time is found to be the time immediately before the first grain leaves the system, and hence the configurations before it are transient as the number of grains in the system always increases with each drive. However $h(t; L)$ at two different times can still be equal as multiple configurations can have the same height. The configurations after $t_c(L)$ are recurrent, as the pile spans the system and grains can now leave it.

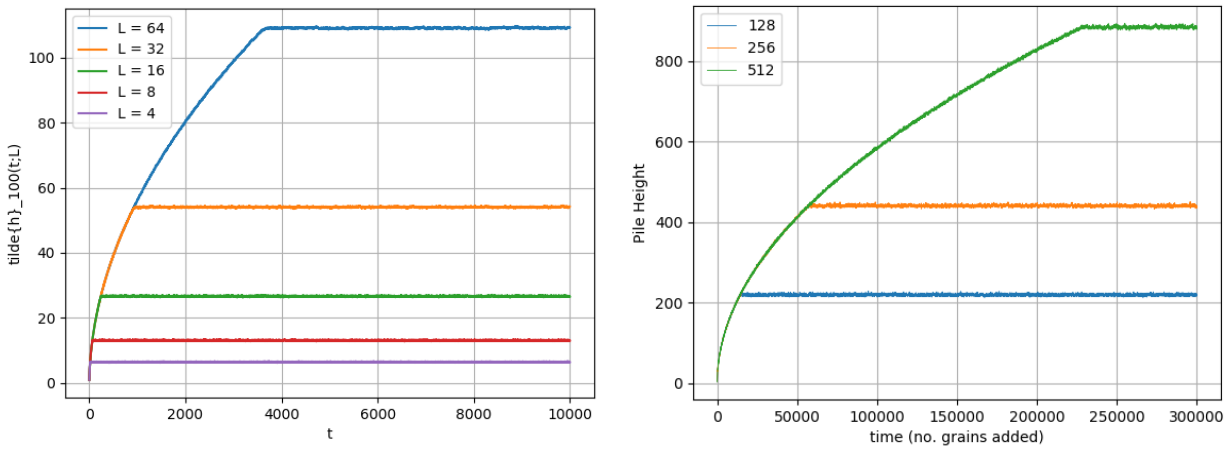

Figure 2: The processed height of the pile over time for different system sizes. Left shows $L = 4$ to $L = 64$ averaged over 100 realisations and right shows $L = 128$ to $L = 512$ averaged over 10 realisations. The pile height scales with time up to some crossover and is then roughly constant.

2.b

The crossover time was measured numerically simply by noting the system time when the $i = L$ site has first been relaxed. This value was averaged over 10 realisations for each system size L and plotted against L. This is shown in Fig. 3 and fitted to a quadratic of the form $f(x) = ax^2$ using the *scipy.optimise.curve_fit* function, yielding a value for a of $0.858 \pm 0.018 \times 10^{-7}$ (though this uncertainty is associated only with the goodness of fit).

2.c

Theoretically, the crossover time is indeed expected to scale to first order as L^2. Considering a system of size $L >> 1$ with an average slope of $\langle z \rangle$ at $t = t_c$, the system can be approximated to a right-angled triangle with horizontal length L and height $\langle z \rangle_{t=t_c} L$. Because no grains have yet left the system, the crossover time (averaged over M realisations) is equal to the number of grains in the system, given by the area of the triangle:

$$\langle t_c(L) \rangle_M = \frac{1}{2} \langle z \rangle_{t=t_c} L^2 \tag{1}$$

Furthermore, the system will continue to look like a staircase after the crossover time so the same geometrical arguments can be used to show that the pile height $h(t > t_c; L)$ continues to scale linearly with L.

The measured crossover time was compared to the theoretical crossover time for a single realisation using the found value of $a = 0.858$. This is shown in Fig. 4 and shows deviations to prediction for small system sizes, suggesting that corrections to scaling may be needed.

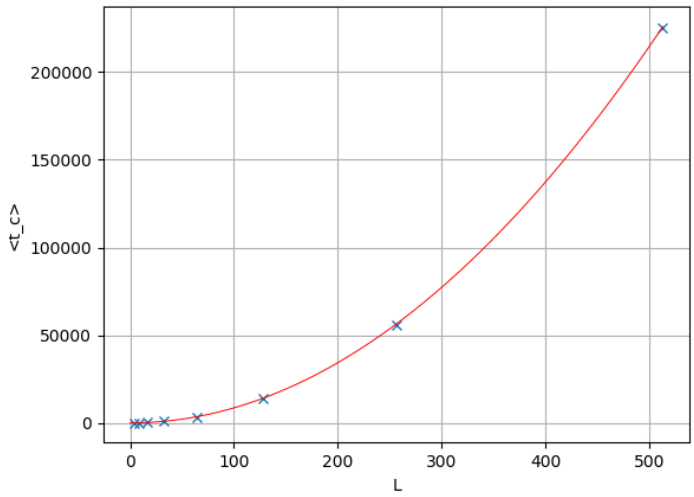

Figure 3: The crossover time was measured numerically as the first time a grain leaves the system. This was averaged over 10 realisations for different system sizes and plotted (x). A fit of ax^2 was found with $a = 0.858$ (-).

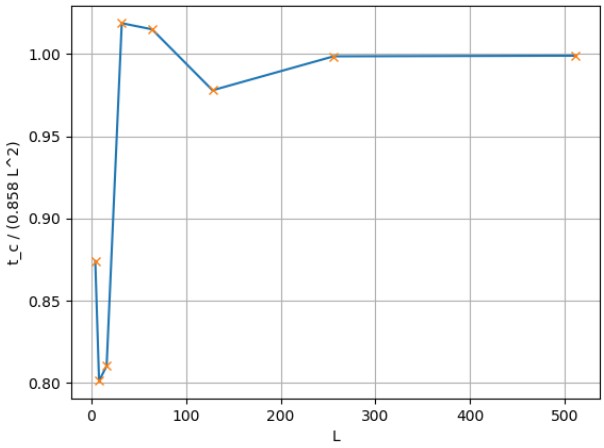

Figure 4: The ratio of the measured crossover time to the first order prediction of $t_c \propto L^2$. The large dips for small system sizes suggest that corrections to scaling may be required.

This can be found by a more careful analysis. It can be shown that the number of grains in the system is

$$N = \sum_{i=1}^{L} iz_i$$

so the crossover time can be found as

$$t_c(L) = \sum_{i=1}^{L} iz_i = L \langle iz_i(L) \rangle \tag{2}$$

As z_i is chosen randomly it is independent of i:

$$t_c = L\,\langle i z_i(L)\rangle = L\langle i\rangle\langle z(L)\rangle = L\left(\frac{1}{L}\sum_{1}^{L} i\right)\langle z\rangle = \frac{\langle z\rangle}{2}L(L+1) = \frac{\langle z\rangle}{2}L^2(1+\frac{1}{L}) \qquad (3)$$

which shows that some correction to scaling is required.

2.d

Using these scaling relations for t_c and $h(t > t_c; L)$ with L, the scaling ansatz Eqn. can be suggested:

$$\tilde{h}_N(t; L) = L\,\overline{\mathcal{F}}\left(\frac{t}{t_c}\right) = L\,\mathcal{F}\left(\frac{t}{L^2}\right) \qquad (4)$$

where the constant constant of proportionality between t_c and L^2 has been absorbed into $\mathcal{F}$.

The scaling function $\mathcal{F}$ must behave differently before and after the crossover time. After the crossover time, the pile height is roughly constant so $\mathcal{F}(x)$ is constant for $x >> 1$. Before the crossover time it can be noted that the pile height is independent of L. In order for the functional form of $\tilde{h}_N(t < t_c; L)$ to be independent of L, $\mathcal{F}(\frac{t}{L^2})$ must scale as $\frac{1}{L}$ so the function must scale as $x^{0.5}$ for $x << 1$:

$$\mathcal{F}(x) = \begin{cases} \text{const} \times x^{0.5} & ; \quad x << 1 \\ \langle z\rangle \quad ; & x >> 1 \end{cases} \qquad (5)$$

Eqn. 4 can be used to produce a data collapse by multiplying the measured heights $\tilde{h}_N(t; L)$ by L^{-1} and plotting against $\frac{t}{L^2}$ for several L. This is shown in Fig. 5.

3 Steady-State Height Distribution

2.e

Although the data collapse is reasonably well-fitted for $t > t_c$, the deviations for $t < t_c$ suggest that some corrections to scaling may be required. The assumption that the steady state height is proportional to L was modified to consider the next highest power of L:

$$\langle h(t > t_c; L)\rangle_t = a_0 L \;\rightarrow a_0\,L\,(1 - a_1\,L^{-w_1}) \qquad (6)$$

where $\langle h(t > t_c; L)\rangle_t$ is the height of pile in steady state averaged over 100,000 time steps for one realisation.

This was first plotted against L to confirm the scaling relation to first order, and then multiplied L^{-1} to make the correction more apparent as shown in Fig. 6. The *curve_fit*

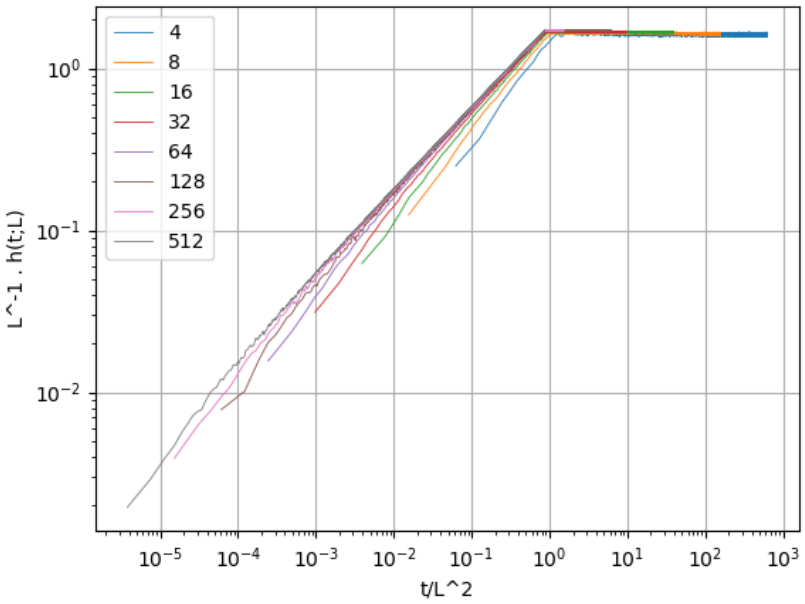

Figure 5: The collapsed height distribution that reveals the form of $\mathcal{F}(x)$ as a power law with exponent 0.5 for $x < 1$ and a constant for $x > 1$.

algorithm was used again to determine values for a_0, a_1, and ω_1 and a line of best fit is plotted to the data.

It was found that $a_0 = 1.733 \pm 0.001$, $a_1 = 0.168 \pm 0.01$, and $\omega_1 = 0.551 \pm 0.08$, though the off-diagonal elements of the covariance matrix were also of similar order.

The value of a_0 is consistent with predictions, as it is equal to $\langle z \rangle$. Although this would suggest that $a_0 \simeq 1.5$ as the threshold probability is 0.5, this does not take into account that sites with threshold 2 are less likely to relax. Hence there are more sites with slope 2 than 1 and so a_0 should be between 1.5 and 2.

2f

The standard deviation of the steady-state height of the pile $\sigma_h(L)$ was calculated numerically as $\sqrt{\langle h(t;L)^2 \rangle_t - \langle h(t;L) \rangle_t^2}$ and can be assumed to have the scaling form

$$\sigma_h(L) = aL^b \tag{7}$$

The values of a and b were found to be 0.528 ± 0.002 and 0.255 ± 0.003 respectively, which suggests that $\sigma_h(L)$ may theoretically scale with $L^{0.25}$. If this is the case, then $\sigma_{\langle z \rangle} = \frac{\sigma_h}{L}$ which will scale as $L^{-0.75}$. So for $L \to \infty$, $\sigma_{\langle z \rangle} \to 0$ and $\langle z \rangle$ will tend to a constant value for all large L.

2g

The probability distribution of heights $P(h;L)$ for $t > t_c$ was defined as:

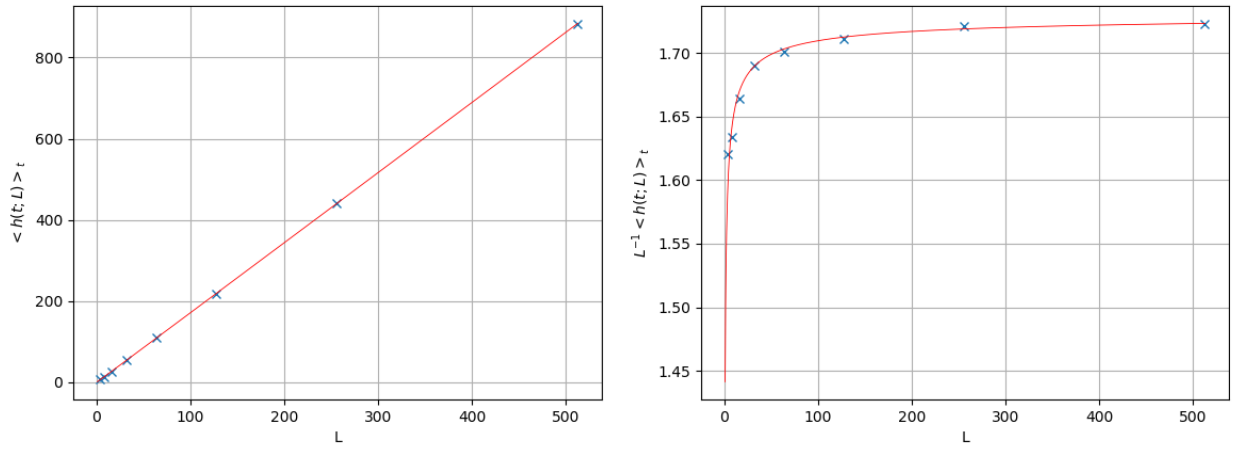

Figure 6: The steady state height as a function of system size (left) plotted with the predicted linear fit. The corrections to scaling are shown in the right subfigure by assuming the correction is a factor of the form $(1 - a_1 L^{-\omega_1})$. a_1 was found to be 0.168 and ω_1 to be 0.551.

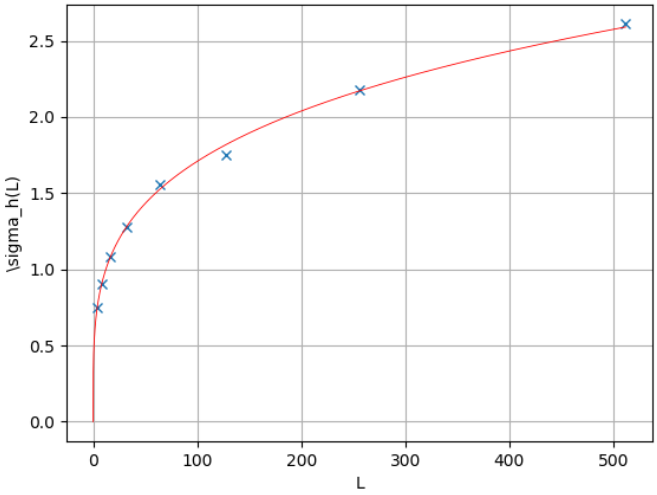

Figure 7: The measured standard deviation of the steady-state height calculated over 100,000 time steps, as a function of system size. It shows a clear power-law behaviour with an exponent found to be 0.255 ± 0.003.

$$P(h; L) = \frac{\text{No. observed configurations with height } h \text{ in system size L}}{\text{No. configurations sampled}} \tag{8}$$

From [1] it is seen that as the system size increases, the total number of unique recurrent configurations $\mathcal{N}_R$ increases drastically. Furthermore, because the configurations are recurrent there is no guarantee that considering $\mathcal{N}_\mathcal{R}$ samples of the height will sample each configuration. However an infeasibly large number of samples would have to be considered for this so

it should be noted that noise in the data could be due to an under-sampling of configurations.

The same array of data from questions 2.e and 2.f was used so that the total number of samples for each system size was 100,000. $P(h; L)$ was found by initialising an array with a length equal to the maximum sampled height. The height array is then iterated through and for a height $h = j$, the j^{th} position of the distribution array is increased by 1. Finally the distribution is divided by the number of samples and is checked that it is indeed normalised.

The height of the pile can be expressed as the sum of all slopes, and each slope is an independent random variable. Because any function of a random variable is itself a random variable and by the Central Limit Theorem, $P(h; L)$ for a sufficiently large sample size is expected to be Gaussian. The Gaussian should have a mean of $a_0 L$ and width given by the results of question 2.f as the same data was used to find $\sigma_h(L)$ for both questions.

$P(h; L)$ is presented as a function of h for several L in Fig 8 with a Guassian fit to each system size.

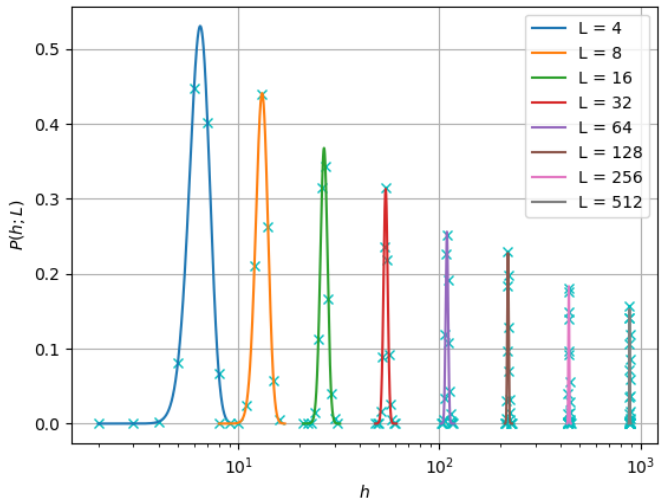

Figure 8: The height distribution of various system sizes in steady-state. Because the heights are assumed to be randomly-distributed, their distribution for a large number of samples is predicted to be Gaussian for each system. This is fit to the data and the fit parameters generally agree with results from previous sections.

The measured values of $\langle h \rangle(L)$ and $\sigma_h(L)$ can be used to produce a data collapse by rescaling the Gaussian function fit to each curve. The expected Gaussian fit for the height distribution is defined as

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

By multiplying each side by the measured $\sigma_h(L)$ for each system size and defining

$$h' = \frac{(h - \langle h \rangle)^2}{\sigma_h^2}$$

the data is expected to collapse onto a Gaussian function of mean 0 and width 1. This is plotted in Fig. 9. also plotted is the expected Gaussian fit and the actual fit, which has $\langle h' \rangle = 6.035 \times 10^{-7}$ and $\sigma = 1.000$ which agree with the prediction. However it can be seen qualitatively that there may be a positive skewness which suggests that the assumption of independence of slope is not correct (not correction to scaling as it is present for all L).

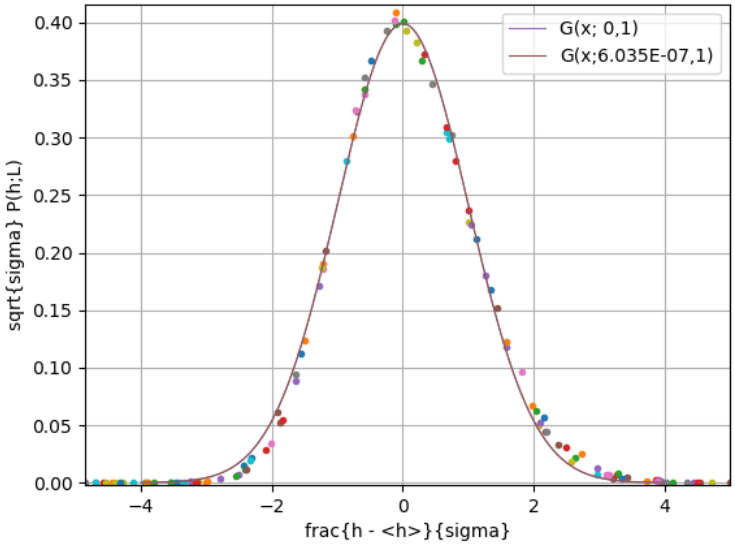

Figure 9: Data collapse for the height distribution of systems of different L. Although the fit parameter to a combined array of data for all L provides a fit that agrees with the prediction, a slight skewness can be seen, suggesting that the assumption of the slopes being independent may not be correct.

4 Avalanche-size Probability

An avalanche was defined in this model to be the number of relaxations after one drive (avalanches of size zero are considered for moment analysis but not when plotting the avalanche-size distribution). The avalanche-size distribution is defined similarly to the pile height distribution:

$$P_N(s; L) = \frac{\text{Number of avalanches of size } s}{\text{Total number of sampled avalanches, } N} \tag{10}$$

again measured after the system has passed the crossover time.

3.a

Although the avalanche-size distribution is defined in the same manner as the height distribution, presenting the distribution with linearly-spaced bins would cause loss of information for the larger-sized avalanches as their frequency is so comparably low compared to the smaller-sized avalanches.

This was solved using a logarithmically-increasing bin sizes [1] (written code was supplied) where each bin is a user-inputted factor a larger than the previous (for $a > 1$). Rather than considering the number of avalanches of size s, the distribution is given as the number of avalanches that fall within a bin $a^j \rightarrow a^{j+1}$. The normalisation factor must now also be multiplied by the number of possible avalanche sizes in bin j which is $(1 + a^{j+1} - a^j)$. This gives the logbinned avalanche-size distribution as:

$$\tilde{P}_N(s^j; L) = \frac{\text{No. avalanches in bin } s = a^j \rightarrow a^{j+1}}{N\,(1 + a^{j+1} - a^j)} \tag{11}$$

The bin j is then associated with the geometric mean $\sqrt{a^j \,.\, a^{j+1}}$.

$\tilde{P}_N(s^j; L)$ is plotted as a function of s^j for different L using $a = 1.5$ (a value taken from trialing) in Fig. 10

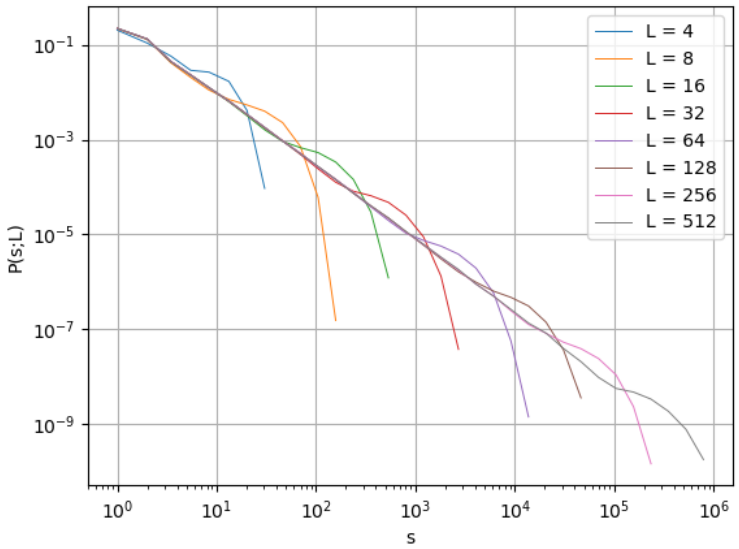

Figure 10: The logartihmically-binned data for the avalanche-size distribution for different system sizes. All systemsfollow the same power law up to some cutoff size, after which there is a bump and rapid decay.

Qualitatively, $\tilde{P}_N(s; L)$ follows a power law independent of system size up to a cutoff size which is dependent on system size. At this cutoff size, there is a bump after which the probability decreases rapidly.

Because the bump and decay are dependent on L, they must be related to some finite system effect. This can be explained by noting that avalanches cannot have a size greater than L, hence the rapid decay. The reason for the bump is because all avalanches that would be counted in an infinite system are instead present here as system-spanning avalanches before the decay.

3.b

Fig. 10 has two key features: the scaling region up to some cutoff size, and the subsequent bump and rapid decay. Because the bumps seem equally-spaced in L, it can be assumed that the cutoff size is a power law in L, leading to the finite scaling ansatz:

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

This can be manipulated to form a data collapse multiplying both sides by s^{τ_s} and rescaling the avalanche size $s \to \frac{s}{s_c}$.

τ_s was found using curve_fit to fit a power law function to the scaling region of the $L = 512$ data over an appropriate range. This yielded a value of $\tau_s = 1.561$ as shown in Fig. 11. Plotting this against s produced Fig. 12 , after which the value of D was varied manually until a data collapse was achieved. This occured for $D = 2.25$ and is presented in Fig. 13.

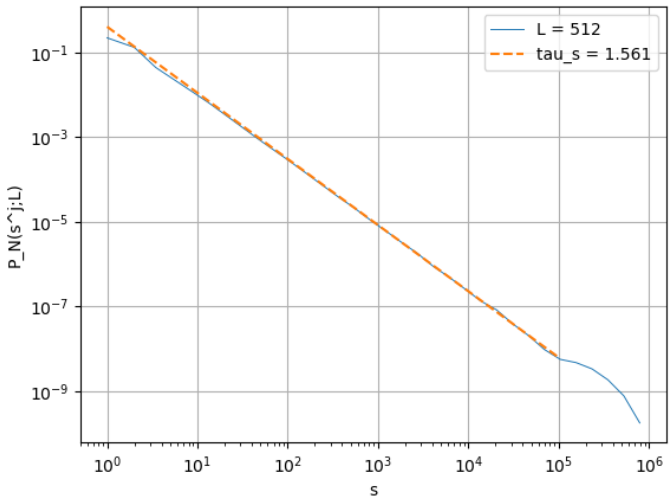

Figure 11: A fit of the form $y = as^{-\tau_s}$ to the avalanche-size distribution for $L = 512$. The function was fit to the data to find an optimal value of τ_s (shown in the plot).

3.c

The kth moment of $\langle s^k \rangle$ was measured numerically as:

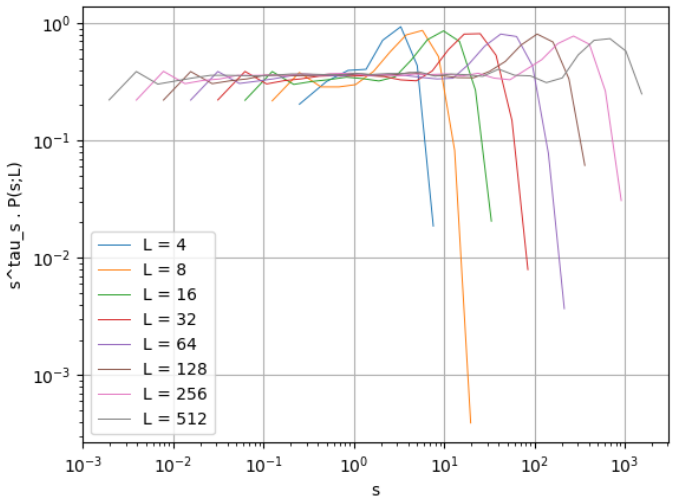

Figure 12: The partially-collapsed avalanche-size distribution for different system sizes found by multiplying the distribution by $\tau_s = 1.561$. This makes the scaling region horizontal which allows the lining up of the "bumps" easier to carry out by eye.

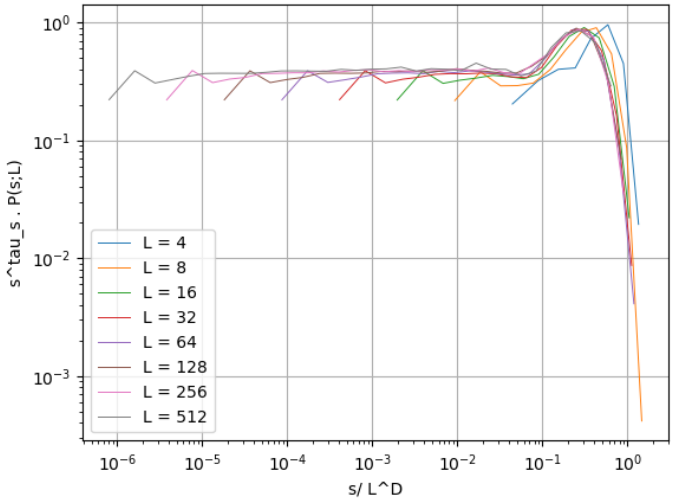

Figure 13: Data collapse formed by multiplying the avalanche-size distribution by $\tau_s = 1.561$ and rescaling s to $\frac{s}{L^{2.25}}$. It can be seen that the $L = 4$ data do not collapse well, suggesting that corrections to scaling may be required.

$$\left\langle s^k \right\rangle = \lim_{T\to\infty} \frac{1}{T} \sum_{t=t_0+1}^{t_0+T} s_t^k \tag{13}$$

where s_t is the avalanche size at a given time. Mathematically,

$$\langle s^k \rangle = \sum_{s=1}^{\infty} s^k P(s;L) = \sum_{s=1}^{\infty} s^{k-\tau_s} \mathcal{G}\left(s/L^D\right) \tag{14}$$

where Eqn. 12 has been used in the last step.

Approximating the sum to an integral and making an appropriate substitution it can be shown [1] that

$$\left\langle s^k \right\rangle = L^{D(1+k-\tau_s)} \int_{1/L^D}^{\infty} u^{k-\tau_s} \mathcal{G}(u) du \quad \propto L^{D(1+k-\tau_s)} \tag{15}$$

where the rapid decay of $\mathcal{G}$ guarantees the convergence of the integral if $1 + k > \tau_s$.

$\langle s^k \rangle(L)$ was found for $k = 1, ..., k = 12$ for different L and each moment was plotted against L separately in Fig. 14 (left). The gradient of the line in log-log space is equal to $D(1 + k - \tau_s)$ and this was measured and plotted against k in Fig. 14 (right).

Plotting grad $= D\,k + D(1 - \tau_s)$ provided the value for D of 2.175 ± 0.006 from the gradient of the plot and $\tau_s = 1.563 \pm 0.004$ from the intercept (though the uncertainty is from the final fit parameters only). The data collapse using these values is shown in Fig. 15 .

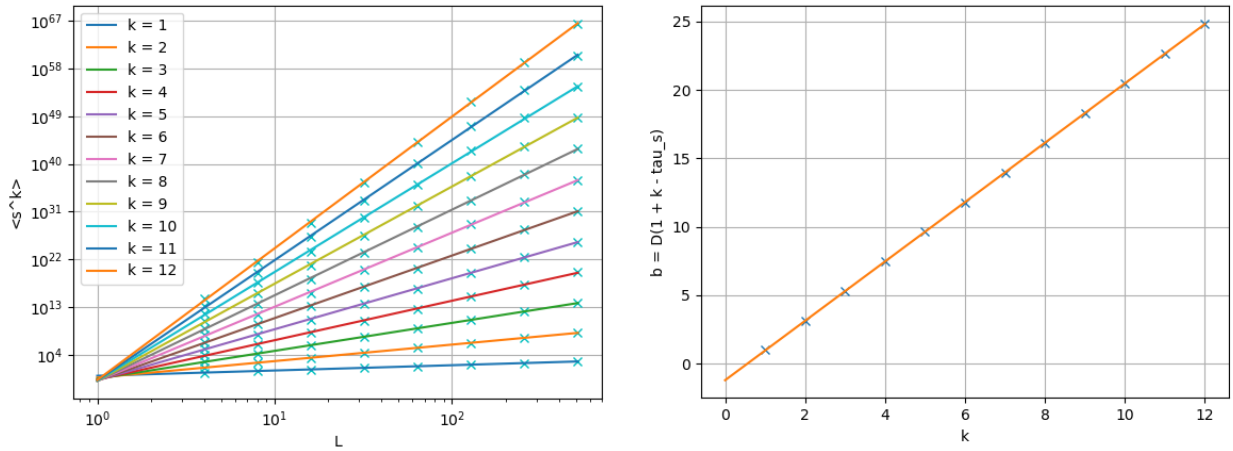

Figure 14: The k^{th} moment of the avalanche-size distribution as a function of L (left). The gradient in log-log space is equal to $D(1+k-\tau_s)$ and so can be used to find the fit parameters τ_s and D. This is done in the right subfigure tof ind $\tau_s = 1.563$ and $D = 2.175$.

Because the $L = 4$ collapsed data does not lie on the function $\mathcal{G}$ as the other systems do, corrections to scaling may be expected. These can take the form of a better fit to $\langle s^k \rangle$, considering higher powers to fit the data to rather than just a linear gradient. However this was not achieved.

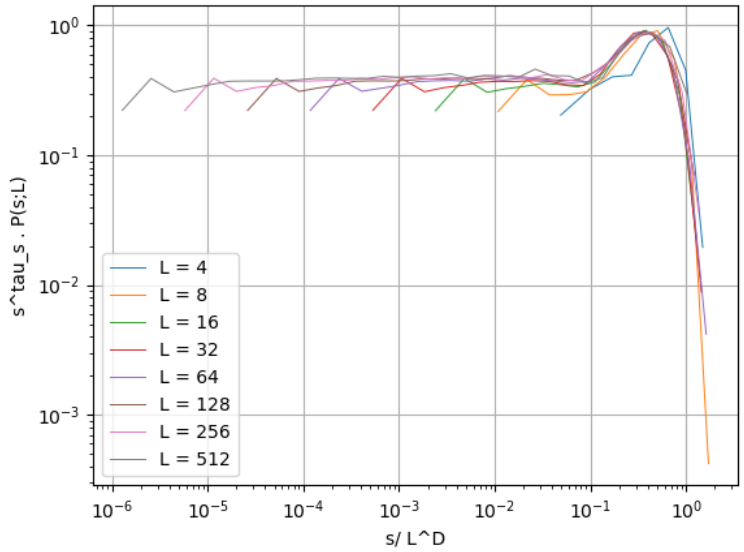

Figure 15: The avalanche-size data collapse using $\tau_s = 1.563$ and $D = 2.175$ as found from analysing the first 12 moments of the distribution.

5 Conclusion

The Oslo Model was shown to follow several scaling relations for $L = 4$ to $L = 512$, an indication that it displays criticality. The height of the pile was found to scale with $\tilde{h}(t; L) = L\mathcal{F}(t/L^2)$ where $mathcalF(x)$ is constant for $x > 1$ and scales approximately as $x^{0.5}$ for $x << 1$. Corrections to scaling were made for both the crossover time and justified and predicted for the average steady-state height.

The steady-state height distribution was found to fit a Gaussian function around $a_0 L$, where a_0 is consistent with the previous result. The measured standard deviation of the Gaussians was found to scale as $\sigma_h(L) = 0.53 \times L^{0.255}$.

The avalanche-size distribution was found to scale as $P(s; L) = s^{-\tau_s}\, \mathcal{G}\left(\frac{s}{L^D}\right)$. The first estimates yielded values of $\tau_s = 1.561$ and $D = 2.25$. The second estimate used moment analysis to yield values of $\tau_s = 1.563$ and $D = 2.175$.

6 References

1. Christensen, K. and Moloney, N. (2005). Complexity and criticality. London: Imperial College Press.