

# Complexity Project

## The Oslo Model: A Self-Organised Criticality Model

CID: 01714841

18th February 2022

**Abstract:** The Oslo Model is a model used to simulate the behaviour of grains of sand that are placed at a single location in a system of size  $L$  such that through performing relaxations to the model, a pile of sand is formed. The Oslo Model Algorithm models the sand-pile system in terms of columns with  $h_i$  grains of sand and defines the slope of the column the difference in number of grains of sand between neighbouring slopes. The algorithm adds grains to the first column and relaxes a column when the slope at any given column reaches a randomly distributed threshold value. Using this model allows us to investigate scaling of the model in terms of the height and avalanche size probability distributions. By modelling the system computationally can we study the nature of these distributions that give insight into the functioning of the model.

**Word count:** 2496 words in report (excluding front page, figure captions, table captions, acknowledgement, equations and bibliography).

# 1 Implementation of The Oslo Model

## 1.1 Testing the Programme [Task 1]

There are numerous ways of testing the implementation of the Oslo model to check if the programme is working as intended. As stated in the project notes, one test is measuring the average height of the  $L = 16$  and  $L = 32$  states which have average heights equal to 26.5 and 53.9 once the system has reached the steady state respectively. To check this, Fig. 1 was plotted and shows the height of the pile  $h$  with time for the two system sizes. From first inspection, one can see that the graph follows a path that one would expect from the Oslo model. The height of the pile increases as more grains are added. The rate at which the height increases decreases as the system requires more grains to increase  $h$  and eventually levels off around a mean height  $\langle h \rangle$  as the system reaches the steady state and grains begin to fall out of the system. The mean height during the steady state can be calculated by taking 1000 samples of  $h$  during the steady state and indeed  $\langle h \rangle$  was calculated to 26.53 and 53.97 - inline with the theoretical values.

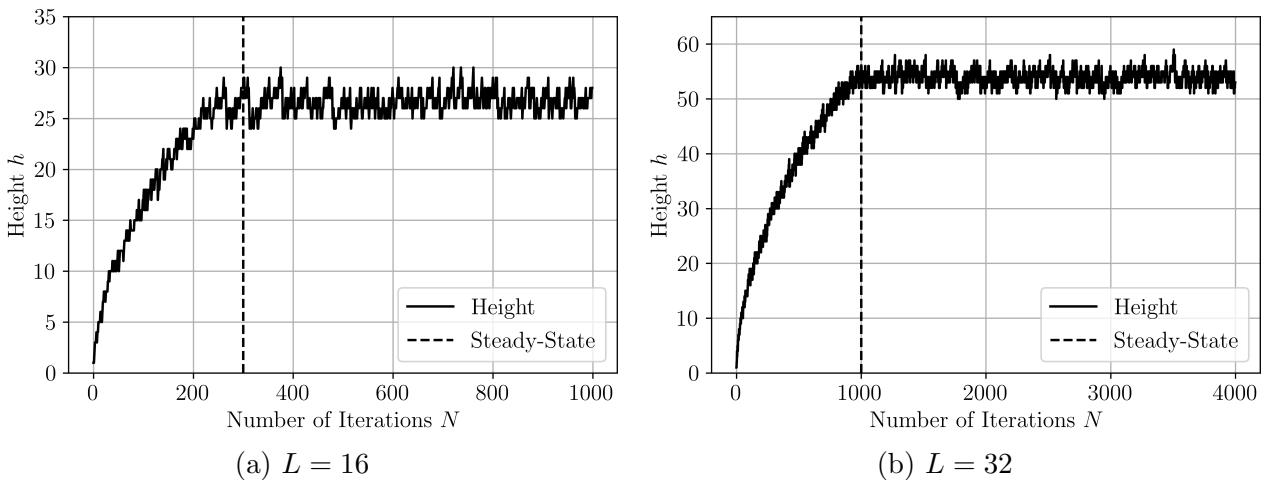


Figure 1: Height of the pile against number of iterations showing how the height of the pile increases in the transient phase and plateaus during the steady state. Vertical Lines indicate the time when the system reaches the steady state.

A second way of testing the model is changing the variable  $p$  that is responsible for changing the threshold values of each column. Setting  $p$  to extreme values of 1 and 0 respectively gives greater insight. Specifically  $p = 0$  sets all threshold values to 2 and  $p = 1$  sets all threshold values to 1 - which we famously know as the BTW model. We can theorise a number of things about these 2 models. Firstly, the steady state height for  $p = 0$  should be twice that for  $p = 1$  and twice the system size  $L$ ; secondly, the steady state height should also be constant as the thresholds are not changing (grains added fall out instantly) and finally, the time it takes to reach the steady state should for  $p = 1$  be double that of  $p = 0$ . Fig. 2 shows the height against time graph for a system of size 16. As expected, the height is 16 and 32 in the steady state for  $p = 1$  and  $p = 0$  respectively. The height is indeed constant in the steady state and finally the number of iterations it takes to reach the steady state for  $p = 0$  is 272 which is twice that for  $p = 1$  - which is 136.

Finally, a more intuitive way of making sure the program works is by adding a visualisation tool. This is a qualitative rather than quantitative way of testing the model but is an easy way to ensure things are running as predicted. Fig. 3 shows the states of the a system of size  $L = 4$  after 5 individual iterations for  $p = 0$  and  $p = 1$ . Firstly, we note that the drive of the model

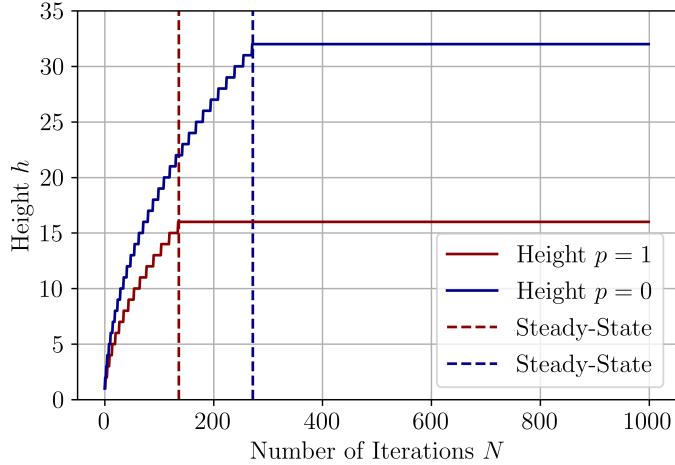


Figure 2: height versus number of iterations for the two extreme models -  $p = 0$  and  $p = 1$  for the  $L = 16$  and  $L = 32$  system sizes. Dashed vertical lines indicate when the system reaches the steady state.

is correct as the number of grains in the system is equal to the number of iterations - this is only valid before the steady state which is the case. Secondly, the maximum  $z$ -value (that is the difference between neighbouring columns) never exceeds the threshold - that is 1 for  $p = 1$  and 2 for  $p = 0$ . Finally, we can compare the systems for  $p = 0$  and  $p = 1$  such that where  $z = 2$  in the  $p = 0$  system, a relaxation has occurred in the  $p = 1$  system as a threshold is met in one but not the other. This can be seen clearly in the second iteration.

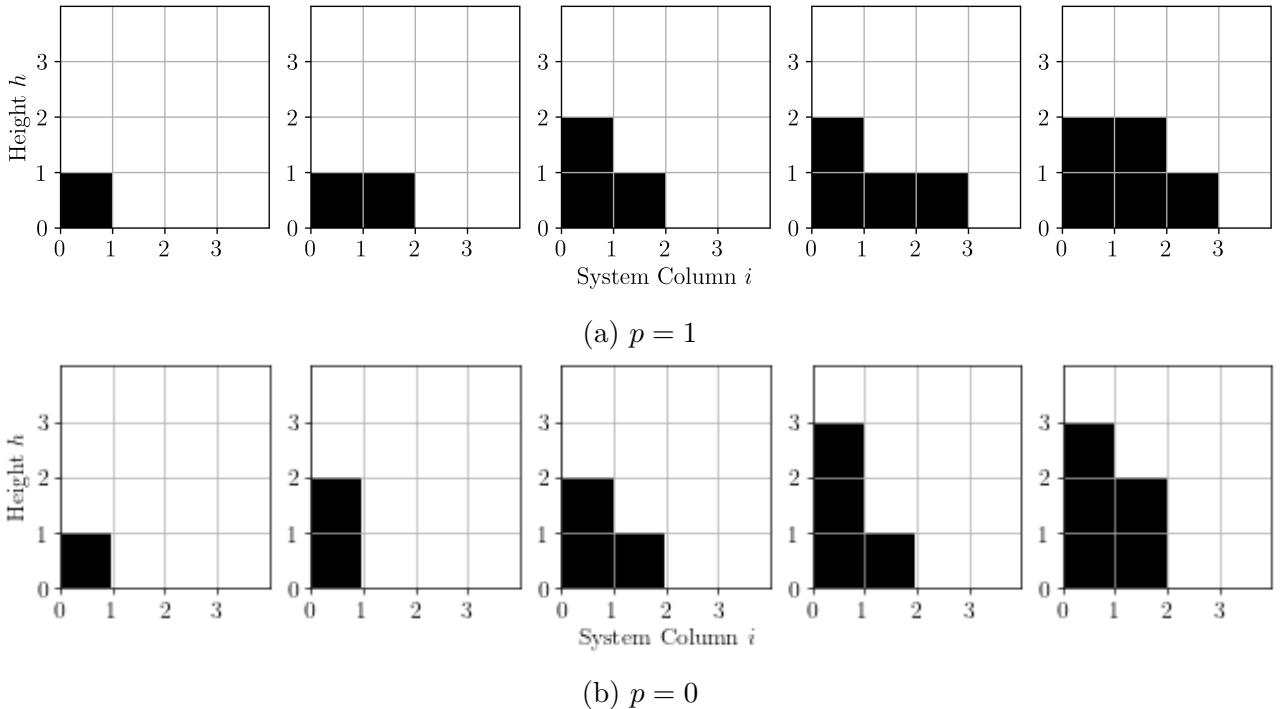


Figure 3: First 5 configurations of the model when  $p = 0$  and when  $p = 1$  at the end of every iteration for a system of size  $L = 4$ - i.e. after drive and relax has been performed.

## 2 The Height of the Pile $h(t; L)$

### 2.1 Height versus Time [Task 2a]

By measuring the height of pile as the height of the column  $i = 1$ , the height of the pile was plotted for systems of size  $L = 4, 8, 16, 32, 64$  and  $128$ . Fig. 4 shows the results. There are 2 features of the graphs that are of interest, that is the height during the transient phase and during the recurrent phase. Firstly, the transient configurations for all  $L$  fall on the same line on the logarithmic plot, this not only suggests the transient behaviour can be modelled as a power law, but also that models of different system sizes share transient states. To be specific, a system of size  $L_0$  shares transient configurations with systems of size  $L < L_0$  but not  $L > L_0$ . That is because at a certain point the system fills up all its columns and grains start to fall out of the system and it reaches its steady state. Therefore, it does not have access to transient configurations of larger system sizes. Once the model reaches the steady state, it experiences recurrent configurations and the height levels off. The recurrent configurations are mutually exclusive to that system size. For example, the set of recurrent configurations for a system size  $L = 8$  has no overlap with the set of configurations for  $L = 16$ .

### 2.2 Cross-Over Time with System Size [Task 2b]

We define the cross-over time  $t_c$  as the number of grains in a system or equivalent number of iterations until a grain leaves the system for the first time.  $t_c(L)$  can be measured for the various system sizes and the average cross over time  $\langle t_c \rangle$  can be measured by taking the average of  $M t_c(L)$  data points. Fig. 5 shows the resulting data for  $M = 10$ . By plotting on a logarithmic scale,  $\langle t_c \rangle$  falls on a straight line suggesting it can be modelled by a power law:  $\langle t_c \rangle = \alpha L^\beta$  where  $\alpha$  and  $\beta$  were measured to  $0.823$  and  $2.007$  - implying a quadratic relationship. The error bars on the plot are too small to plot due to  $M$  being large enough. For  $L \gg 1$ ,  $t_c$  will be very large and will increase with the square of  $L$  (as  $\beta \sim 2$ ).

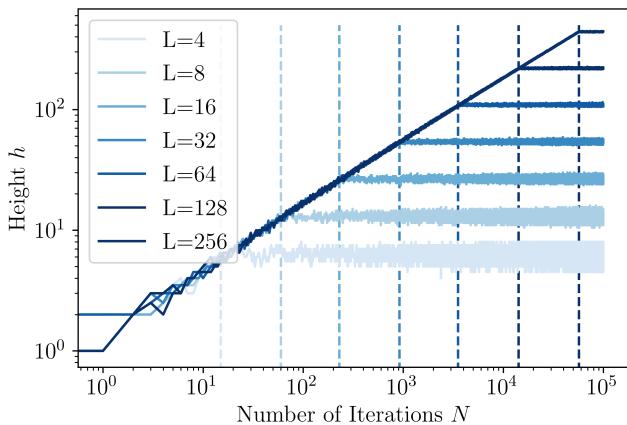


Figure 4: Logarithmic Plot of the pile height of the model with time for different system sizes  $L$ . Dashed vertical lines indicate when the system reaches the steady state.

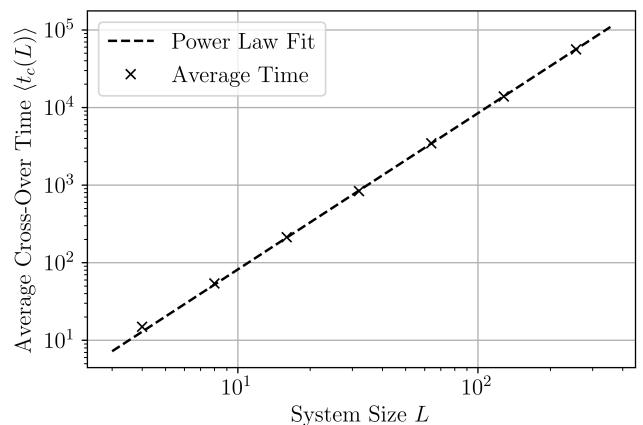


Figure 5: Logarithmic Plot of the average cross overtime for different system sizes with a power law fitted to the data.

## 2.3 Theoretical Explanations [Task 2c]

As we have seen in Task 2b, the cross-over time scales with  $L^2$ . A theoretical understanding of why it follows a quadratic relationship is in terms of the maximum and minimum number of grains in a system at the steady state. The maximum and minimum  $t_c$  can be calculated by assuming  $p = 1$  for minimum  $t_c$  and  $p = 0$  for maximum  $t_c$  where the thresholds are all 1 or all 2 respectively.

$$\begin{aligned}\min(t_c) &= \sum_{n=1}^L n = \frac{L}{2}(L+1) \\ \max(t_c) &= \sum_{n=1}^L 2n = L(L+1)\end{aligned}\tag{1}$$

By taking the average of these quantities, which is reasonable assuming the thresholds fall somewhere in an equal distribution of 1 and 2, we can calculate the average cross-over time to be proportional to  $0.75L^2$  which matches the data from Task 2b.

$$\langle t_c \rangle \approx \frac{1}{2}(\min(t_c) + \max(t_c)) \sim \frac{3}{4}L^2\tag{2}$$

We can devise a similar theoretical procedure to prove that the average height  $\langle h \rangle$  of the model in the steady state scales with the system size  $L$ . The mathematics is much simpler, we can define the height of a pile as the sum of its thresholds hence the average height is the average of the sum of the thresholds which is proportional to  $L$ .

$$h = \sum_{i=1}^L z_i \Rightarrow \langle h \rangle = \left\langle \sum_{i=1}^L z_i \right\rangle \sim L\tag{3}$$

## 2.4 Height Data Collapse [Task 2d]

We have described theoretically in Task 2c that the height scales with  $L$  and the cross-over time scales with  $L^2$ . This allows us to produce a data collapse for the height versus time curves that we saw in 4. This can be done by introducing the following scaling function:

$$\tilde{h}(t; L) = L\mathcal{F}(t/L^2)\tag{4}$$

where  $\tilde{h}(t; L)$  is the height data that has been smoothed out i.e. running the model  $M$  times and measuring the height for different random numbers that determine the threshold values. Fig. 6(a) shows the resulting data collapse when the data is inputted into the scaling function. It is clear from the graph that the respective scaling functions for height and cross-over time are correct as the data collapse has accurately superimposed the data for all system sizes. The graph can also tell us about the behaviour of the scaling function  $\mathcal{F}(x)$ . Specifically, that when  $x \ll 1$ ,  $\mathcal{F}(x)$  is well approximated by a power law and that when  $x \gg 1$ ,  $\mathcal{F}(x)$  is always constant. These domains are separated by a discontinuity at  $x = 1$  which signifies the system reaching the steady state. Fig. 6(b) shows the collapsed data during the transient phase for different sizes. On a logarithmic plot, the transient height form a straight line which means it follows a power law with exponent  $\beta$ . By fitting a power law with exponent  $\beta$  to the transient height data, it was found that  $\beta = 0.493$ . This suggests that  $\tilde{h}(t; L)$  increases with  $\sqrt{t}$  during the transient. The measured value is slightly less than 0.5 due to scaling errors in the height at small system sizes. It can be seen in Fig. 6(b) the data points under the line fit are a lighter blue and hence a smaller system size. These data points from smaller system sizes slightly reduce the expected value of  $\beta$ .

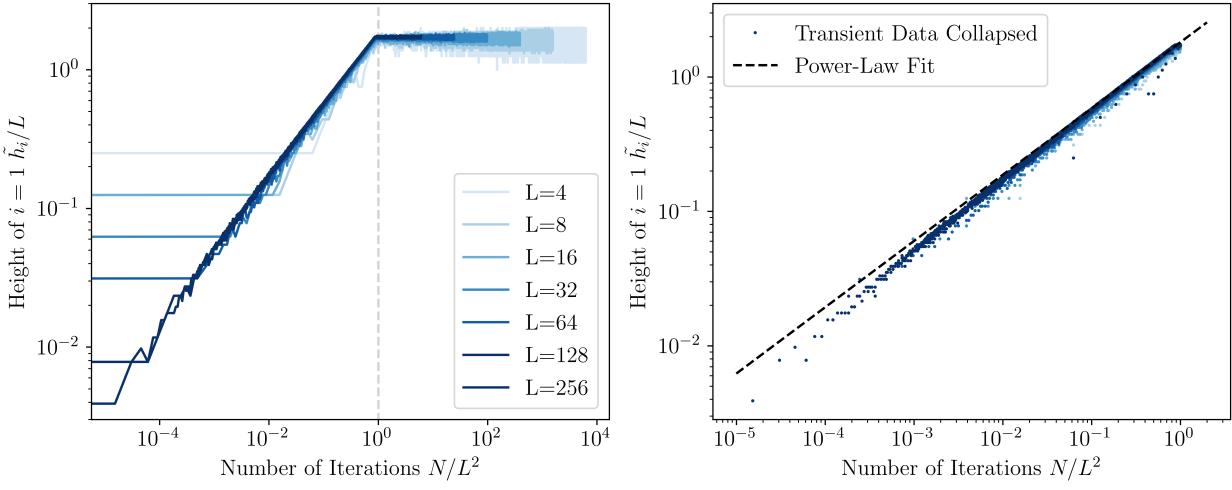


Figure 6: (a) Height versus time data collapse using the scaling function  $\tilde{h}(t; L) = L\mathcal{F}(t/L^2)$ . (b) Transient height versus time after the data collapse plotted to fit a power law to investigate transient behaviour.

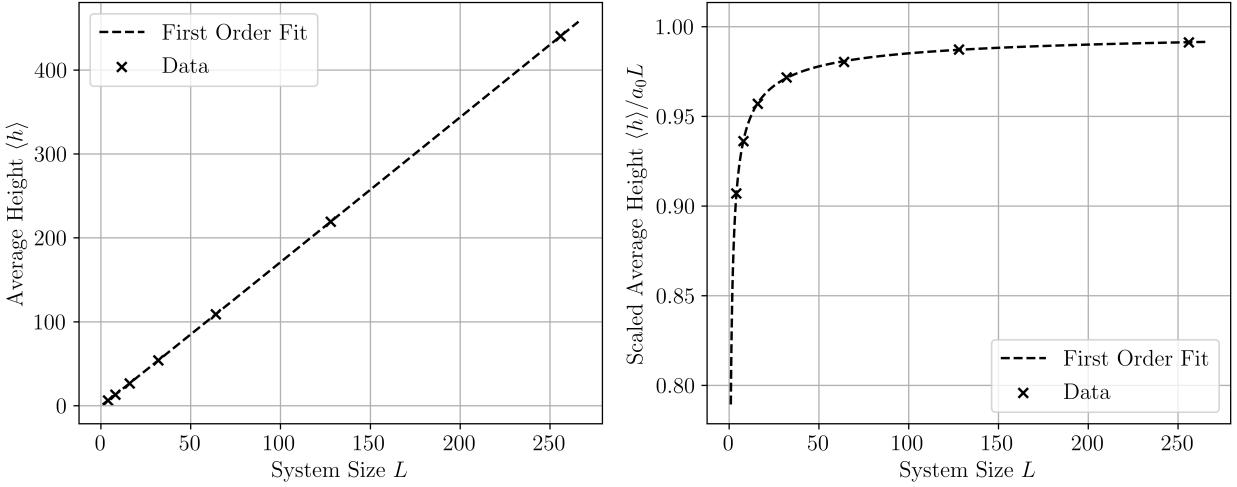


Figure 7: (a) Average height with system size fitted with a first order corrections to scaling curve. (b) Scaled average height with system size to eliminate the linear term which highlights corrections to scaling for small  $L$ .

## 2.5 Height Corrections to Scaling [Task 2e]

In the previous task, we discussed scaling errors that occur in the height of the pile. We will now investigate corrections to scaling in the height with the system size  $L$ . Firstly, we wish to measure the average height of a system in the steady-state  $\langle h \rangle$ . For the following calculations 10,000 samples of the height were used to calculate  $\langle h \rangle$  as this resulted in a high enough accuracy at a low computational cost. The average height was calculated for the same set of system sizes as we have been discussing throughout this task. Firstly, we assume the correction to scaling to take the following form:

$$\langle h(t; L) \rangle_t = a_0 L \left( 1 - \sum_{i=1}^{\infty} a_i L^{\omega_i} \right) = a_0 L (1 - a_1 L^{\omega_1} + a_2 L^{\omega_2} + \dots) \quad (5)$$

where  $a_i$  and  $\omega_i$  are constants that quantify the scaling. To test if there are signs of corrections to scaling, it is as simple as plotting  $\langle h \rangle$  for the various values of  $L$  and performing a fit of the correction function to first order - that is ignoring terms where  $i > 1$ . Fig. 7(a) shows this fit.

As expected, the height scales with  $L$  however signs of corrections at small  $L$  are hard to see. From the first order fit, three constants could be measured, namely,  $a_0 = 1.735$ ,  $a_1 = 0.212$  and  $\omega_1 = 0.573$ . To see signs of scaling, we can divide  $\langle h \rangle$  by the linear term  $a_0L$  to isolate the  $1 - a_1L^{\omega_1}$ . This analysis is performed in Fig. 7(b) where signs of scaling become much more apparent and that systems with sizes less than 20 can deviate by as much as 5% or more when linearly scaled from larger  $L$ .

## 2.6 Standard Deviation $\sigma_h$ with System Size [Task 2f]

The standard deviation in the height of the pile during the steady state  $\sigma_h$  can be calculated from the set of heights used to calculate  $\langle h \rangle$  in Task 2e. Due to the random nature of the Oslo model, the height during the steady state is hardly ever constant but varies around  $\langle h \rangle$ . The size of this variation can be studied as a function of  $L$ . Fig. 8(a) plots this data and fits a power law  $\sigma_h(L) = \alpha L^\beta$  to the straight line in the logarithmic plot. Through the fit,  $\beta$  was calculated to be 0.239 such that the scaling can be determined to be:  $\sigma_h \sim L^{0.239}$ . Fig. 8(b) applies a similar method to Task 2e to check if there are signs of corrections to scaling where we can see there are indeed no signs of corrections to scaling.  $\sigma_h$  can also be used to determine the average slope  $\sigma_z$  of the pile. Using the basic definition of the gradient slope that  $\langle z \rangle = \langle h \rangle / L$ , a similar analogy can be made for the standard deviation such that  $\sigma_z = \sigma_h / L$ . This would mean that using our relation for  $\sigma_h(L)$ ,  $\sigma_z \sim L^{\beta-1} \sim L^{-0.761}$ . Hence, in the limit that  $L \rightarrow \infty$ , the standard deviation of the slope will tend to zero, the average height would tend to  $a_0L$  hence the slope would tend to  $a_0$ .

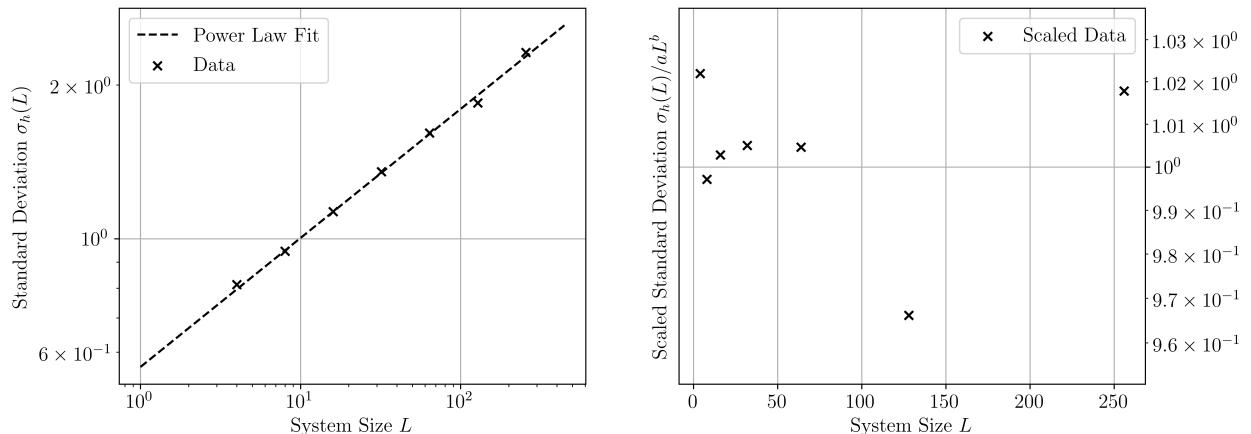


Figure 8: (a) The standard deviation of the height plotted against various system sizes on a logarithmic plot with a fitted power law. (b) Finding the residual data by dividing the data by the power law to check for signs of corrections to scaling.

## 2.7 Height Probability Distribution [Task 2g]

Let us consider that the thresholds of the Oslo model,  $z_i$ , are independent, identically distributed random with finite variance. The random variables  $z_i$  determine the height of the pile by the relation  $h = \sum_{i=1}^L z_i$ . By the Central Limit Theorem (CLT) for significantly large  $L$ , the probability distribution of heights  $P(h; L)$  can be modelled by a Gaussian with mean and standard deviation  $\langle h \rangle$  and  $\sigma_h$  respectively. The CLT will only hold if the thresholds are independently distributed. If there is correlation between the thresholds and there is dependence, the resulting distribution will be poorly approximated by a Gaussian. In the CLT theoretical framework, we expect  $\sigma_h \sim \sqrt{L}$  if  $z_i$  are indeed independent and identically distributed random variables which we know is not the case and requires further investigation.

We now move onto the task of plotting the probability distribution  $P(h; L)$  which we will use to perform a data collapse and test the Gaussian approximation. In order to get the distributions, the height data for each system size  $L$  must be binned such that each height  $h$  has a frequency associated with it and then divide by the total number of height data points. The result is a probability distribution for each value of  $L$  and can be seen in Fig. 9(a). If the theoretical explanation of these probability distributions above is correct the plotted distributions can be treated as Gaussians in order to produce a data collapse. Specifically, this means shifting the distribution by the measured mean  $\langle h \rangle$ , dividing the heights by the standard deviation  $\sigma_h \sim L^\beta$  and multiplying the probabilities by the standard deviation. In essence we are transforming the Gaussian distribution to a normal distribution of mean and standard deviation of 0 and 1. This procedure follows from the mathematical expression where a random variable  $X$  is mapped onto a normally distributed variable  $Z$  by the transformation  $Z = (X - \mu)/\sigma$  and  $P(Z) = \sigma P(X)$ .

The data collapse shown in Fig. 9(b) is an indication that the measured values of  $\langle h \rangle$  and  $\sigma_h$  in previous tasks are correct, however, the fit is not perfect. Notably, the scaled distributions for smaller system sizes have lower peaks and slightly wider distributions than the for the higher values of  $L$ . Secondly, by comparing with the plot of the theoretical normal distribution, it is clear that the curves do not match the expected distribution. This suggests that  $p(h; L)$  cannot be approximated by a Gaussian and that the assumption that the threshold random variables  $z_i$  are not independent and identically distributed but are dependent and correlated in some way. The dependence of the threshold values could arise as a result of the boundary as relaxation of the model depends on the slopes of neighbouring columns which provides a different effect at the boundaries which does not have an effect in the middle of the pile - this effect also leads to the corrections to scaling for the height at lower system sizes. A numerical test that could be implemented to investigate the dependence caused by this effect is measuring the average slope along different regions of the pile. An average slope that changes based on proximity to the boundary would prove the existence of this effect

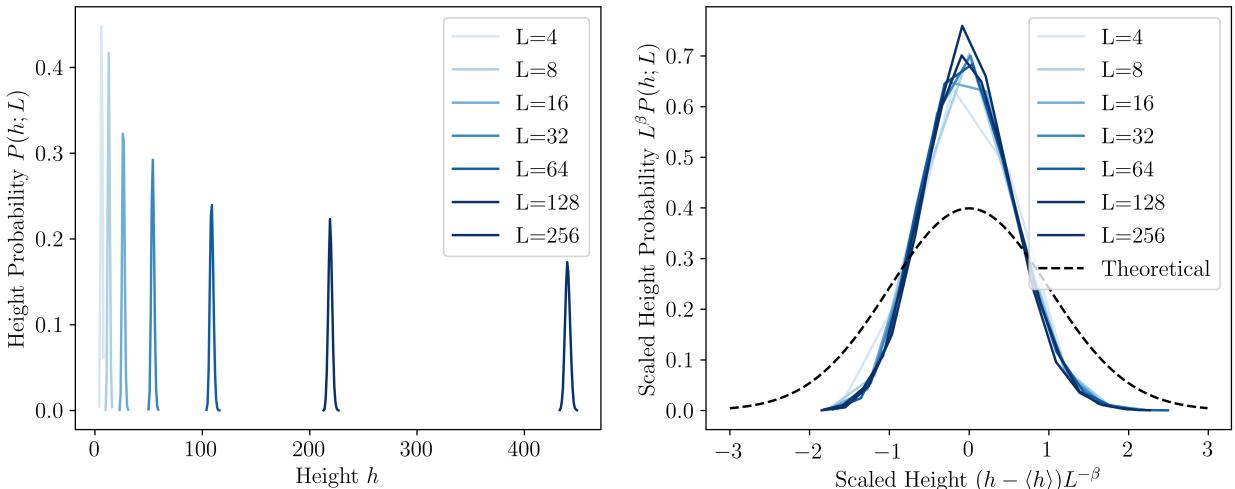


Figure 9: (a) Average height with system size fitted with a first order corrections to scaling curve. (b) Scaled average height with system size to eliminate the linear term which highlights corrections to scaling for small  $L$ .

### 3 The Avalanche-Size Probability $P(s; L)$

#### 3.1 Avalanche Probability Distribution [Task 3a]

The avalanche size is a measure of the total number of relaxations performed in a given iteration. By running the model for 500,000 iterations past the steady state time, enough data was collected to plot the avalanche-size probability for the various system sizes used in Task 2. The probability distribution can be found by performing a log-binning of the avalanche-size data with a scaling parameter  $\lambda$ .  $\lambda$  scales the size of each bin relative to the size of the previous bin such that the sizes of the bins increase exponentially. If  $\lambda = 1$ , then the result is the frequency of each unique integer avalanche size as the bin size does not increase from the initial value of 1. Fig. 10 shows avalanche-size probability for  $\lambda = 1$  and  $\lambda = 1.3$ . For  $\lambda = 1$ , a scatter graph is produced such that each point represents an individual avalanche size - the larger the avalanche size the lower the avalanche probability. This trend occurs up until a cut-off avalanche size  $s_c$  where the probabilities become very small. These minimum probability values occur for the lowest frequency of avalanche sizes. The lowest line occurs due to avalanche sizes of frequency 1, the next line above of frequency 2 and so on. A large proportion of avalanches occurring with a low frequency causes the dips that we see for  $\lambda = 1.3$  where we can clearly see a power law decrease in the probability distribution. As expected, the larger the system size, the larger the avalanche probability distribution hence the smaller probabilities for these larger avalanches with a higher  $t_c$ .

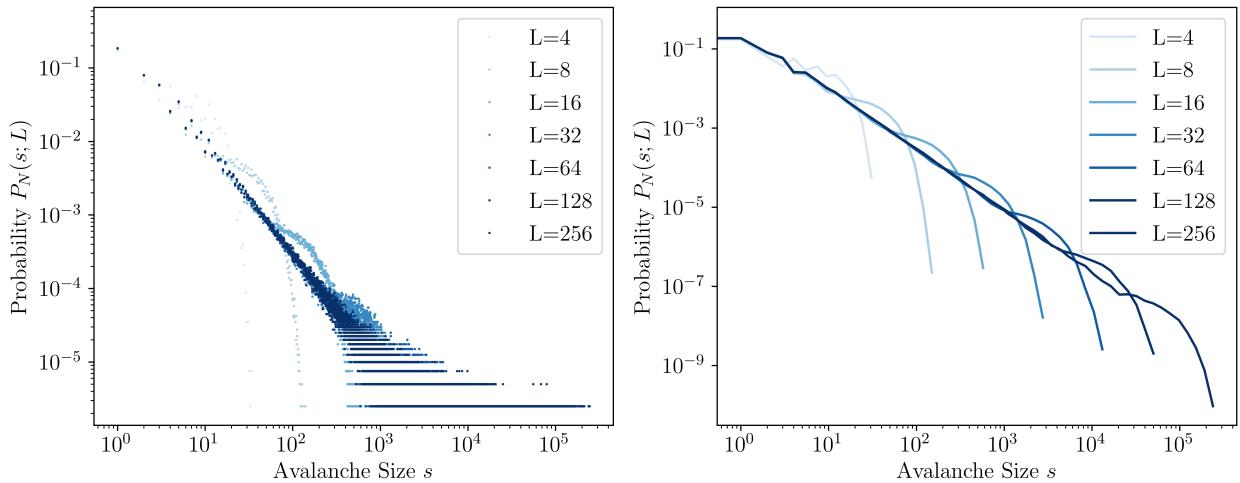


Figure 10: Avalanche-size probability distributions for various system sizes  $L$  with scaling parameters (a)  $\lambda = 1$  and (b)  $\lambda = 1.3$ .

Two parameters were chosen to produce the data in this graph to ensure high quality. That is the number of samples  $N$  and the bin-scale parameter  $\lambda$ . As stated,  $N = 400000$  was decided to be large enough such that enough large avalanche sizes occur so that the dip in the probability curve is smooth. Additional improvements in the smoothness of the curve were very computationally expensive. Secondly, the scaling parameter  $\lambda$  was chosen to be 1.3. In order for the graph to be smooth,  $\lambda$  could not be too large as to reduce the number of bins, but similarly not too small to pick up higher frequency noise. Fig. 11 shows the probability distribution for  $L = 128$  with different values of  $\lambda$  and how the noise is reduced for increases in the value of  $\lambda$ .

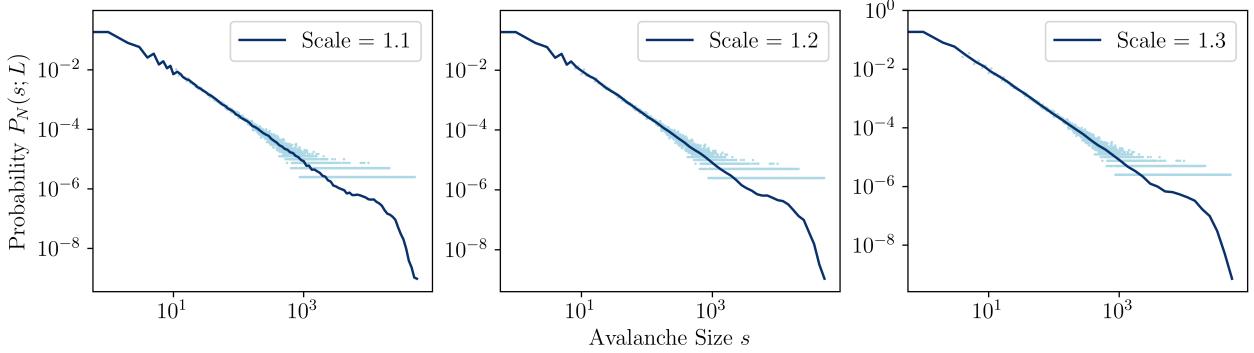


Figure 11: Variations in the avalanche-size probability distributions for a system of size 128 for bin-scaling parameters 1.1, 1.2 and 1.3. For lower values of  $\lambda$ , noise is picked up that follows the data from  $\lambda = 1$ .

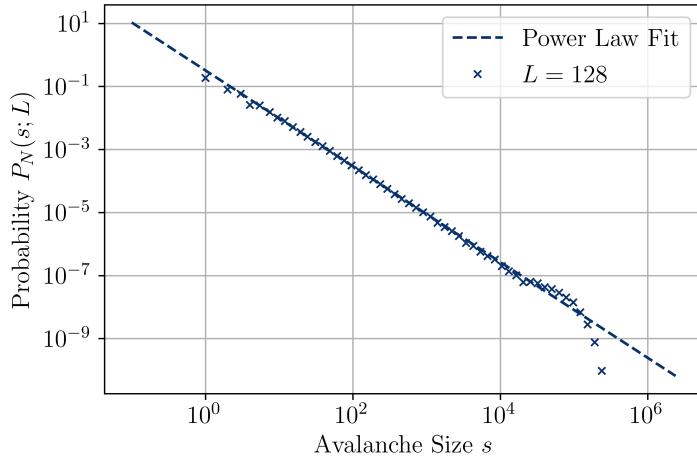


Figure 12: The avalanche-size probability for a system of size 256 with a power law approximating the straight line behaviour

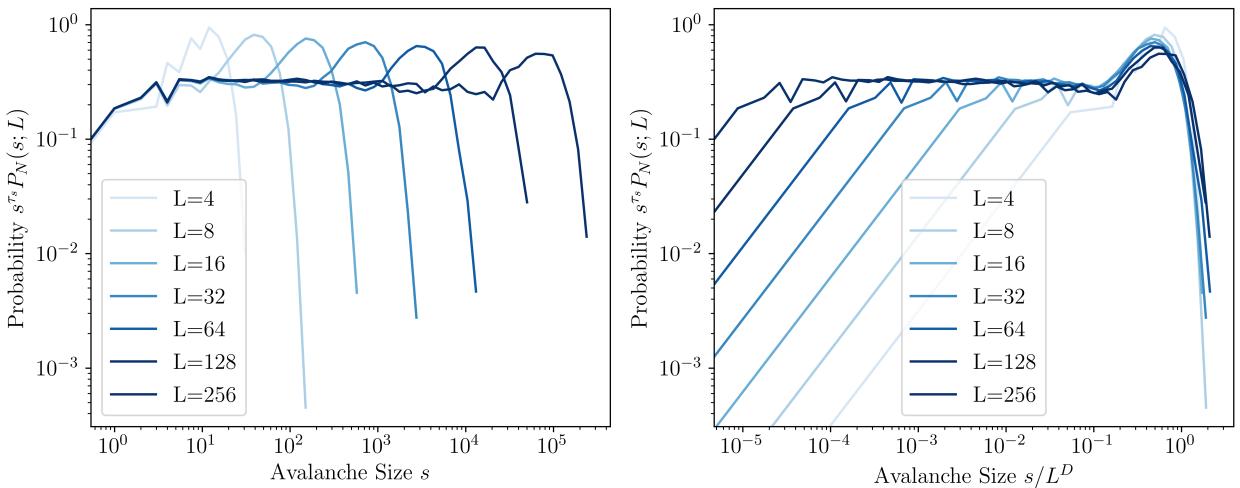


Figure 13: Data collapse for the avalanche-size probability distributions. Initially scaling the probabilities by  $s^{\tau_s}$  (a) and then the avalanche sizes by  $1/L^D$  (b).

We now consider the finite-size scaling ansatz:

$$P_N(s; L) \propto s^{-\tau_s} \mathcal{G}(s/L^D) \quad L \gg 1, s \gg 1, \quad (6)$$

and wish to check whether the probability distributions are consistent with this scaling ansatz and to determine the values of the parameters  $\tau_s$  and  $D$ . As we have discussed and seen in Fig. 10, the straight line section of the graph can be modelled by a power law that is present in the scaling ansatz. Hence,  $\tau_s$  can be determined by fitting a power law to this section of the curve. The curve with the maximum number of data points will give the most accurate value of  $\tau_s$  therefore, the analysis was conducted on the curve for  $L = 256$  and can be seen in Fig. 12. From this fit  $\tau_s$  was measured to  $1.520 \pm 0.008$ . To test whether the scaling ansatz is correct we can use this measured value of  $\tau_s$  by performing a data collapse. This is done by multiplying the probabilities i.e. the y-axis, by  $s^{\tau_s}$  and can be seen in Fig. 13(a). The data collapse correctly superimposes the straight line behaviour of the probability curves showing that the  $s^{-\tau_s}$  term in the scaling ansatz is consistent. All that is left to do in the data collapse is to collapse the avalanche-sizes by a factor  $1/L^D$ . The value of  $D$  was found by trying to minimise the distances between the peaks of the 'bumps' in the distribution, this value was found through trial and improvement to be  $D = 2.1$ . Fig. 13(b) shows the final result of the data collapse which proves the consistency of the scaling ansatz as the curves are relatively superimposed.

### 3.2 $k$ -Moment Analysis [Task 3b]

A more quantitative way of determining the scaling ansatz parameters is using a  $k$ -moment analysis. The  $k$ 'th moment is defined as:

$$\langle s^k \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=t_0+1}^{t_0+T} s_t^k \quad (7)$$

where  $t_0$  is the time that the system reaches the steady state. The  $k$ 'th moment data can be found rather easily using the avalanche size data from Task 3a and finding the average of  $s^k$ . In my analysis I used 4 values of  $k$  specifically,  $k = 1, 2, 3$  and  $4$ . The  $k$ 'th moment for various system sizes can be seen in Fig. 14. The  $k$ -moments increase with system size  $L$  because as previously mentioned the larger system size, the larger the avalanche sizes. The  $k$ -moment analysis to determine  $D$  and  $\tau_s$  is as follows. From the graph of  $k$ -moments against system size  $L$  (Fig. 14) we calculate the gradient of the graph  $\log \langle s^k \rangle$  versus  $\log L$  for each value of  $k$ . From the definition of the  $k$ 'th moment a scaling relation can be derived for  $\langle s^k \rangle$  and  $L$  where we can take the logarithm of both sides to get it into a more useful form.

$$\langle s^k \rangle = L^{D(1+k-\tau_s)} \int_{1/L^D}^{\infty} u^{k-\tau_s} \mathcal{G}(u) du \quad (8)$$

$$\log \langle s^k \rangle = D(1 + k - \tau_s) \log L + \log \left( \int_{1/L^D}^{\infty} u^{k-\tau_s} \mathcal{G}(u) du \right) \quad (9)$$

Hence, the gradients of the  $\log \langle s^k \rangle$  versus  $\log L$  graph is equal to  $D(1 + k - \tau_s)$ . By plotting the  $k$ -slope gradients against the  $k$  moments, we are also plotting  $D(1 + k - \tau_s)$  against  $k$  which should produce a straight line with gradient  $D$  and a  $y$ -intercept of  $D(1 - \tau_s)$ . This analysis was performed and can be seen in Fig. 15. By performing a linear fit, the value of  $D$  was calculated to be  $D = 2.13 \pm 0.01$  and using the fitted value of the  $y$ -intercept,  $\tau_s$  was calculated to be  $\tau_s = 1.54 \pm 0.05$ . We can compare our findings from the data collapse in Task 3a. Firstly, the value of  $\tau_s$  found through the  $k$ -moment analysis is less accurate than the one previously found ( $\tau_s = 1.520 \pm 0.008$ ) however this value does fall within the error range of the  $k$ -moment value.

In terms of the parameter  $D$ , it is of course much more accurate via the  $k$ -moment analysis rather than the trial method used in Task 3a, however, the values are similar which gives us confidence that the value is indeed correct.

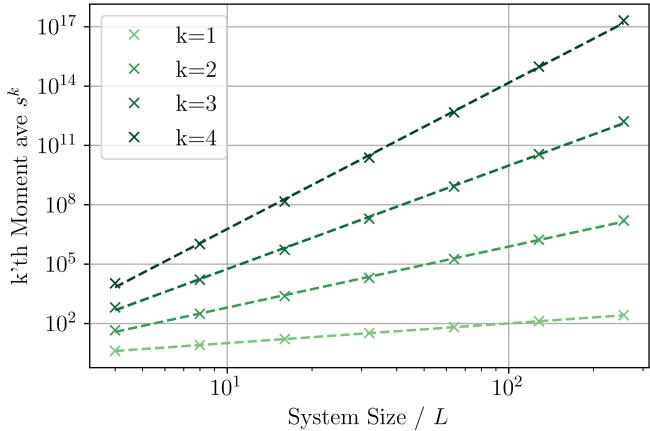


Figure 14: Logarithmic plot of the  $k$ 'th moment for different system sizes with a power law fitted for all values of  $k$ . these power laws form straight lines with gradients equal to  $D(1 + K - \tau_s)$ .

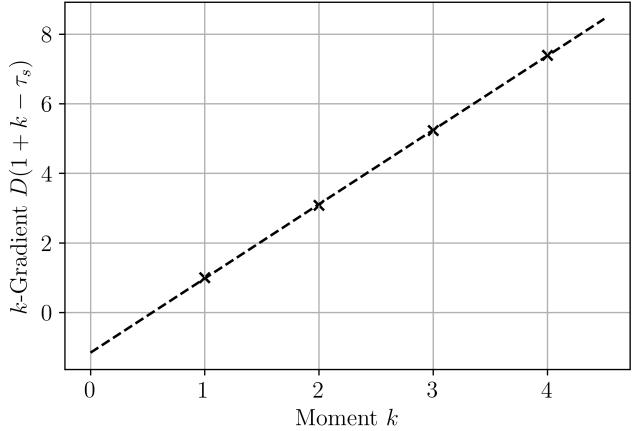


Figure 15: Height of the pile against number of iterations showing how the height of the pile increases in the transient phase and plateaus during the steady state. Vertical Lines indicate the time when the system reaches the steady state

## 4 Summary

This report has shown that a proven working programme implementation of the Oslo Model Algorithm can give enormous insight into the dynamics of the model. Investigations into the height probability has given rise to powerful scaling relations motivated by power laws and have shown dependence of the threshold random variables. As well as the avalanche size study that has proven the validity of the finite-size scaling ansatz and the power of  $k$ -moment analysis to determine scaling parameters.