Complexity Project Report

CID: 00923604, Imperial College London

February 19, 2018

Word Count: 2469

Abstract

In this report, the Oslo model is used to investigate the emergence of scale free invariance in self-organised, critical systems. The system height around the crossover time t_c is shown to scale across system sizes through a data collapse, which is motivated by theoretical arguments. Corrections to scaling are explored and shown to be present in the numerical data for the average steady state height $\langle h \rangle$. Probability distributions for the steady state height and avalanche sizes are examined and also shown to collapse to scaling functions, with corrections to scaling for smaller systems. In the case of the avalanche size probability distribution, the critical exponents for the scaling function are estimated using both a data collapse and moment analysis methods which are shown to give similar results.

1 Introduction & Aims

Complexity science is the study of the emergence of scale free behaviour. Scale free behaviour in systems is characterised by a power law distribution of response sizes, whereby the distribution is identical at different sizes up to a scaling factor. In other words, the system has no characteristic time or length scale. Scale free behaviour is interesting as it is frequently observed in natural systems, and finding mechanisms that generate this behaviour may provide insight into how nature operates. Examples of scale free behaviour include the distribution of earthquake sizes, shown to follow a power law by Gutenberg & Richter (1944), and the occurrence of the characteristic $1/f$ 'flicker noise' in a wide variety of observed power spectra (Bak et al. 1988).

A property that is known to give rise to scale invariance is self-organised criticality (Christensen & Moloney 2005), displayed by systems that settle into a highly susceptible steady state without the fine-tuning of external parameters. This project implements and analyses a numerical version of the Oslo model first presented by Christensen et al. (1996), a system which exhibits self-organised criticality. The Oslo model describes a 1 dimensional rice pile system, slowly driven by the addition of single rice grains to the pile. The system relaxes periodically when the gradient at a site exceeds some threshold, redistributing the grains throughout the system. After some time in a transient phase, the system 'self-organises' into a steady state where the system heights cycle through a set of recurrent states. During this steady state, the relaxations that are observed span several orders of magnitude. The system is said to be in a 'critical' state as the same slow driving force can result in both large and small relaxations, hence the property of 'self-organised criticality'. This project investigates the scaling properties of the Oslo model, in particular relating to the steady-state height of the pile and the distribution of avalanche sizes.

2 Oslo Model Algorithm

The Oslo model consists of a 1 dimensional rice pile with L sites, open at one end, containing h_i grains of rice in each site, where $h_{L+1} = 0$. Each site can be characterised by two values: the gradient at that point $z_i = h_i - h_{i+1}$ and the threshold gradient z_i^{th}. The system is initialised by setting $z_i = 0$, $\forall i$ and initialising the threshold gradients with a probability p, through

$$z_i^{th} = \begin{cases} 1, & \text{with probability } p \\ 2, & \text{with probability } 1-p, \end{cases} \tag{1}$$

The system is slowly driven by adding a single grain of rice to the first site in the system. If the gradient at a site exceeds the threshold gradient, the site topples and $h_i \to h_i - 1$,

$h_{i+1} \rightarrow h_{i+1}+1$, and z_i^{th} is reset for the toppled site according to equation (1). The algorithm is presented in pseudocode for a system of size L run for N iterations, in terms of the site gradients, in Algorithm 1.

Algorithm 1 Oslo Model

procedure $\text{OSLO}(L, N)$ ▷ System of size L run for N iterations.
  $z_i \leftarrow 0$ ▷ Initialise the gradients to 0.
  $z_i^{th} \leftarrow \{1, 2\}$ ▷ Initialise the threshold values.
  $j = 0$
  while $j < N$ **do** ▷ Iterate N times.
    $z_1 \leftarrow z_1 + 1$ ▷ Drive the system.
    while $z_i > z_i^{th}$ **do** ▷ Relax any sites above threshold.
      if i=1 **then**
        $z_1 \leftarrow z_1 - 2$
        $z_2 \leftarrow z_2 + 1$
      end if
      if i=2,3,..L-1 **then**
        $z_i \leftarrow z_i - 2$
        $z_{i\pm1} \leftarrow z_{i\pm1} + 1$
      end if
      if i=L **then**
        $z_L \leftarrow z_L - 1$
        $z_{L-1} \leftarrow z_{L-1} + 1$
      end if
      $z_i^{th} \leftarrow \{1, 2\}$ ▷ Initialise the threshold values of relaxed sites.
    end while
    $j \leftarrow j + 1$
  end while
end procedure

3 Results

3.1 Scaling behaviour of system height

As an initial test for the implementation of the Oslo model algorithm, particularly simple values of $p = 0$ and $p = 1$ are selected, resulting in observed gradients of 1 and 2 for every site, respectively. The system with $p = 0.5$ was tested by comparing the average height of site 1 obtained in the steady state to the results quoted in the literature (Christensen, K. 2017), which were found to be in close agreement. The system with $p = 0.5$ was also tested by running the code for a small system size, printing out the threshold gradients and system heights, and ensuring that the sites topple or build up as expected by the algorithm.

Having established the correct implementation of the model, the variation of system height

$h(t, L)$ with time (defined as the height of site 1, $h_1(t, L)$) is now investigated. In order to smooth out the frequent fluctuations in system height for a better visual representation, the heights are temporally averaged over a window of size $W = 25$, defining the processed height,

$$\tilde{h}(t, L) = \frac{1}{2W + 1} \sum_{t'=t-W}^{t+W} h(t', L). \tag{2}$$

The behaviour of $\tilde{h}(t, L)$ is plotted as a function of time for system sizes $L = 8, 16, 32, 64, 128, 256$, each run for $N = 10^5$ iterations. Figure 2 clearly shows the transition between the transient and recurrent phases, also known as the crossover time, t_c. This is defined as the time where the first grain leaves the system and average number of grains entering the system equals the average number leaving the system: $\langle \text{influx} \rangle = \langle \text{outflux} \rangle$. A striking feature of figure 2 is that in the transient phase $\tilde{h}(t, L)$ increases identically for all system sizes before reaching a steady state height $\langle h(t, L) \rangle_{steady}$. This can be understood qualitatively by considering that at the start of the transient phase, a system can rapidly increase in height, whereas later on each subsequent site will have to become stable before the overall system height can increase, resulting in a slow down in height increase. The only difference that exists for larger systems is the time taken to reach the crossover time t_c, due to the larger number of sites that must be reached before this point. The relationship between $\langle h(t, L) \rangle_{steady}$ and system size L can be derived theoretically. In the steady state, the system can be characterised by an average gradient $\langle z \rangle$, as illustrated in figure 1. The site heights h_i can be approximated by the equation $h_i = \langle z \rangle (L - i)$ where i is the site position, implying a relationship of the form $\langle h(t, L) \rangle_{steady} \propto L$ by matching the y intercepts. The dependence of $\langle t_c \rangle$ on system size can again be obtained using this average gradient, by considering that the steady state is first reached when every site up to the end of the system has been filled, each with an average gradient $\langle z \rangle$. Finding the average crossover time is then reduced to calculating total area occupied by the system, given by $\langle t_c \rangle = \sum_{i-1}^{L} h_i$, shown in figure 1. Using the standard expression for triangular numbers results in the following expression

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

which reduces to $\langle t_c \rangle \propto L^2$ for large system sizes.

This dependence suggests the following scaling form

$$\tilde{h}(t, L) = L \mathcal{F} \left(\frac{t}{L^2} \right) \tag{4}$$

$$\frac{\tilde{h}(t, L)}{L} = \mathcal{F} \left(\frac{t}{L^2} \right), \tag{5}$$

where $\mathcal{F}$ is a scaling function that describes the general behaviour displayed by all systems. This scaling behaviour allows for a data collapse, which is produced in figure 3. In complexity

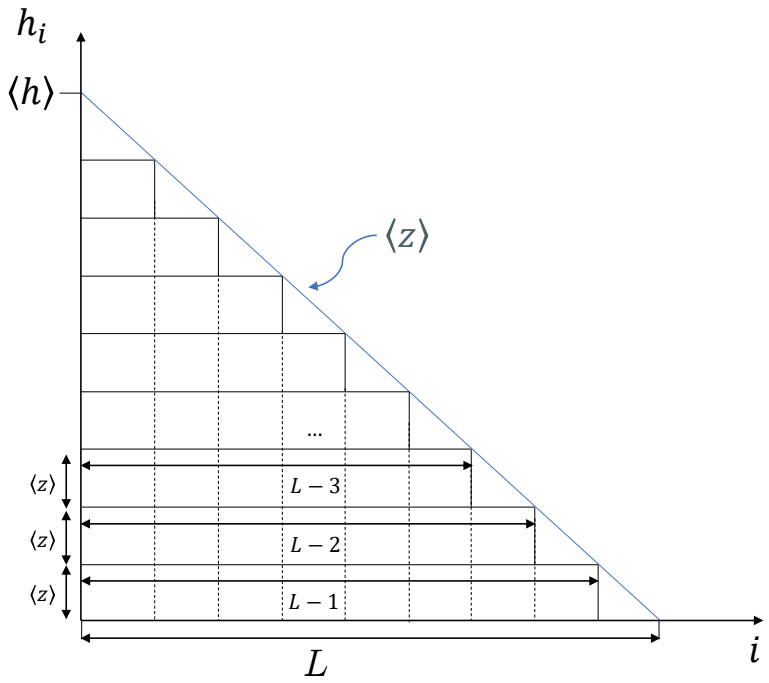


Figure 1: Diagram of the Oslo model in an idealised recurrent phase with every site i at average gradient $\langle z \rangle$, illustrating the origin of the $\langle h \rangle_{steady} \propto L$, $\langle t_c \rangle \propto L^2$ relations.

science, a data collapse is an important indication that different systems are governed by the same underlying physical mechanism (see Bak et al. (2002) for an application in the context of unifying the causes of earthquakes and aftershocks). In figure 3a only the scaling for $\tilde{h}(t, L)$ is applied, which scales $\langle h \rangle_{steady}$ to the same value for all system sizes. In figure 3b the scaling is applied to both axes, which also shifts the value of t_c to the same point for all values of L. It can be seen in figure 3b that larger systems collapse onto the scaling function $\mathcal{F}$, whereas smaller systems deviate from this function. This suggests that the general trend that $\mathcal{F}$ describes only applies in the limit $L \gg 1$. This is the first indication of 'corrections to scaling' for small system sizes, a notion which appears often throughout this report. Plotting the transient region of figure 3b on a log-log plot, the increase of the system height can be shown to obey $\tilde{h}(x) \propto x^{1/2}$ where $x = t/L^2$ is the scaled time variable, by plotting this function alongside the data.

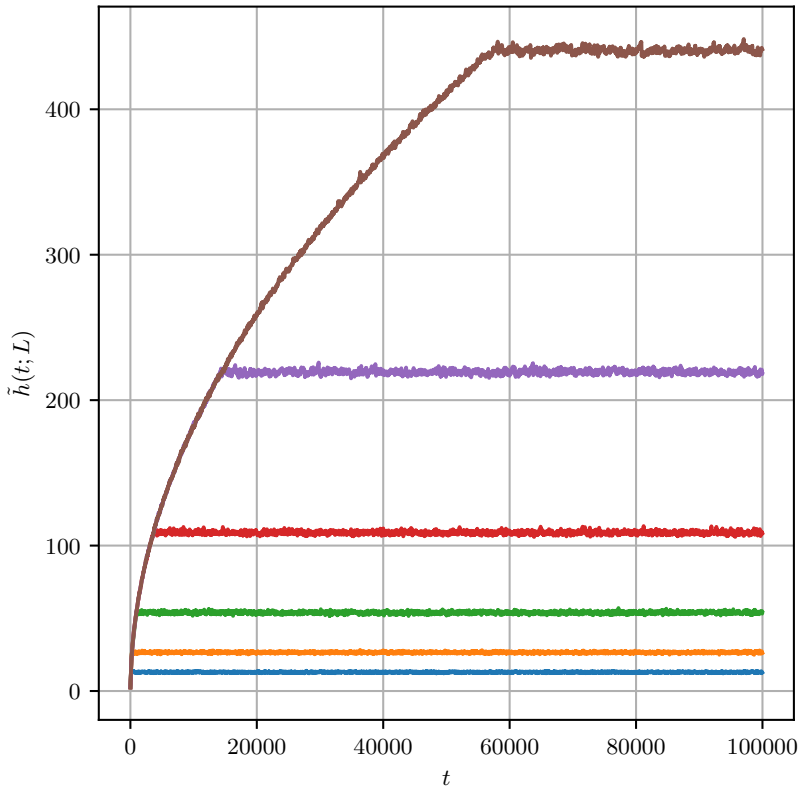
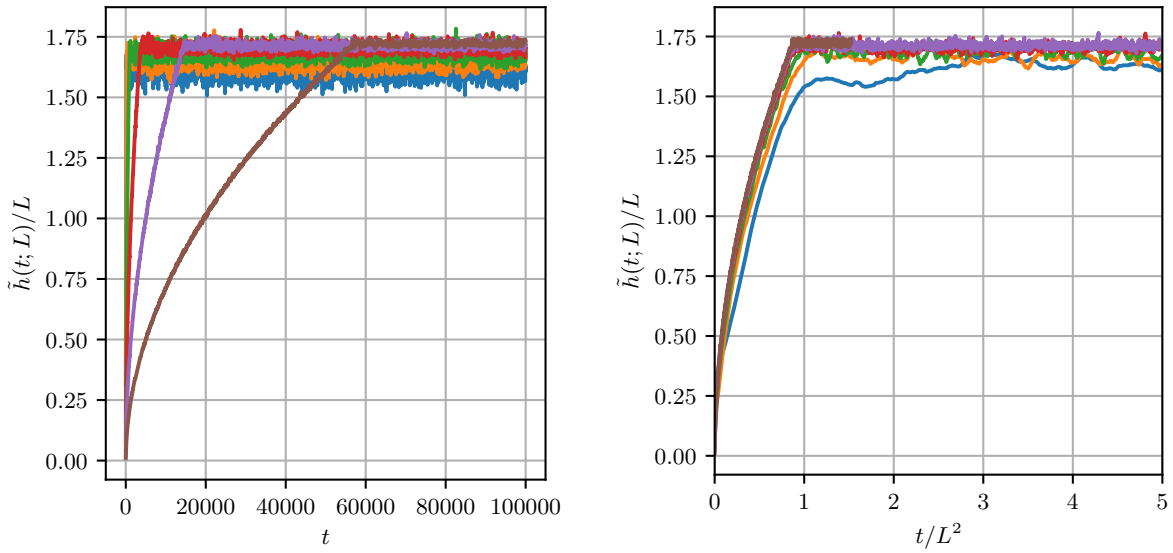


Figure 2: Temporally smoothed heights $\tilde{h}(t, L)$ for system sizes $L = 8$ (blue), 16 (orange), 32 (green), 64 (red), 128 (purple), 256 (brown), each run for $N = 10^5$ iterations, clearly displaying the transition between transient and recurrent phases.



(a) Data collapse applied to the height of the pile through $\tilde{h}(t, L)/L$.

(b) Data collapse applied to both the height of the pile, $\tilde{h}(t, L)/L$ and the time scale, t/L^2.

Figure 3: Data collapse of the temporally smoothed heights $\tilde{h}(t, L)$ for system sizes $L = 8$ (blue), 16 (orange), 32 (green), 64 (red), 128 (purple), 256 (brown).

To examine the cross-over time more closely, the theoretical expression given by equation (3) is compared to the observed value given by numerical simulations. In figure 4, plotting the theoretical and observed crossover times, averaged over 10 runs, shows that the theoretical expression matches the experimental value to within the standard deviation measured. In particular it can be seen that the standard deviation on larger system sizes increases, which would be expected due to the greater number of recurrent states for larger systems, as demonstrated by the greater fluctuations in system height observed for larger systems in figure 2.

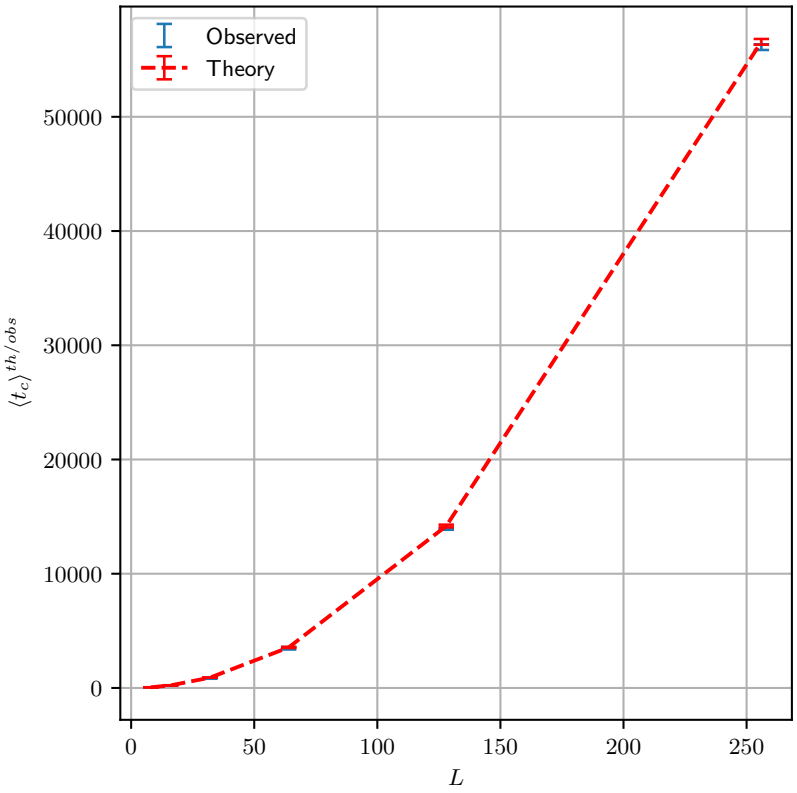


Figure 4: Mean theoretical (red dashed line) and observed (blue points) cross-over time $\langle t_c \rangle^{th/obs}$ and standard deviation (error bars), for system sizes $L = 4, 8, 16, 32, 64, 128, 256$, averaged over 10 runs for $N = 10^5$ iterations.

3.2 Corrections to scaling for system height

A theoretical argument was presented in section 3.1 for the average system height in the steady state scaling as $\langle h \rangle_{steady} \propto L$, which was supported by the data collapse produced in figure 3. In this section, the possibility of corrections to scaling will be explored, that is, the possibility that the true scaling behaviour of $\langle h \rangle_{steady}$ diverges from this simple form. For this purpose, the average steady state height is defined as

$$\langle h(L) \rangle = \lim_{T \to \infty} \frac{1}{T} \sum_{t=t_c+1}^{t=t_c+T} h(t; L), \tag{6}$$

where the subscript is dropped for convenience and t_c is the cross-over time. To explore corrections to scaling, the following form of $\langle h(L) \rangle$ is assumed,

$$\langle h(L) \rangle = a_0 L(1 - a_1 L^{-\omega_1}), \tag{7}$$

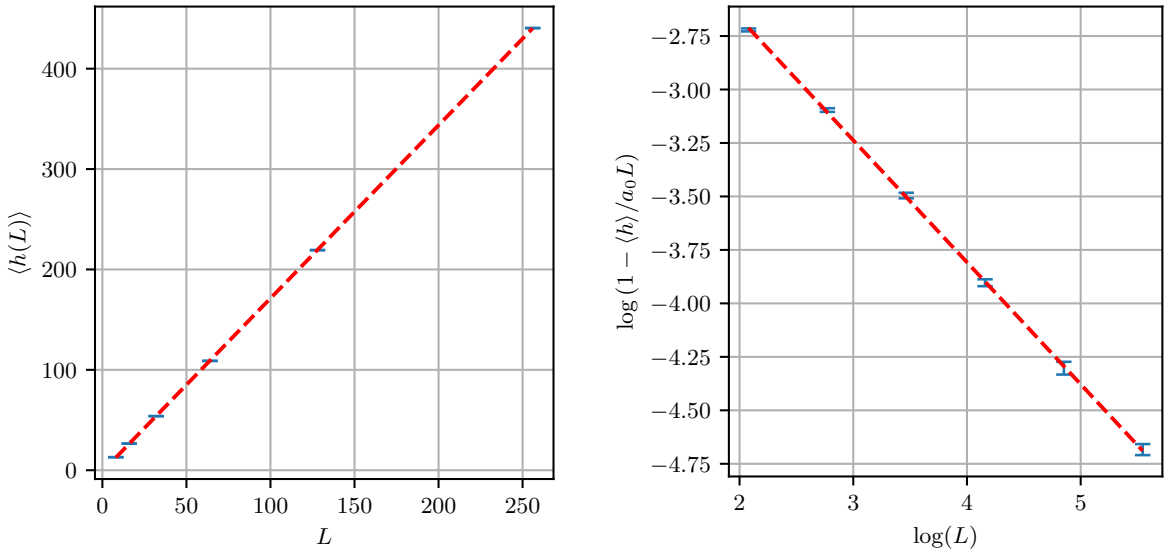
where $\omega_1 > 0, a_i$ are constants. In order to extract the parameter values, this expression can be manipulated into the following form,

$$\log\left(1 - \frac{\langle h \rangle}{a_0 L}\right) = -\omega_1(\log(a_1) + \log(L)). \tag{8}$$

Although the ω_1, a_0 cannot be directly obtained by plotting this expression, the LHS and the second term on the RHS can be plotted against each other, for different values of a_0. An initial guess for a_0 can be found by assuming no corrections to scaling and performing a linear regression for $\langle h(L) \rangle$ plotted as a function of L, resulting in an average gradient implying $a_0 \approx 1.72$, as shown in figure 5a. Then, a linear regression is performed for equation (8) for a range of values around this estimate, and each value of a_0 is assessed by the R-value obtained in each case, describing the goodness of fit of the linear regression (Barlow 1989). For each run, the value of a_0 corresponding to the minimum R-value or most linear fit, is selected. Using this method averaged over 10 runs, parameter values of $a_0 = 1.7366 \pm 0.0007$, $\omega_1 = 0.5704 \pm 0.0067$ are obtained where the uncertainty is the standard deviation on the parameter. This shows that corrections to scaling are indeed significant, at least in the average steady state height. Having examined the scaling form of the average system height, it is natural to investigate the L dependence of the standard deviation of the system height, $\sigma_{h,L}$. Performing a linear regression on a log-log plot of these variables reveals a relationship of the form $\sigma_{h,L} \propto L^{\gamma}$, where $\gamma = 0.240 \pm 0.01$, also averaged over 10 runs.

3.3 System height probability

In this section, the scaling properties of the system height probability distribution $P(h; L)$ are investigated. The expected form of the system height probability distribution can be derived by considering the system height as a sum over the site gradients, $h = \sum_{i=1}^{L} z_i$. If it is now assumed that the site gradients are independent, identically distributed, random variables with a finite variance, the central limit theorem states that the observed system heights should follow a Gaussian distribution, for $L \gg 1$. Figure 6a shows the height probability distributions $P(h; L)$ obtained for system sizes $L = 8, 16, 32, 64, 128, 256$, showing the expected Gaussian form and normalised to unity. In particular, larger system sizes have a greater average steady state height $\langle h(L) \rangle$ and a greater spread of heights resulting in a larger $\sigma_{h,L}$. The increased spread of values obtained for larger system sizes would be expected as larger systems have a greater number of recurrent states available in the steady state. The form of $P(h; L)$ suggests



(a) Average linear regression over 10 runs (red dashed line) of $\langle h(L) \rangle$ data points (blue points, standard deviation shown as error bars), in order to obtain a first estimate of a_0.

(b) Best fitting linear regression of equation (8) (red dashed line) averaged over 10 runs and data points corresponding to this best fit (blue points, standard deviation shown as error bars).

Figure 5: Estimating the correction to scaling parameters a_0 and ω_1.

the following scaling dependence

$$P(h; L) = \frac{1}{\sigma_h} \mathcal{G}_h \left(\frac{(h - \langle h \rangle_L)}{\sigma_{h,L}} \right) \tag{9}$$

$$\mathcal{G}_h(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \tag{10}$$

where the scaling function $\mathcal{G}_h(x)$ is a unit Gaussian. This allows for a data collapse as rearranging equation (9) results in

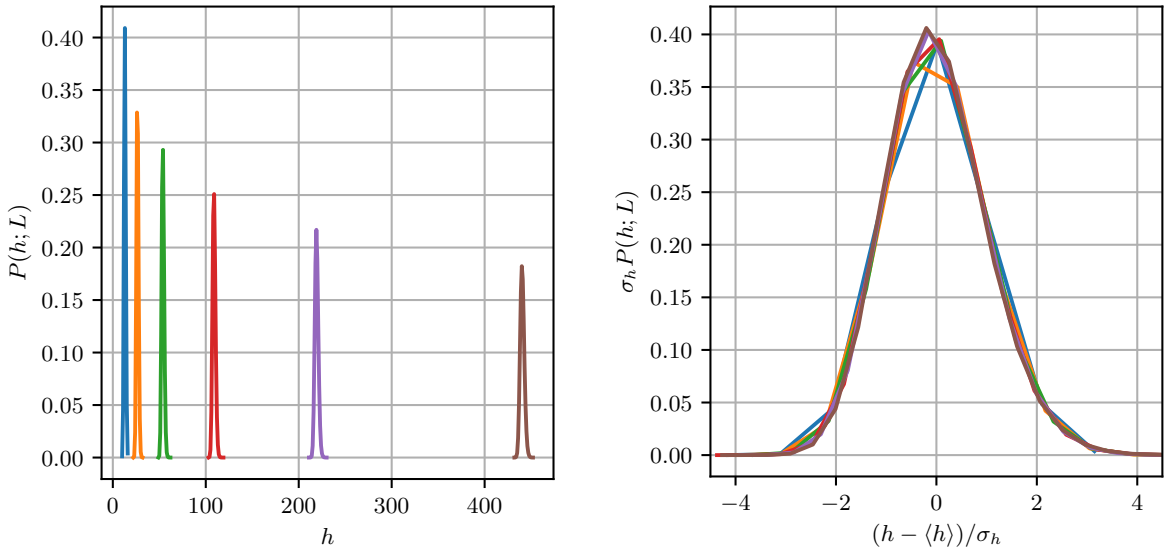
$$\sigma_h P(h; L) = \mathcal{G}_h \left(\frac{(h - \langle h \rangle_L)}{\sigma_{h,L}} \right), \tag{11}$$

which is plotted in figure 6b for the relevant system sizes. The data collapse quality can be seen to be good, with larger system sizes being nearly indistinguishable while the smaller systems display corrections to scaling as expected.

3.4 Avalanche size probability and moments

This section focuses on measuring the avalanche size probability and associated moments. The normalised avalanche-size probability is defined as

$$P_N(s; L) = \frac{\text{No. of avalanches of size } s \text{ in a system of size } L}{\text{Total no. of avalanches } N}, \tag{12}$$



(a) Steady state height probability distribution $P(h;L)$.

(b) Data collapse of steady state height probability.

Figure 6: Data collapse of the steady state height probability distribution, onto a unit Gaussian distribution, for system sizes $L = 8$ (blue), 16 (orange), 32 (green), 64 (red), 128 (purple), 256 (brown).

where measurements are made after the cross-over time t_c. Due to the finite statistics used to sample the underlying avalanche size probability distribution, a logarithmic binning technique introduced in Christensen & Moloney (2005) is applied, whereby the bin size increases for larger avalanches, in order to reduce the noise present in large avalanche sizes. In figure 7, $P(s;L)$ is plotted on a log-log scale for system sizes $L = 8, 16, 32, 64, 128, 256, 512, 1024$, for $N = 10^7$. It is seen that the distribution follows a power law for large avalanche sizes, before reaching a cut-off point related to the system size L. Just before the cut-off point, a bump is observed which is related to system-spanning avalanches. At small avalanche sizes, a characteristic zig-zag pattern is also observed in all system sizes, which can be related to odd and even avalanche sizes through exact calculation of the Oslo model occupation probabilities (Corral 2004).

The procedure followed to carry out the data collapse of the avalanche size probability distribution follows that used in Christensen & Moloney (2005). The cut-off avalanche size s_c would physically be expected to be related to the system size, and so a scaling relation of the form $s_c \propto L^D$ is introduced. Furthermore, $P_N(s;L)$ follows a power law distribution to a good approximation for large avalanche sizes prior to cutoff which can be expressed as $P_N(s;L) \propto s^{-\tau_s}$. This motivates the following form of the finite-size scaling ansatz:

$$P_N(s;L) \propto s^{-\tau_s} \mathcal{G}(s/L^D), \tag{13}$$

where τ_s and D are known as 'critical exponents', commonly encountered in statistical physics as universal quantities that describe a system near critical phase transition points. Performing a data collapse based on this ansatz allows the critical exponents D and τ_s to be estimated. Assessed by eye, the best results for data collapse occurred for exponent values of $D = 2.25$ and $\tau_s = 1.55$, shown in figure 8.

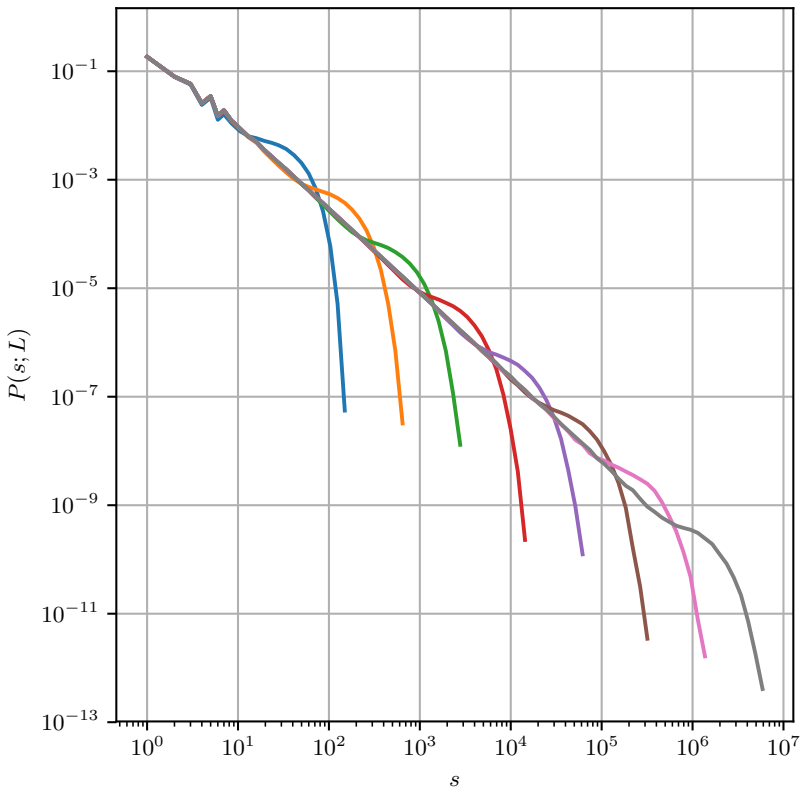
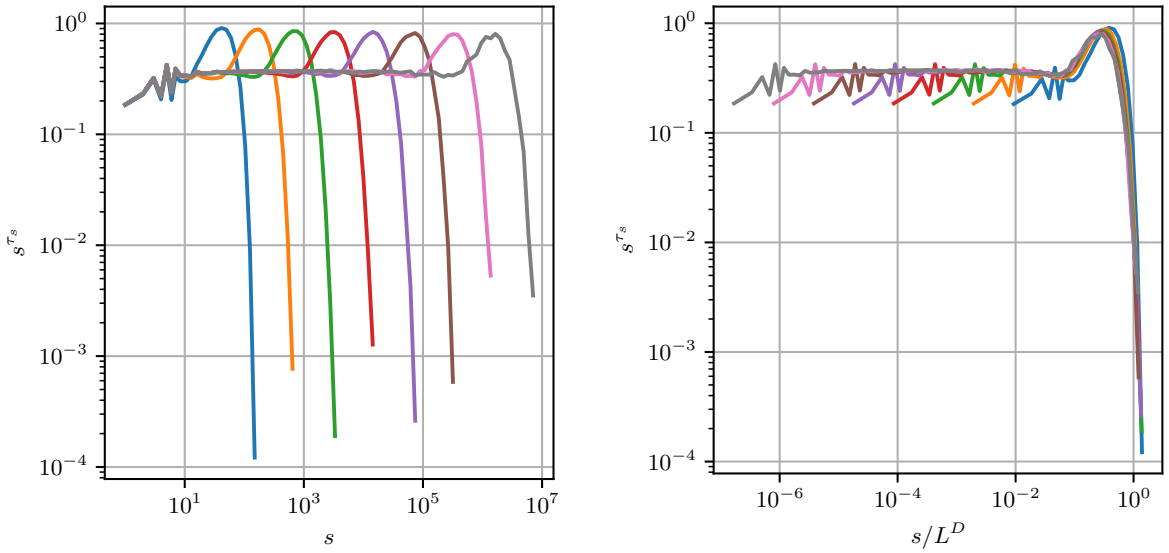


Figure 7: Log binned avalanche size probability distribution for $N = 10^7$ iterations, for system sizes $L = 8$ (blue), 16 (orange), 32 (green), 64 (red), 128 (purple), 256 (brown), 512 (pink), 1024 (grey).

The critical exponents can also be measured directly through moment analysis. The kth moment of the avalanche size distribution is measured as

$$\langle s^k \rangle = \lim_{T \to \infty} \frac{1}{T} \sum_{t_c+1}^{t_c+T} s^k, \tag{14}$$

where the avalanche sizes s are only measured during the recurrent phase, that is, after t_c. Theoretically, the kth moment can also be derived through the underlying probability distribution given by equation 13



(a) Transformed avalanche size probability $s^{\tau_s}P(h;L)$ plotted against actual avalanche size s.

(b) Transformed avalanche size probability $s^{\tau_s}P(h;L)$ plotted against rescaled avalanche size s/L^D.

Figure 8: Data collapse of the log-binned avalanche size probability distribution for system sizes $L = 8$ (blue), 16 (orange), 32 (green), 64 (red), 128 (purple), 256 (brown), 512 (pink), 1024 (grey), for $N = 10^7$, using critical exponents $D = 2.25$, $\tau_s = 1.55$.

$$\langle s^k \rangle = \sum_{s=1}^{\infty} s^k P(s;L) \tag{15}$$

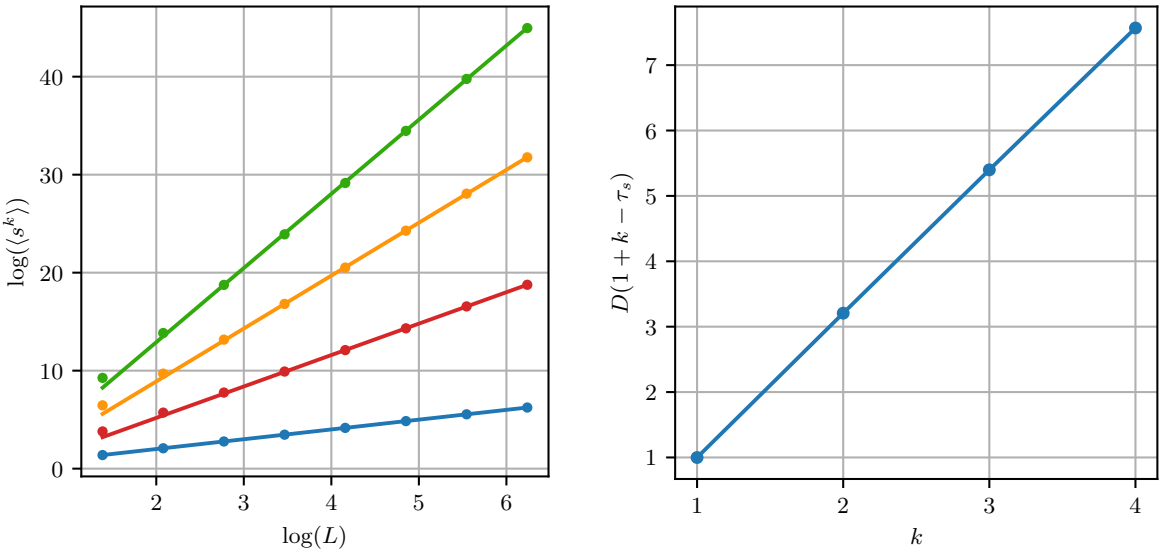
$$= \sum_{s=1}^{\infty} s^{k-\tau_s} \mathcal{G}(s/L^D). \tag{16}$$

Approximating the summation by an integral, it can be shown (Christensen & Moloney 2005) that the kth moment satisfies the following scaling relation

$$\langle s^k \rangle \propto L^{D(1+k-\tau_s)}, \tag{17}$$

for $L \gg 1$, $k \geq 1$. Hence, plotting the kth moment measured through equation (14) against system size L should result in a straight line with gradient $D(1 + k - \tau_s)$. However, it is important to take into account corrections to scaling by estimating the gradient using moments calculated for system sizes satisfying $L \gg 1$. In figure 9a, the gradients are calculated through a linear regression of the moments corresponding to the three largest system sizes ($L = 128, 256, 512$). Plotting this line of best fit in figure 9a reveals the corrections to scaling for smaller system sizes as they do not follow this trend. In figure 9b $d\langle \log(s^k) \rangle / d\log(L)$ is plotted as a function of k, allowing the values of the critical exponents D and τ_s to be extracted through the gradient and intercept of the resulting linear fit. This procedure shown in figure

9, resulting in critical exponent estimates of $D = 2.254 \pm 0.024$ and $\tau_s = 1.559 \pm 0.006$. Clearly both of these values are in good agreement with the estimates made using the data collapse.



(a) Measured kth moment of the avalanche size probability distribution as a function of L, plotted alongside $d\langle \log(s^k) \rangle / d\log(L)$ estimated from the 3 largest systems.

(b) $D(1 + k - \tau_s)$ estimated from the values of $d\langle \log(s^k) \rangle / d\log(L)$, as a function of k for $k = 1, 2, 3, 4$, allowing D, τ_s to be estimated from the gradient and intercept.

Figure 9: Moment analysis used to estimate the critical exponents D, τ_s, applied to system sizes $L = 4, 8, 16, 32, 64, 128, 256, 512$, for $N = 10^6$ iterations.

4 Conclusion

In this report, the concepts of scale invariance and self organised criticality were explored using the Oslo model. In section 3.1, the scaling behaviour of the system height in the transient phase and at the start of the steady state phase was investigated, initially through a theoretical argument which was subsequently shown to fit the numerical results through a data collapse. Here, the concept of 'corrections to scaling' was encountered where smaller system sizes deviated from the expected scaling behaviour. This concept was further explored in section 3.2, where a theoretical expression taking into account corrections to scaling for $\langle h \rangle$ was introduced. The numerical data was shown to be consistent with this form, confirming that corrections to scaling are significant for the system sizes considered. In section 3.3, the scaling behaviour of the probability distributions of system heights was investigated. The probability distributions were shown to collapse onto a unit Gaussian as expected due to the underlying site gradients being identically distributed, independent random variables. Here, corrections to scaling were again observed in the distributions. Finally, in section 3.4 the

scaling properties of the probability distributions of avalanche sizes were examined. It was observed that the distributions collapse onto the finite size scaling form given by Christensen & Moloney (2005), which allows the estimation of the critical exponents D and τ_s. These exponents were also estimated through moment analysis, where the theoretical form of the underlying avalanche size probability distribution was used to relate the kth moment to the critical exponents, resulting in similar values for D and τ_s. Once again, evidence for corrections to scaling was observed in the moment analysis.

References

Bak, P., Christensen, K., Danon, L. & Scanlon, T. (2002), 'Unified scaling law for earthquakes', *Phys. Rev. Lett.* **88**, 178501.
URL: *https://link.aps.org/doi/10.1103/PhysRevLett.88.178501*

Bak, P., Tang, C. & Wiesenfeld, K. (1988), 'Self-organized criticality', *Phys. Rev. A* **38**, 364–374.
URL: *https://link.aps.org/doi/10.1103/PhysRevA.38.364*

Barlow, R. (1989), *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences*, Wiley.

Christensen, K. (2017), 'Complexity project notes'.

Christensen, K., Corral, A., Frette, V., Feder, J. & Jøssang, T. (1996), 'Tracer dispersion in a self-organized critical system', *Phys. Rev. Lett.* **77**, 107–110.
URL: *https://link.aps.org/doi/10.1103/PhysRevLett.77.107*

Christensen, K. & Moloney, N. (2005), *Complexity and Criticality*, Imperial College Press.

Corral, A. (2004), 'Calculation of the transition matrix and of the occupation probabilities for the states of the oslo sandpile model', *Phys. Rev. E* **69**, 026107.
URL: *https://link.aps.org/doi/10.1103/PhysRevE.69.026107*

Gutenberg, B. & Richter, C. F. (1944), 'Frequency of earthquakes in california', *Bulletin of the Seismological Society of America* **34**(4), 185–188.