Complexity Project
The Oslo Model

CID: 01702898

18th February 2022

Abstract: For this report we implemented the Oslo model for systems sizes L = 4, 8, 16, 32, 64, 128, 256, 512. We used a data collapse to identify the finite-size scaling ansatz for the height of a pile, obtaining $\tilde{h}(t;L) = L\mathcal{F}(t/L^2)$, and investigated the behaviour of the scaling function for $t/L^2 \ll 1$ and $t/L^2 \gg 1$. We also used data collapse and moment analysis methods to confirm the finite-size scaling ansatz for the avalanche size probability and retrieve critical exponents. For the finite size ansatz of $\tilde{P}_N(s;L) \propto s^{-\tau_s}\mathcal{G}(s/L^D)$, the data collapse method retrieved $D = 2.12$ and $\tau_s = 1.56$, and the moment scaling analysis retrieved $D = 2.19$ and $\tau_s = 1.58$. These values are close to the theoretical values of $D = 2.25$ and $\tau_s = 1.55$.

Word count: 2412 words in report (excluding front page, figure captions, table captions, acknowledgement and bibliography).

1 Introduction

The Oslo model is one of the simplest models to display self-organised criticality (SOC), which is the property that slowly driven systems exhibit avalanches and have power law correlated interactions [4]. Avalanches here mean a sudden recurrent modification in the internal energy of a system. Many slowly driven non-equilibrium systems exhibit SOC, they approach their critical state with no external fine-tuning and exhibit scale-free behaviour. These systems are prevalent in nature as real systems not truly isolated and therefore may be being slowly driven. Examples of systems that exhibit SOC are diverse, from sandpiles to earthquakes to the motion of magnetic flux lines in superconductors, etc. [1]. Another example is a slowly driven pile of rice, which was first investigated at the University of Oslo in 1996 [3].

The boundary-driven Oslo model serves to model the ricepile experiment, which is discussed in [2]. For a model of system size L, the sites where a grain may exist are denoted $i = 1, 2, 3, ...L$, each site i has a slope threshold value that is a random choice of either 1 or 2. As the model is driven at site $i = 1$, the grains pile. For any site i, if the slope z_i is greater than the threshold value for the site, i.e. the slope is supercritical: $z_i > z_{th,i}$, the site at i relaxes. The relaxation of a site may cause a neighbouring slope to increase beyond the threshold slope, triggering an avalanche. Once a site has relaxed, a new slope threshold value is chosen for it, again either 1 or 2.

For this report the models implemented had system sizes of $L = 4, 8, 16, 32, 64, 128, 256, 512$.

1.0.1 Report Structure

This contains four key sections

- Implementation of the Oslo Model.
- Analysis of the height of the pile *h(t; L)*.
  - Task 2a
  - Task 2b
  - Task 2c
  - Task 2d
  - Task 2e
  - Task 2f
  - Task 2g
- Analysis of the avalanche size probability*P(s;L)*
  - Task 3a
  - Task 3b
- Conclusion

1.1 Implementation of the Oslo model.

Using the Oslo model algorithm outlined in [5], the model was implemented in `python`. As the time taken for model realisations increased with L, the `Numba` package was implemented in order to increase the effiency of simulations. The package uses a just-in-time compiler to optimise `python` functions, by using the LLVM library to generate machine code equivalent functions[6].

Tests were implemented to ensure that the model ran as required. Tests included checking that:

1. $0 \leq z_i \leq 2$
2. $L \leq h_{i=1}$ in recurrent regime $\leq 2L$
3. $\langle h \rangle \approx 26.5, 53.9$ for $L = 16, 32$, respectively
4. $\langle s \rangle \approx L$, where $\langle s \rangle$ is the time averaged avalanche size in the recurrent regime
5. implementing the boundary-driven BTW model (d=1) returns the expected staircase recurrent configuration.

Test 1 must be satisfied as a slope of 3 is supercritical for any choice of $z_{th} \in \{1, 2\}$, and so the site will always be relaxed before a new grain can be added to the site. Moreover, the slopes can never be negative, leaving viable slope values of $z_i \in \{0, 1, 2\}$. Test 2 must be satisfied as the minimum and maximum $h_{i=1}$ are achieved for $z_i = 1 \forall i$ and $z_i = 2 \forall i$, respectively. Since $h = \Sigma_{i=0}^{L} z_i$, we obtain $h_{i=1,min} = L$ and $h_{i=1,max} = 2L$. Test 3 was included as part of a hint in [5]. Test 4 must be satisfied as in the recurrent state $\langle$ influx of grain$\rangle$ = $\langle$ outflux of grain $\rangle$, for this to be true, on average for each grain added a single grain must leave. The number of relaxations required for this to occur is L, hence $\langle s \rangle \approx L$. This relation is shown for $k = 1$ for $\langle s^k \rangle$ in Figure12, where the gradient is measured to be 1.001.

Test 5 was implemented by manipulating the z_{th} choosing function in the `python` Oslo Model class, such that $z_{th} \in \{1\}$. Therefore, the initial set of threshold slopes were all 1 also. Recording the z_i for recurrent configurations revealed $z_i = 1 \forall i$, giving rise to the expected staircase configuration.

2 The height of the pile *h(t; L)*

The height of the pile, h(t; L), is defined as the height at site i = 1 at time t once the pile has come to rest after adding the t^{th} grain. $h(t; L)$, is calculated by summing the slope values at all the sites:

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

2.1 Task 2a

For each L, the plot of height against time exhibits two regions, as shown in Figure1, a region where the height trend is increasing and a region where the heights vary about a constant value. The increasing trend in height is representative of the transient regime for each system. In this regime, as grains are added, more sites become available, so the sum over slopes to calculate height is over a larger number of sites. As a result, height increases. As each system is slowly driven, they approach the attractor of the dynamics, which is the set of recurrent configurations. For these configurations, the number of grains in the system approaches a steady state. For systems of all sizes, the transient follows the same trend, only systems with smaller L branch off into the recurrent configurations earlier in time. The transient indicates the scale-free nature of the Oslo model, whereas the recurrent regime is a result of the finite-size of the systems being simulated.

The time at which a system enters the set of recurrent configurations from the set of transient configurations is know at the cut-off time, t_c and is discussed further in subsection 2.2.

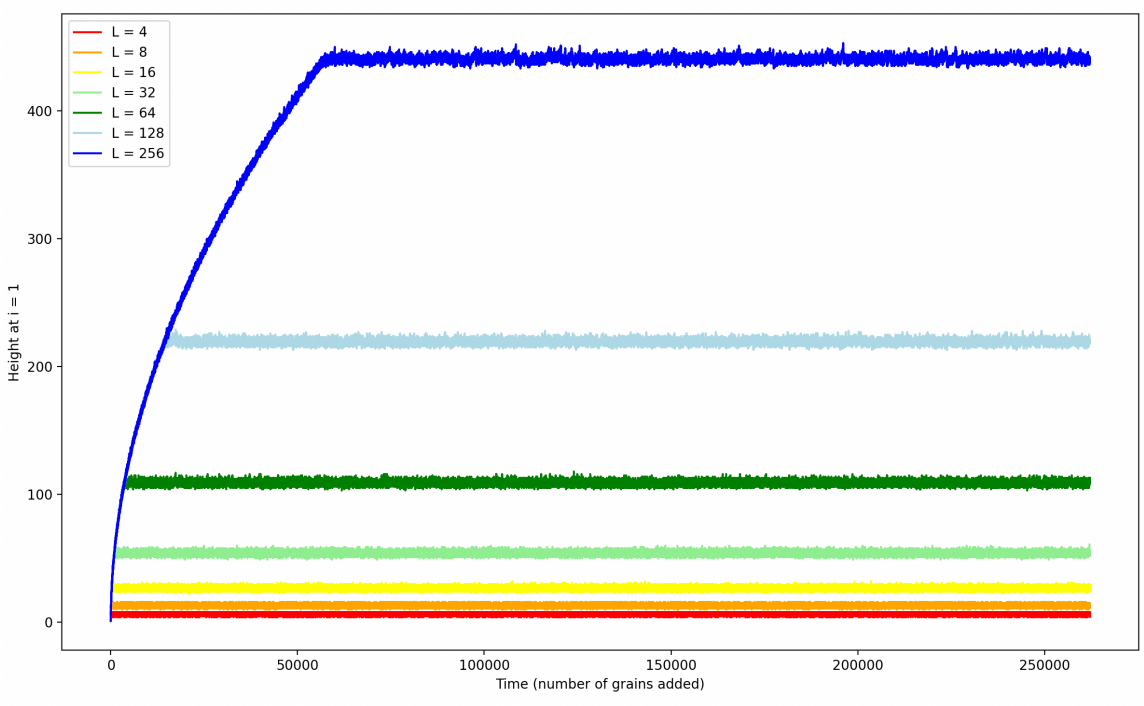

Figure 1: A plot of the height of pile at site i = 1 over time, for system sizes L = 4,8,16,32,64,128,256

2.2 Task 2b

The cross-over time t_c is defined as the number of grains added to a system before an added grain induces a grain to leave the system for the first time. We can approximate the average cut-off time for a system of size L, $< t_c >$ by running multiple realisations of the Oslo model for each system size, and recording the cut-off time for each realisation and then taking the arithmetic mean of these values. For each system size we simulated 10 realisations. In Figure2 the $< t_c >$ for each of the system sizes is plotted with log-scaled x and y axis. This reveals the scaling of $< t_c >$ with L by calculating the gradient of the log-log plot. The gradient was calculated using a linear least-squares fit for $L \geq 64$ and the resulting scaling relation was determined to be $< t_c > \propto L^{2.006}$. The error in the exponent of L was determined to be 0.2×10^{-6}. Therefore, the expected value of 2 is not within the measured range. This is suspected to be due to remaining scaling corrections due to the finite-size of the system; using even larger system size would produce a result more consistent with our expectation.

2.3 Task 2c

For $L \gg 1$ we can ignore potential corrections to scaling and aim to determine the scaling behavior for average height of the pile in the steady state, $h(t > t_c; L)$ and for the average cross-over time, $\langle t_c \rangle$.

Since $h(t > t_c; L)$ is determined as expressed in (1) which can be rewritten as

$$h(t; L) = \bar{z}(t)L, \tag{2}$$

where $\bar{z}(t)$ is the arithmetic mean slope for the configuration at a given time t. The height, $h(t > t_c; L)$, scales linearly with L under the assumption that $\bar{z}(t)$ is independent of L for $L \gg 1$.

The average cross-over time, $\langle t_c \rangle$ scales as L^2. The cross-over time can be calculated by counting the number of grains in the first critical pile with supercritical site $i = L$, since no

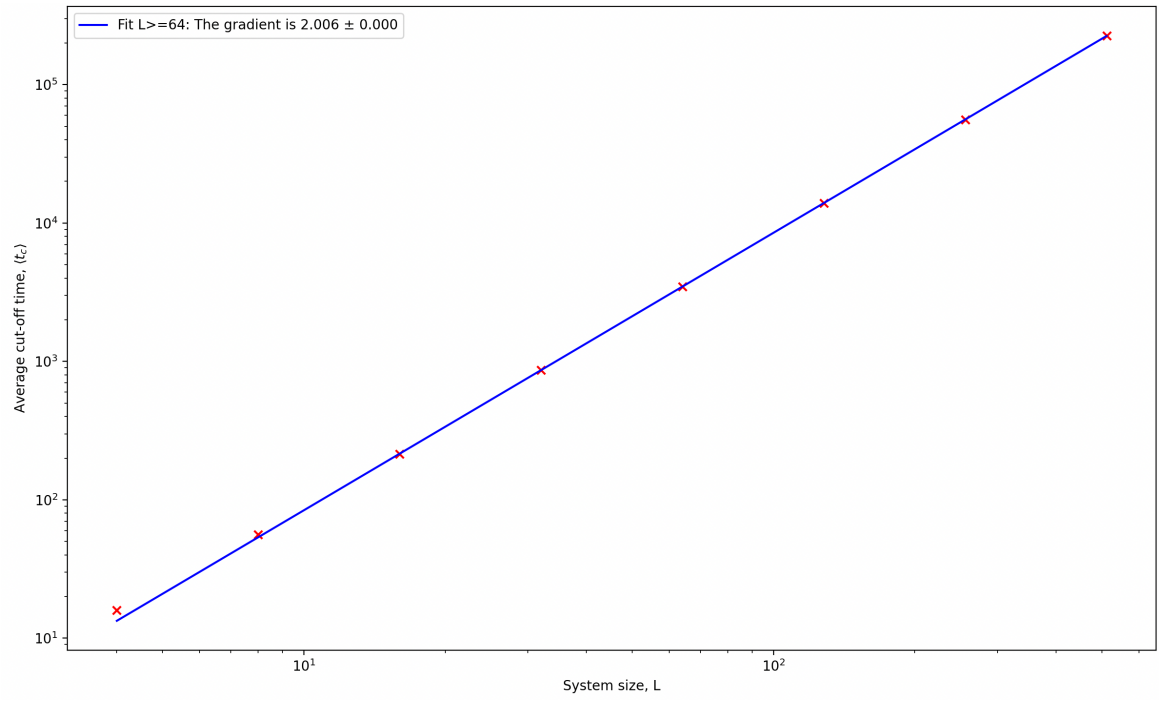

Figure 2: A plot of the average cut-off time for each system size with log-scaled axes.

grains would have left the pile in the transient configuration. The number of grains is in essence given by the 2D area of the pile, which is approximately formed of a right angled triangle with a 'rough' surface, where the relaxations occur. The area of the triangle is determined by $(base \times height)/2$, where the base must be length L, since for a grain to fall out, the site at i = L must be supercritical, so must have a slope of at least 1. As argued above, height must scale linearly with L, therefore the area of the triangle scales with L^2. For larger system sizes, the proportion of the pile area occupied by the surface in comparison to the right angled triangle becomes negligible, hence the area of the pile, and therefore t_c, tends to the area of the triangle which scales as L^2.

2.4 Task 2d

Firstly, the height data for each system size was smoothed by taking the average of the height at each time across multiple realisations i.e.

$$\tilde{h}(t;L) = \frac{1}{M}\Sigma_{j=1}^{M} h^j(t;L),$$

where $h^j(t;L)$ is the height at time t in the j^{th} realisation of a system of size L. We used 10 realisations for each system size. A data collapse was produced by scaling the processed heights and the time for each system size in the following way: $\tilde{h} \to \frac{\tilde{h}}{L}$, $t \to \frac{t}{L^2}$, given the scaling relations theoretically argued in subsection 2.4. This could be expressed mathematically as

$$\tilde{h}(t;L) = L\mathcal{F}(x),$$

where

$$x = \frac{t}{L^2}.$$

For $x \ll 1$, each system is in the transient regime. Earlier, in subsection 2.1, we noted the scale-free behaviour of the transient regime by observing Figure1. This indicates that

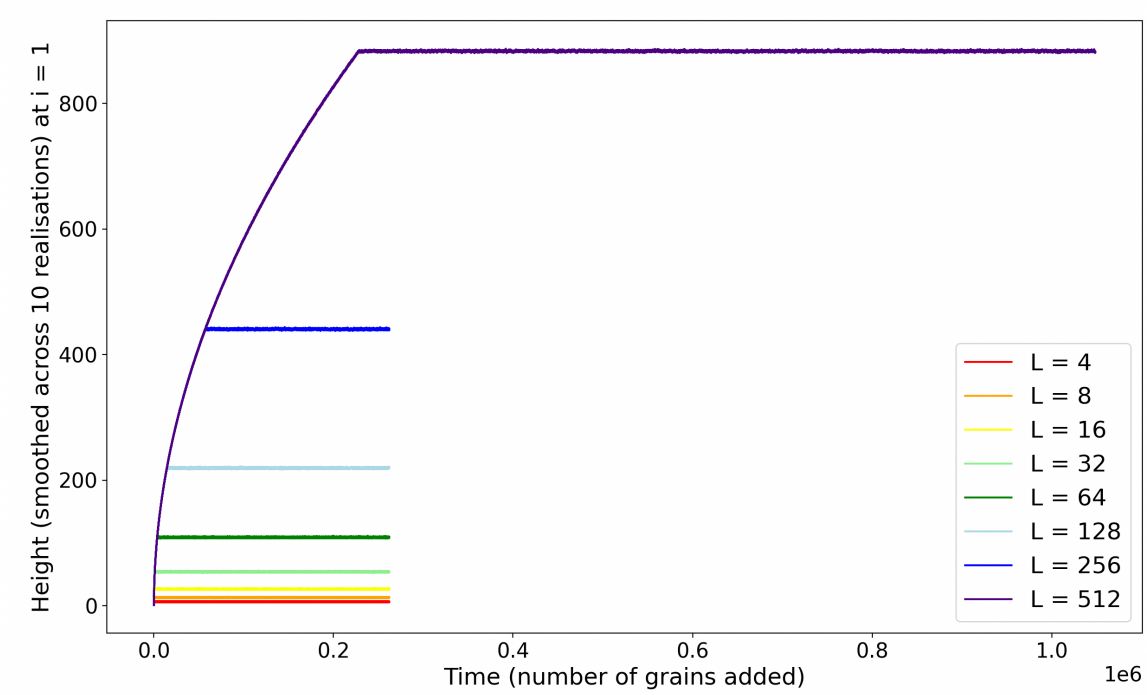

Figure 3: A plot of heights at site i = L, averaged over 10 realisations, against time. Vertical lines are the measured t_c for each system size.

$\tilde{h}(t < t_c; L)$ is independent of L, therefore, $\mathcal{F}(x) \propto 1/L \propto x^{1/2}$. As a result, we determine that $\tilde{h}(t < t_c; L) \propto t^{1/2}$ for the transient, by using the definition of x.

For $x \gg 1$, each system is within the recurrent regime. From the data collapse we note that the scaled steady state heights $h(t > t_c; L)/L$ for all systems tend to the same constant as x increases. This indicates that $\mathcal{F}(x)$ is constant.

2.5 Task 2e

We also investigated the scaling of time-averaged height, $\langle h(t; L)\rangle$, in the recurrent state. We assume the following corrections to scaling;

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

Ignoring higher order corrections with $i > 1$, and rearranging, we obtain the following:

$$y = 1 - \frac{\langle h(t; L)\rangle_t}{a_0 L} = a_1 L^{-\omega_1},$$

or equivalently,

$$log(y) = log(1 - \frac{\langle h(t; L)\rangle_t}{a_0 L}) = log(a_1) - \omega_1 log(L).$$

For sufficiently large L, we assumed all corrections of scaling to be negligible. Therefore, by plotting $\langle h(t; L)\rangle$ against L and recording the gradients for least-squares linear fits for $L \geq 64$, we determined $a_0 = 1.729 \pm 0.001$. This is demonstrated in Figure4.

The value of a_0 and its error could then be used to calculate y, and by measurement of the gradient of the curve $log(y)$ against $log(L)$ we determined $\omega_1 = 0.605 \pm 0.005$. The gradient was measured using a linear least-squares fit for $L \geq 64$. Only large values of L were fitted because smaller values of L are more susceptible to higher order scaling terms that we ignored. This method is shown in Figure5.

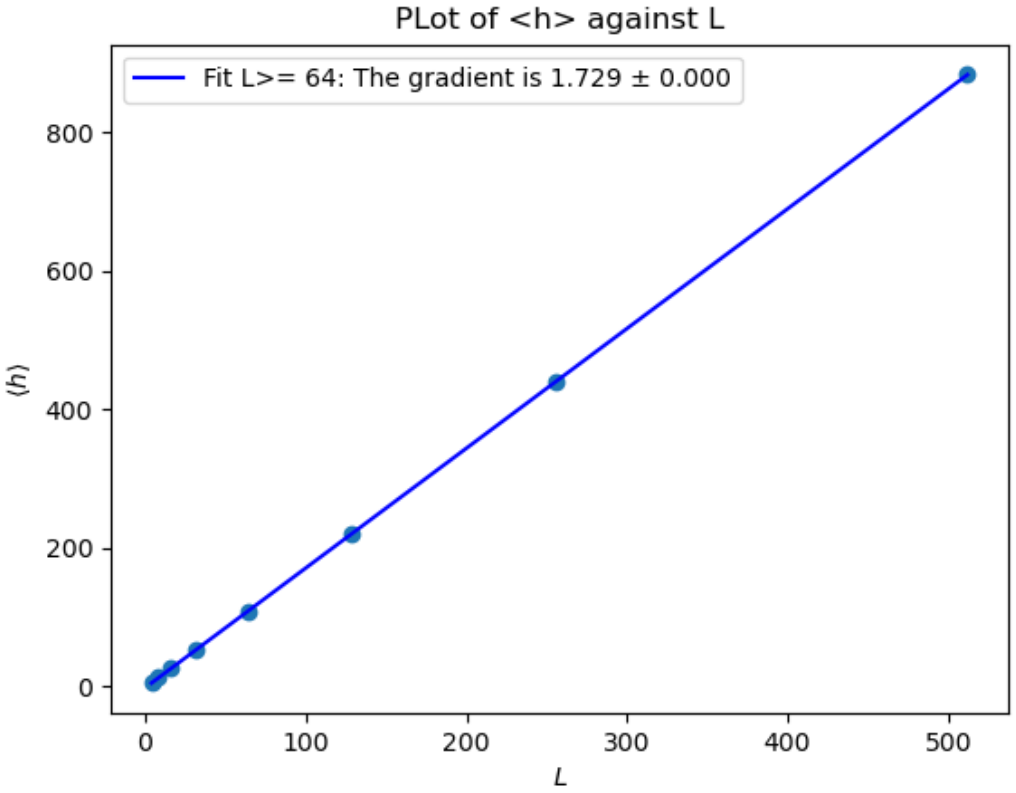

Figure 4: Plot of $\langle h(t;L)\rangle_t$ against L.

2.6 Task 2f

The steady state heights for each system vary about the value $\langle h(t;L)\rangle_t$. The standard deviations of the steady state heights, $\sigma_h(L)$, as defined in equation (6) of [5], were also measured for systems with varying L. A plot of $\sigma_h(L)$ against L is shown in Figure6. Gradient measurement of the log-scaled plot revealed the scaling of $\sigma_h(L)$ to be $\propto L^{0.237}$. The error in the gradient determined from the covariance matrix of the linear fit was 1.57×10^{-6}.

In the limit $L \to \infty$, corrections to scaling vanish and the time averaged steady state height can be written as

$$\langle h(t;L)\rangle = a_0 L = \langle \bar{z}(t)\rangle_t L, \tag{3}$$

where $\langle \bar{z}(t;L)\rangle_t$ is the time-averaged value of $\bar{z}(t;L)$ in the recurrent regime, and where the second relationship is determined from (2). That is $\langle \bar{z}(t;L)\rangle_t$ will tend to a_0. The standard deviation of the average slope, $\sigma(\bar{z}(t;L))$, as it is measured over time, will scale as $\sigma(\bar{z}(t;L)) \propto L^{0.237}/L = L^{-0.763}$, as determined by using the property of standard deviation for multiplied distribution, namely $\sigma(AX) = A\sigma(X)$, where X is a random variable and A is a constant, and the relation in (2).

2.7 Task 2g

By assuming that z_i are independent, identically distributed random variables with finite variance, we expect the probability of measuring a given h, $P(h;L)$ to approach a normal distribution for $L \gg 1$, as a result of the Central Limit Theorem.

The assumption yields that $\sigma\bar{z}$ scales as $\frac{1}{L^{1/2}}$, since L acts as the sample size, as it is the size of the configuration. σ_h therefore scales as $L \times \frac{1}{L^{1/2}} = L^{1/2}$, by using the relation (2).

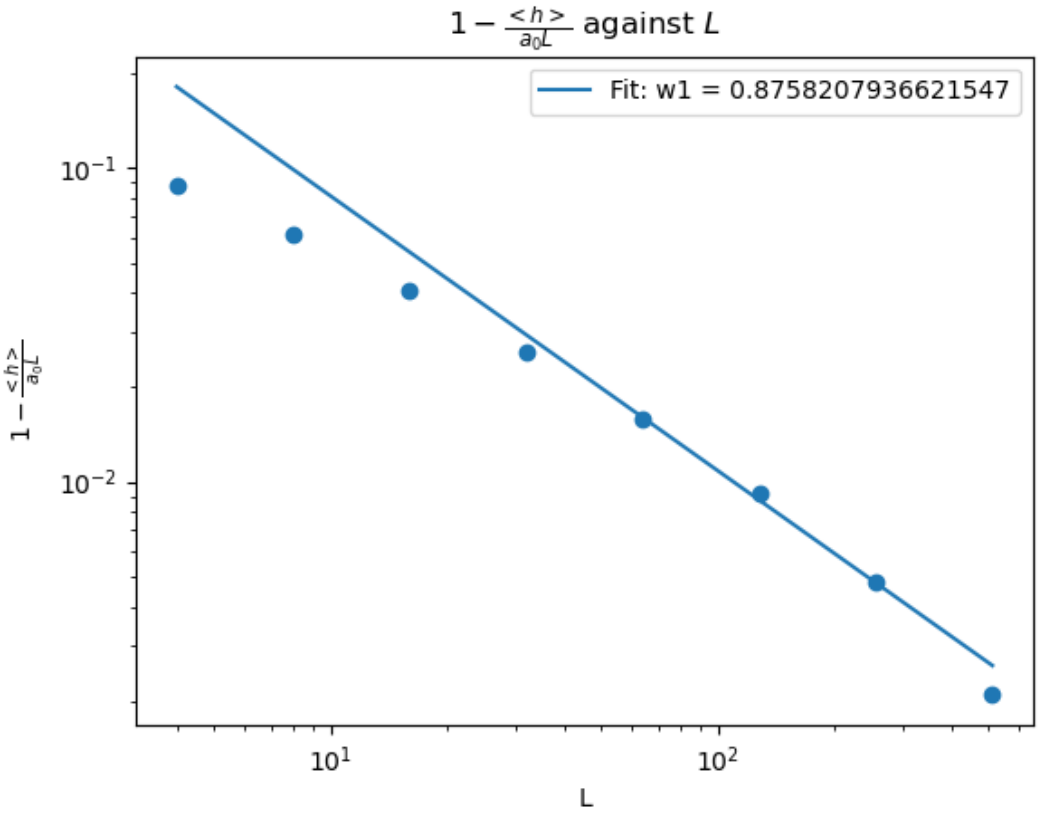

Figure 5: Plot of $1 - \frac{\langle h(t;L)\rangle_t}{a_0 L}$ against L, with log-scaled x and y axes

The measured height probabilities for different system sizes are shown in Figure7. By assuming the theoretically determined form of $P(h;L)$, given the above assumptions for the distribution of z_i, a data collapse for the varying system sizes can be accomplished by using the standard transformation of a Gaussian with mean μ and standard deviation σ to the canonical Gaussian with $\mu = 0$ and $\sigma = 1$. This involves transforming h and $P(h;L)$ in the following way: $h \to \frac{h-\langle h\rangle}{\sigma_h}$ and $P(h;L) \to \sigma_h P(h;L)$, in order to realise the underlying Gaussian scaling function shown below:

$$P(h;L) = \frac{1}{\sigma_h\sqrt{2\pi}} exp(-\frac{1}{2}\frac{(h-\langle h\rangle)^2}{\sigma_h^2}).$$

The resulting data collapse is shown in Figure8, with the canonical Gaussian also plotted.

The above data collapse does not trace out the Gaussian function that would be expected from the assumption that z_i are independent, identically distributed random variables with finite variance. Although similar in shape, the actual probability distribution exhibits negative skew. Moreover, we measured that $\sigma_h \propto L^{0.237}$, which disagrees with the expected Gaussian scaling of $\sigma_h \propto L^{0.5}$. As a result we may assume that z_i are not independent and identically distributed, but have some dependence on their neighbours.

A numerical test of the normality of the distribution would be to calculate the area under the data collapsed probability distribution for a system with $L \gg 1$, in the range $\langle h\rangle - \sigma_h < h < \langle h\rangle + \sigma_h$, the area should be approximately 0.682.

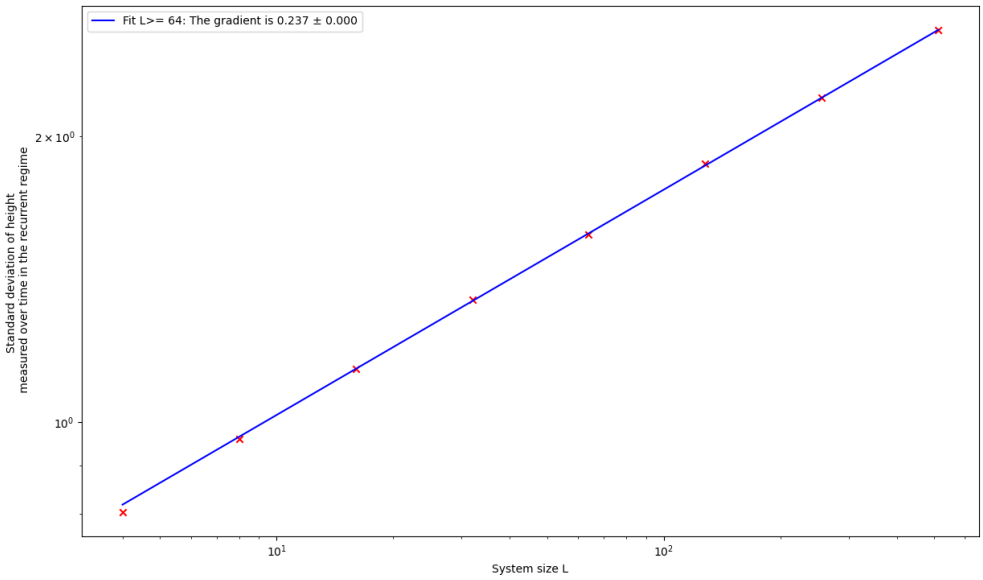

Figure 6: Plot of standard deviation of measured height over time in the recurrent regime against L

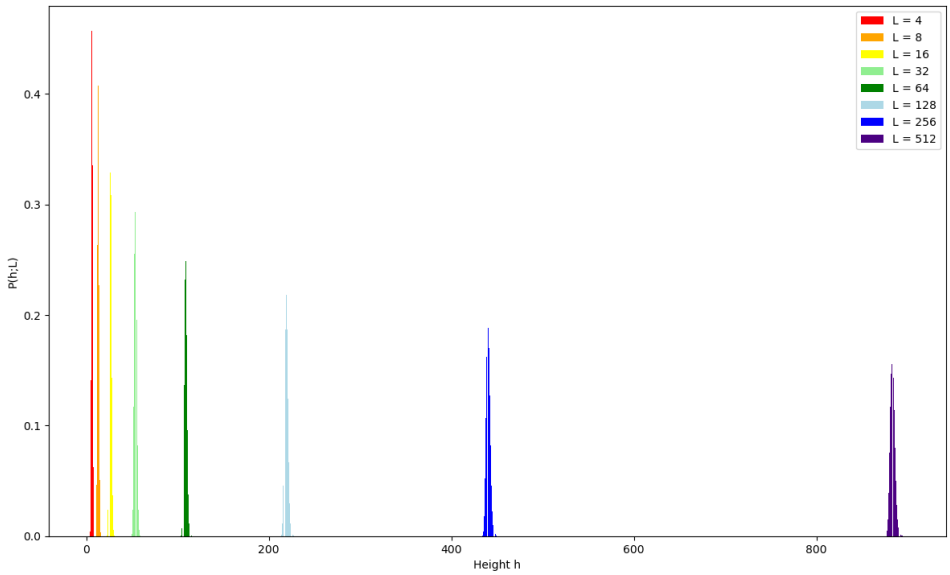

Figure 7: Plot of $P(h; L)$ against h for system sizes L = 4,8,16,32,64,128,256,512.

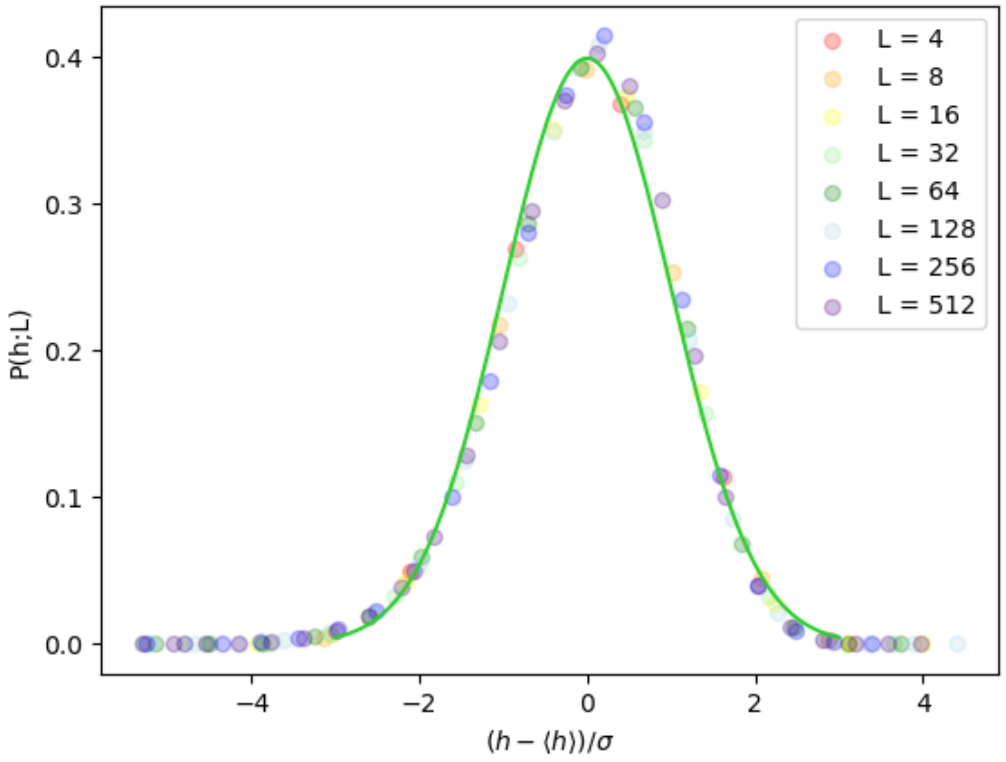

Figure 8: Plot of $\sigma_h P(h;L)$ against $\frac{h-\langle h\rangle}{\sigma_h}$ for system sizes L = 4,8,16,32,64,128,256,512. Normal distribution with $\mu = 0$ and $\sigma = 1$ also plotted.

3 The avalanche-size probability *P(s;L)*

We aimed to estimate the underlying avalanche size probability, $P(s;L)$, in the recurrent regime by sampling a finite series of N avalanches. We measured $P_N(s;L)$ as defined in equation (8) of [5].

3.1 Task 3a

After measurement of $P_N(s;L)$, the data was logarithmically binned in order to reduce scatter and improve the approximation of $P_N(s;L)$ to the underlying probability. Logarithmic binning of the avalanche sizes involves dividing the s-axis into bins that are exponentially increasing in size according to a scale factor j. For each bin, the geometric mean of the bin edges are defined as the bin centres. Then, by noting that avalanche sizes of zero exist, we removed this data bin before plotting $\tilde{P}_N(s;L)$ against s with log-scaled axes, where $\tilde{P}_N(s;L)$ is defined as the avalanche size probability distribution for the log-binned avalanche sizes.

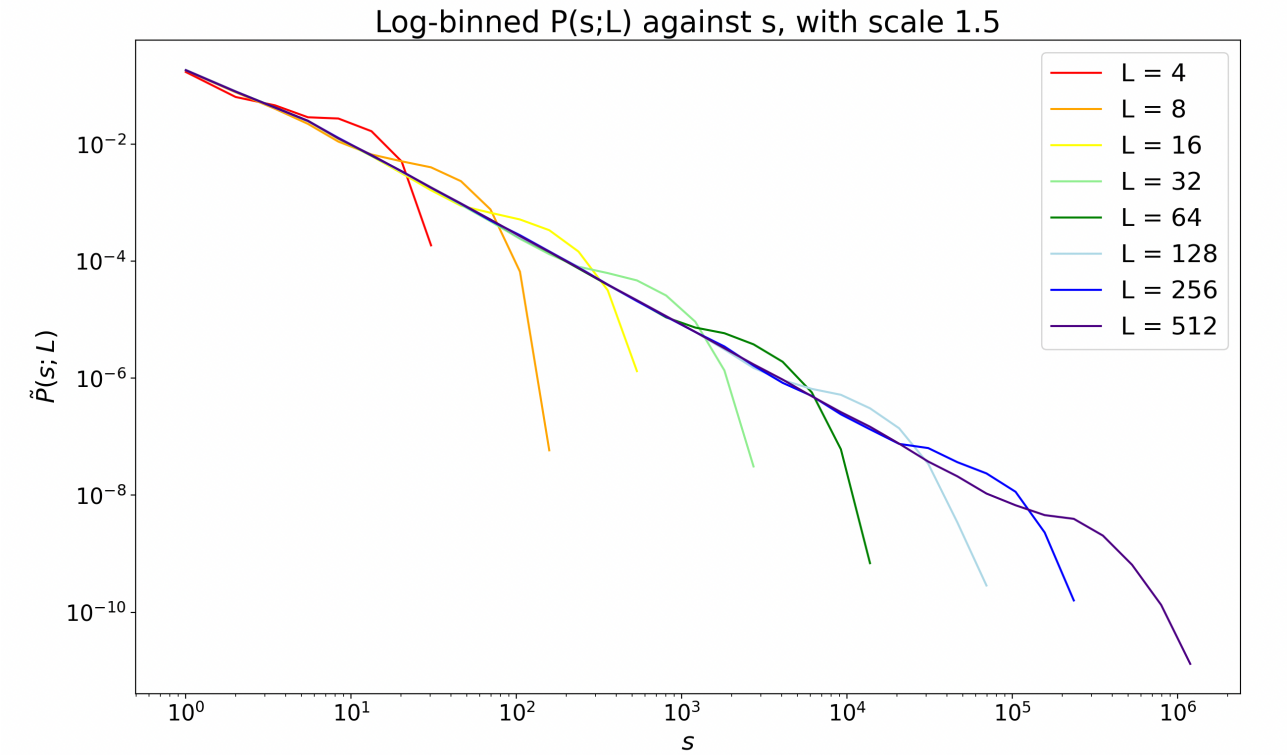

Figure 9: Plot of $P(s;L)$ against s for system sizes L = 4,8,16,32,64,128,256,512.

As $N \to \infty$, $P_N(s;L) \to P(s;L)$. For each system size, the model was run for $t = 4L^2$, and measurement of avalanche sizes began for $t > t_c$. This enabled large enough sample sizes to reveal important behaviour of $P(s;L)$, notably the finite avalanche cut-off size for each system size, s_c and the bump in the linear trend of the log-log plot of $\tilde{P}_N(s;L)$ against s before the avalanche probability rapidly decays, as shown in Figure9. These are observed due the finite size of the systems, otherwise we would expect the linear trend to continue indefinitely.

In order to determine whether the measured $\tilde{P}_N(s;L)$ was consistent with the finite-size scaling ansatz,

$$\tilde{P}_N(s;L) \propto s^{-\tau_s \mathcal{G}(s/L^D)}, \tag{4}$$

for $L \gg 1$ and $s \gg 1$, we attempted to perform a data collapse for the different system sizes. Firstly, we aligned the bumps in the linear trend vertically, by mapping $\tilde{P}_N(s;L) \to s^{\tau_s} P(s;L)$,

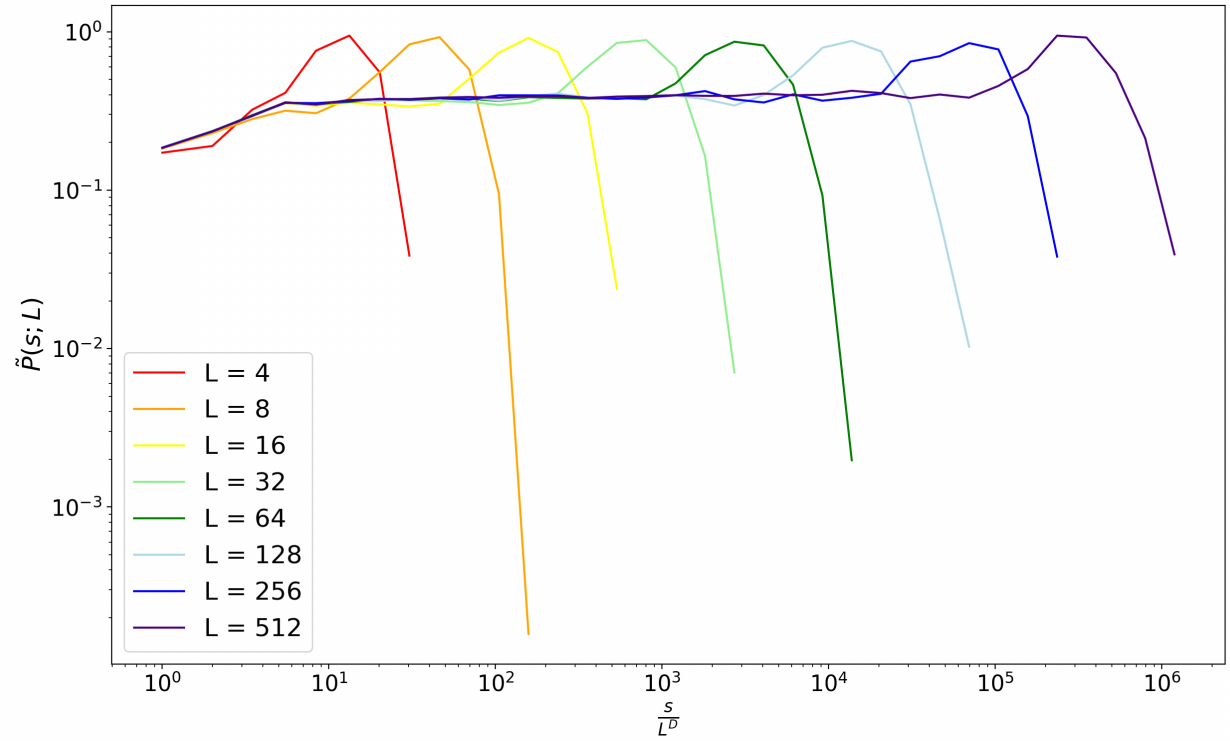

Figure 10: Plot of $s^{\tau_s}P(s;L)$ against s for system sizes L = 4,8,16,32,64,128,256,512.

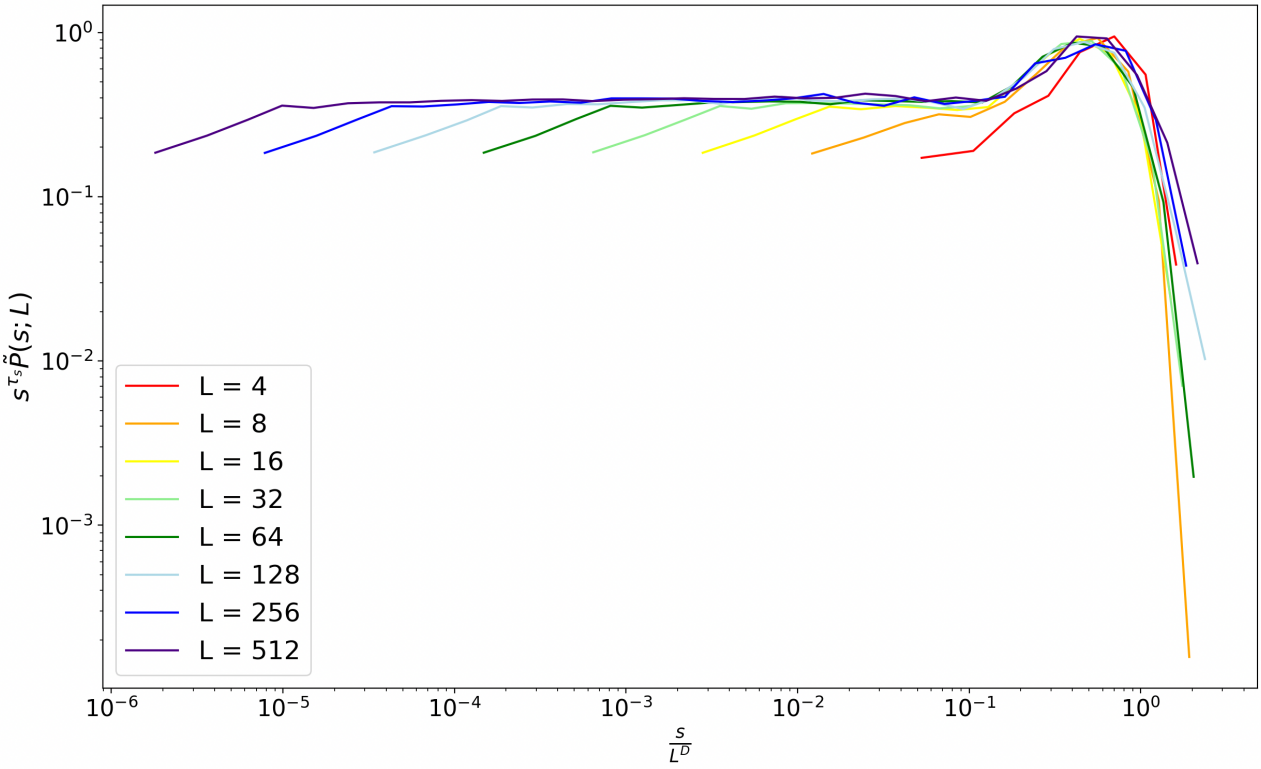

Figure 11: Plot of $s^{\tau_s}P(s;L)$ against s/L^D for system sizes L = 4,8,16,32,64,128,256,512.

and choosing the value of τ_s that best aligns these features, which was $\tau_s = 1.56 \pm 0.01$. The cut-off avalanche sizes of the resulting plot were then aligned horizontally by mapping $s \to s/L^D$. The value of D that resulted in the best data collapse was $D = 2.12 \pm 0.05$. The errors in

D and τ_s were determined by visually observing the effectiveness of the data collapses, and therefore reflect human error. The resulting data collapse, shown in Figure11, demonstrates the agreement of $\tilde{P}_N(s;L)$ with the finite-size scaling ansatz. Note that the data collapse is poorer for systems of lower L, due to the larger impact of corrections to scaling for these systems.

3.2 Task 3b

The k^{th} moment of measured avalanche sizes is defined in equation (10) in [5]. Given the finite-size scaling relation for avalanche size probability expressed in (4), we find $\langle s^k \rangle \propto L^{D(1+k-\tau_s)}$. The scaling of $\langle s^k \rangle$ with L is shown in Figure12.

To demonstrate that our data was consistent with the expected scaling relation, we first measured the gradient of the log-scaled $\langle s^k \rangle$ against L plot for each k using a least-squares fit for $L \geq 64$, this plot is shown in Figure12. We then compared the measured gradients to the expected gradients given by $D(1+k-\tau_s)$ with D and τ_s as determined from the data collapse in subsection 3.1. The percentage difference between the measured and expected gradients are shown in Figure13 as a function of k. We can see that the percentage difference between these gradients is very small for all k, showing agreement between our data and the expected scaling relation.

Using the moment scaling analysis method we determined new values of D and τ_s. We plotted the measured gradient of log-scaled $\langle s^k \rangle$ against L plot, against k. This plot is shown in Figure14 The measured gradient is labelled $D(1+k-\tau_s)$ here to reflect our expectation for the scaling relationship of $\langle s^k \rangle$. The plot exhibits the expected linear fit from which D was determined to be 2.19, and τ_s was determined to be 1.58. These values are similar to those determined from data collapsing the avalanche size probability distribution in subsection 3.1.

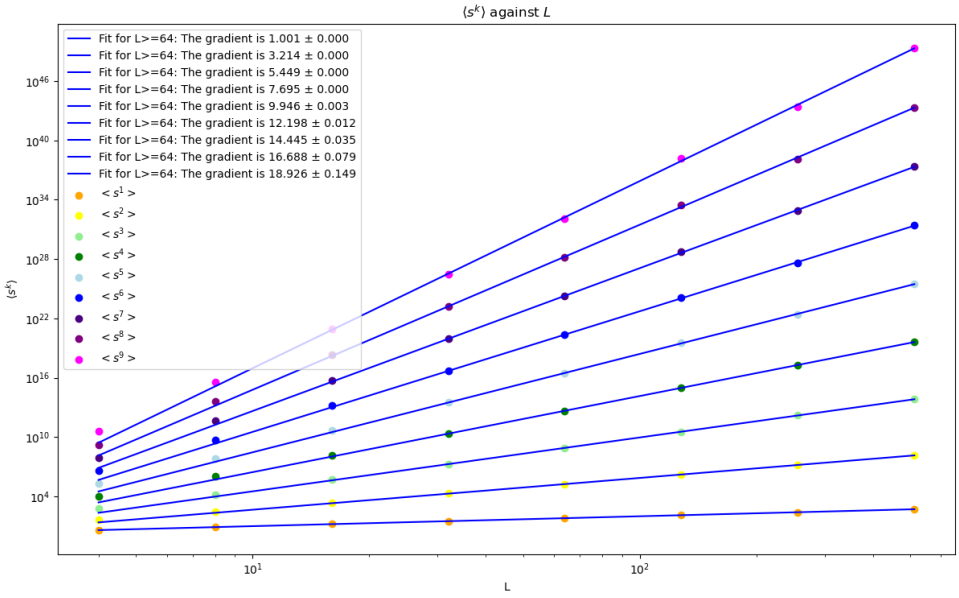

Figure 12: Plot of $\langle s^k \rangle$ against L for system sizes L = 4,8,16,32,64,128,256,512.

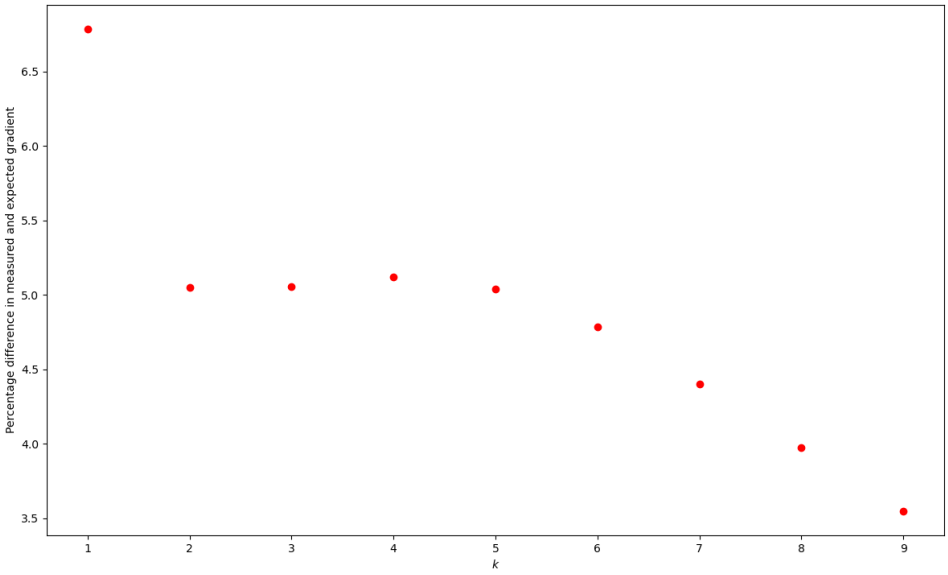

Figure 13: Plot of difference between measured gradient and expected gradient against k. Percentage difference calculated as calculated as $100 \times \frac{grad_{measured} - D(1+k-\tau_s)}{grad_{measured}}$

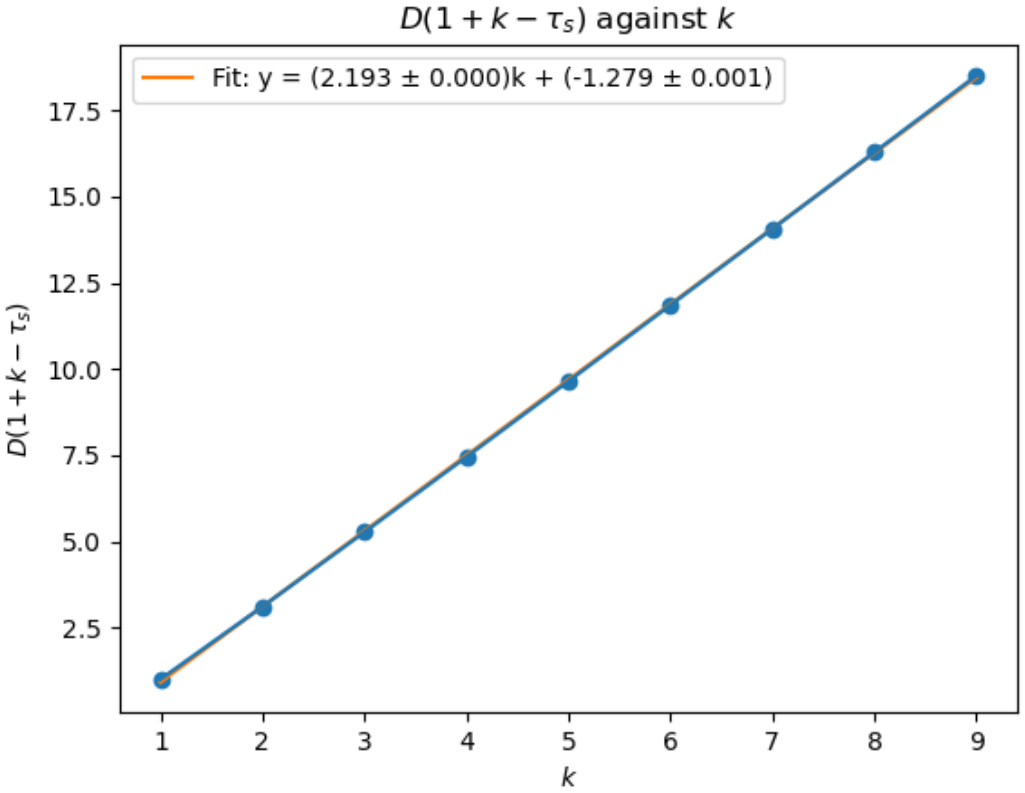

Figure 14: Plot of $D(1+k-\tau_s)$ against L for system sizes L = 4,8,16,32,64,128,256,512.

4 Conclusion

In conclusion, we implemented the Oslo model for system sizes of L = 4,8,16,32,64,128,256,512 and investigated the probability distributions of pile height and avalanche size. We explored

corrections to scaling for finite system sizes, and found critical exponents for ansatz through data collapsing methods and moment analysis techniques.

5 References and Acknowledgements

References

[1] Golyk, Vladyslav A. "Self-organized criticality." (2012).

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

[3] Christensen, Kim, Álvaro Corral, Vidar Frette, Jens Feder, and Torstein Jøssang. "Tracer dispersion in a self-organized critical system." Physical review letters 77, no. 1 (1996): 107.

[4] Kalinin, Nikita, Aldo Guzmán-Sáenz, Yulieth Prieto, Mikhail Shkolnikov, Vera Kalinina, and Ernesto Lupercio. "Self-organized criticality and pattern emergence through the lens of tropical geometry." Proceedings of the National Academy of Sciences 115, no. 35 (2018): E8135-E8142.

[5] Prof. Kim Christensen. "Complexity Project Notes" Condensed Matter Theory Group & Centre for Complexity Science Imperial College London.

[6] Lam, Siu Kwan, Antoine Pitrou, and Stanley Seibert. "Numba: A llvm-based python jit compiler." In Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC, pp. 1-6. 2015.