Neutral Theory and Scale-Free Neural Dynamics

Clara Eminente

Review of the original paper published in 2017 by *Matteo Martinello, Jorge Hidalgo, Serena di Santo, Amos Maritan, Dietmar Plenz, Miguel A. Muñoz*

Quantitative Life Science 2019/2020

Table of contents

GOAL: shed some light on the results obtaind simulating neuronal avalanches using the Millman model.

Introduction

Neuronal avalanches

Human brain is endogenously active i.e. exhibits activity even at rest in the form of **avalanches of neurons activations**

Empirically measured temporal avalanches exhibit **scale-free statistics** both in size (s) and duration (t)

$$P(s) \sim s^{-\tau} \qquad\qquad P(t) \sim t^{-\alpha}$$

Exponents the same of those of the unbiased branching process, $\tau = \frac{3}{2}$, $\alpha = 2$.

$\longrightarrow$ **Critical brain hypothesis**

Critical brain hypothesis

Brain dynamics is poised at the edge of a **critical phase transition**.

Computational advantages: optimal transmission, optimal storage capability, optimal sensibility to signals, etc.

Possible explanation: self-organized criticality (SOC)

Assessing criticality: power laws and critical exponents, crackling noise relation.

Scale-free statistics in absence of criticality:

- Yule processes
- unbiased random walks

Goals

INVESTIGATE:

- alternative scenario for scale-free neural dynamics: **neutral theory**
- robustness of crackling noise relation

Millman model

Millman Model

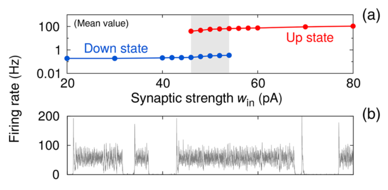
Structure:

- N neurons
- Erdős–Rényi network
- Leaky integrate and fire
- Free parameter: synaptic strength

Avalanches are defined causally (”who triggers who”) and not from temporal proximity

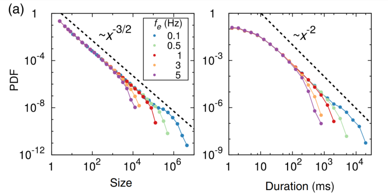
[1]Self-organized criticality occurs in non-conservative neuronal networks during ‘up’ states, *D. Millman, S. Mihalas, A. Kirkwood, E. Niebur*; Nature Physics 6, 801–805(2010)

Results



First result: two distinct phases of activity

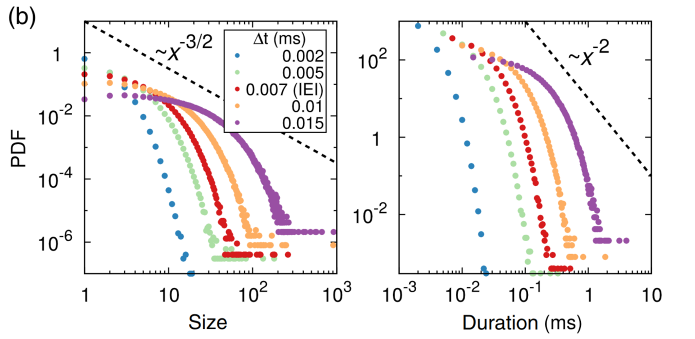
Second result: causal avalanches in the up phase show scale-free statistics



Millman et al. : scale-free behaviour is due to self-organized criticality

Temporal avalanches

Avalanches defined from temporal proximity (as done empirically) do not show typical scale-free behaviour



Empirically found behaviour is not reproduced!

What is the relationship with experimental data?

Authors considerations on Millman et al. Model

Inconsistencies:

- phase transition is not second order
- ”critical” behaviour even far from criticality

Where does scale-free dynamics stem from?

Introduce a minimal model based on neutral theory

Neutral theory and minimal model

Neutral theory

Introduced in molecular evolution by Kimura in 1968.

All cohexisting species (alleles, trees, avalanches..) are **intrinsically equivalent**.

$\longrightarrow$ Changes have **effects purely driven by stochasticity**

$\longrightarrow$ **Marginal propagation**: a specific specie does not have any net tendency to shrink or expand

Successfully used in epidemic outbreaks, microbiome evolution, spreading of neutral memes...

Minimal model

Structure:

- N nodes
- Fully connected network
- A: active node, I: inactive node
- k: label of an avalanche, each A belongs to a k

Dynamics:

- new avalanche starts with rate ϵ
- inactive nodes are activated with rate λ
- active nodes deactivate with rate μ

Notes on the minimal model

$$I \xrightarrow{\epsilon} A_{M(t)+1} \tag{1a}$$

$$I + A_k \xrightarrow{\lambda} A_k + A_k \tag{1b}$$

$$A_k \xrightarrow{\mu} I \tag{1c}$$

- Neutrality: rates do not depend on k
- for $\epsilon \to 0$: normal contact process $\longrightarrow$ critical phase transition for $\lambda_c = \mu$
- differences with a contact process: scale-free behaviour is exhibited in all the ”up” phase ($\lambda > \lambda_c$)

Simulations

Simulations were run by means of the **Gillespie algorithm**.

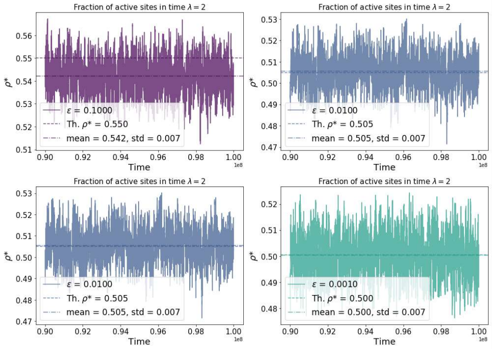
parameter	value
N	10^4
μ	1
λ	$2 = 2\lambda_c$
ϵ	$10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}$

s	total number of neurons activated by an avalanche
t	time elapsed from first activation to extinction
ρ(t)	number of active sites at time t

Results i

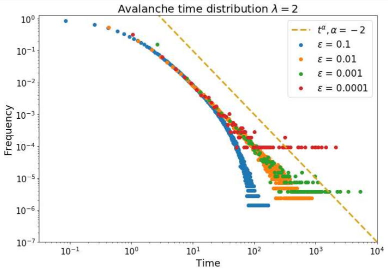
First result: up and down phases appear for different λ

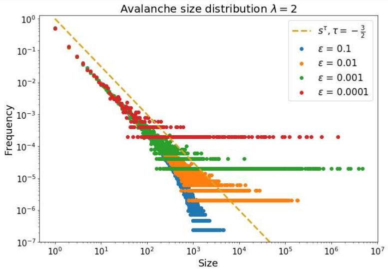
Second result: total number of active sites at stationarity, $\rho*$, is compatible with expected value.



Results ii

Third result: sizes and durations exhibit power law statistics also far from λ_c





Associated Langevin equation:

$$\dot{\rho}_k = A(\rho_k) + \sqrt{B(\rho_k)}\xi_k(t) = \\ = \rho_k\Big[\lambda(1-\rho) - \mu\Big] + \sqrt{\frac{\rho_k[\lambda(1-\rho)+\mu]}{N}}\xi_k(t) \quad (2)$$

In the **down state at stationarity** $\rho = \rho* = \frac{\epsilon}{(\mu-\lambda)}$.

For $\epsilon \to 0$, $\rho* \to 0$

Analytical study ii

In the **up state at stationarity** $\rho = \rho* = 1 - \frac{\mu}{\lambda} + \epsilon\frac{\mu}{\lambda(\lambda-\mu)}$ and for $\epsilon \to 0$ the deterministic term vanishes:

$$\dot{\rho}_k = \sqrt{\rho_k(t)}\xi_k(\hat{t}) \tag{3}$$

Explaining "critical" behaviour: in the up state the dynamics is a random walk (rw) regardless of λ

Explaining exponents: avalanche size and time distributions can be computed by means of **first return to the origin** of a rw.

Crackling noise relation

Power-law statistics and exponents are not sufficient to identify criticality.

Crackling noise relation (CNR): if size and duration distributions **of a critical process** follow:

$$P(s) \sim s^{-\tau} \qquad\qquad P(t) \sim t^{-\alpha} \tag{4}$$

then average size scales with duration as:

$$\langle S \rangle(t) \sim t^{\gamma} \qquad \text{with:} \quad \gamma = \frac{\alpha - 1}{\tau - 1} \tag{5}$$

In our minimal model scale-free statistics is not due to criticality.

Does the crackling noise relation fail?

	$\epsilon = 0.01$	$\epsilon = 0.001$	$\epsilon = 0.0001$	theory
α	1.94 ± 0.01	1.98 ± 0.02	1.95 ± 0.02	2
τ	1.52 ± 0.03	1.49 ± 0.01	1.49 ± 0.01	1.5
γ (from fit)	1.83 ± 0.02	1.93 ± 0.03	1.90 ± 0.03	2
γ (from $\frac{\alpha-1}{\tau-1}$)	1.8 ± 0.1	2.00 ± 0.05	1.92 ± 0.05	2

CNR appears to hold even in absence of criticality.

Conclusions

About the dynamics

Empirical neuronal avalanches defined from **temporal proximity** exhibit scale-free statistics.

Millman et al. proposed a model for **causal avalanches** .

Millman explains these results by means of **SOC** even in **absence of a critical phase transition**.

Scale-free behaviour is a consequence of the intrinsic neutrality of the process.

About the crackling noise relation

Exponents and scale-free behaviour are not a reliable methods to asses criticality.

Usually also CNR is investigated.

Fitting non critical causal avalanches show that CNR is satisfied.

Destexhe and Touboul[2] argued that CNR is highly depending on fitting choices.

Further investigations needed:

- different fitting methods
- more data
- different values of λ

[2] Comment on "Criticality Between Cortical States", *Alain Destexhe and Jonathan D. Touboul*

About the connection with empirical data

Empirical data vs causal avalanches

Still lacking clear connection between **temporal and causal** avalanches.

Benefits of causal information

Some learning mechanisms involve causal information in the activation of neurons (synaptic timing dependent plasticity)

Thank you

Appendix

Appendix i

Millman model for causal avalanches:

$$\dot{V}_i = -\frac{V_i - V_r}{RC} + \sum_k \frac{I_{e_i}^k(t)}{C} + \frac{1}{C} \sum_{\substack{i' \in n.n.(i) \\ j,k}} \Theta[p_r U_{i'j}(t_{s_{i'}}^k) - \zeta_{i'j}^k] I_{\mathrm{in}_{i'}}^k(t),$$

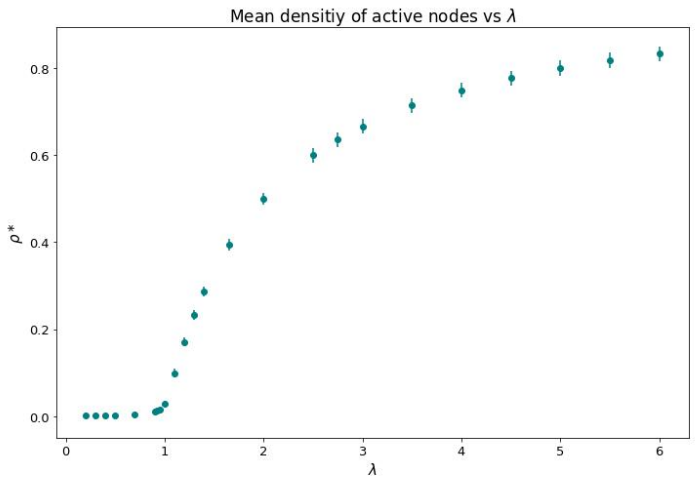
with:

$$I_{e/in\,i}^k = w_{e/in} exp\Big(-\frac{t - t_{s\,i}^k}{\tau_S}\Big)$$

and:

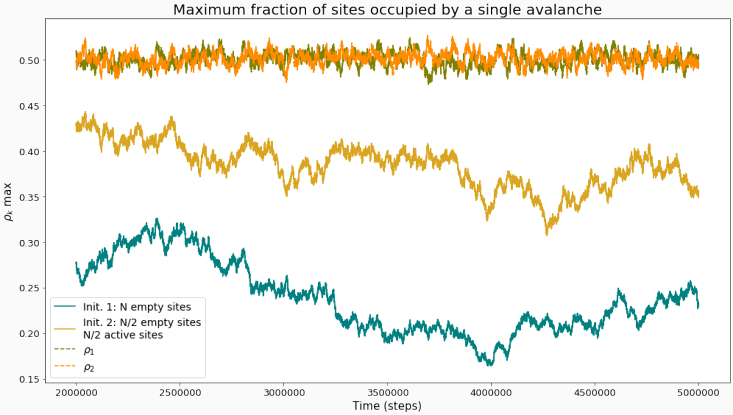
$$\dot{U}_{ij} = \frac{1 - U_{ij}}{\tau_R} - \sum_k U_{ij} \Theta(p_r - \zeta_{ij}^k) \delta(t - t_{s_i}^k).$$

Appendix ii

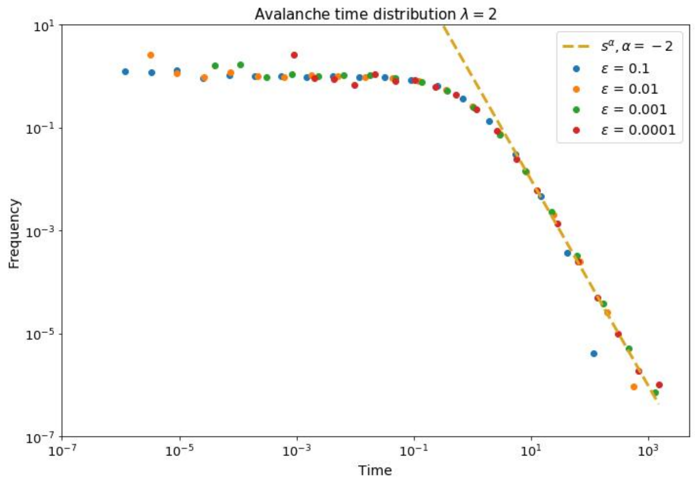


Mean densitiy of active nodes vs λ
ρ^*
0.0
0.2
0.4
0.6
0.8
0
1
2
3
4
5
6
λ

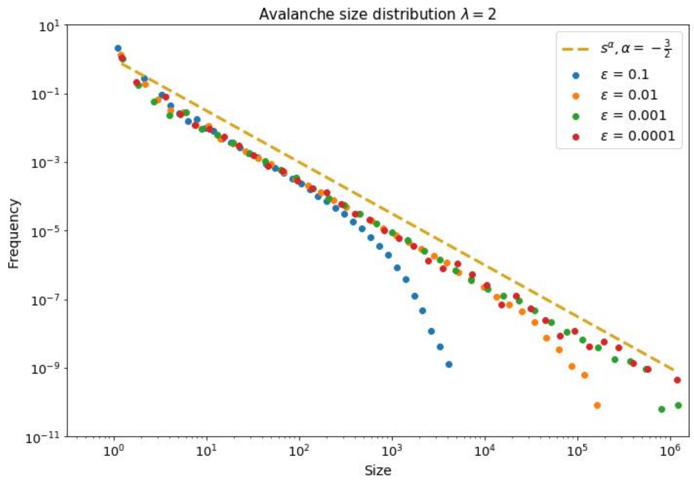
Appendix iii



Appendix iv

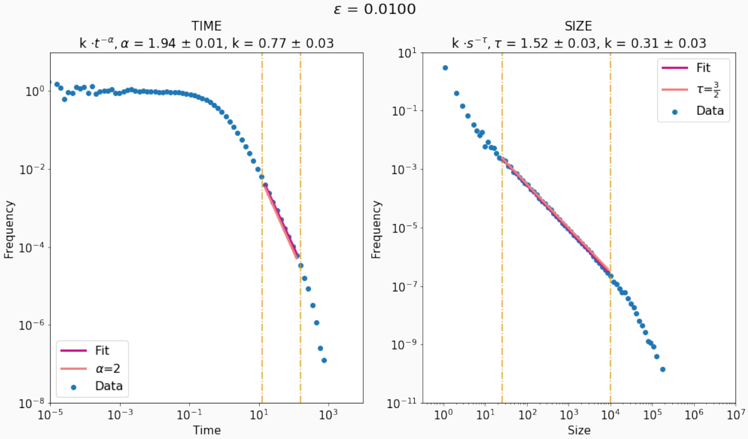

Avalanche time distribution $\lambda = 2$
$s^{\alpha}, \alpha = -2$
ε = 0.1
ε = 0.01
ε = 0.001
ε = 0.0001
Frequency
Time
10^{1}
10^{-1}
10^{-3}
10^{-5}
10^{-7}
10^{-7}
10^{-5}
10^{-3}
10^{-1}
10^{1}
10^{3}

Appendix v



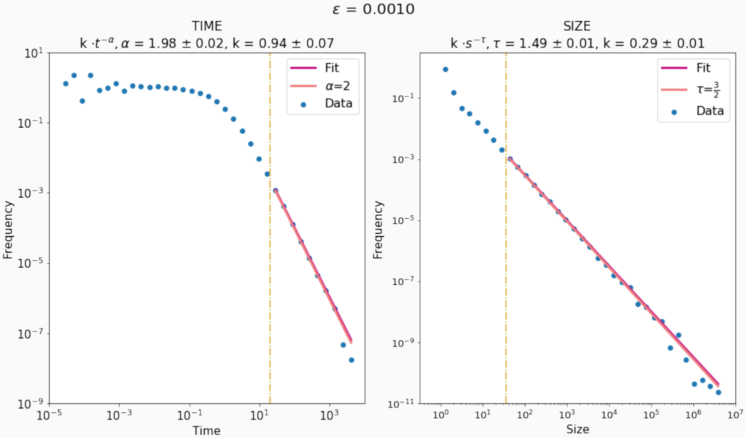
Avalanche size distribution $\lambda = 2$
$s^{\alpha}, \alpha = -\frac{3}{2}$
ε = 0.1
ε = 0.01
ε = 0.001
ε = 0.0001
Frequency
10^{1}
10^{-1}
10^{-3}
10^{-5}
10^{-7}
10^{-9}
10^{-11}
10^{0}
10^{1}
10^{2}
10^{3}
10^{4}
10^{5}
10^{6}
Size

Appendix vi

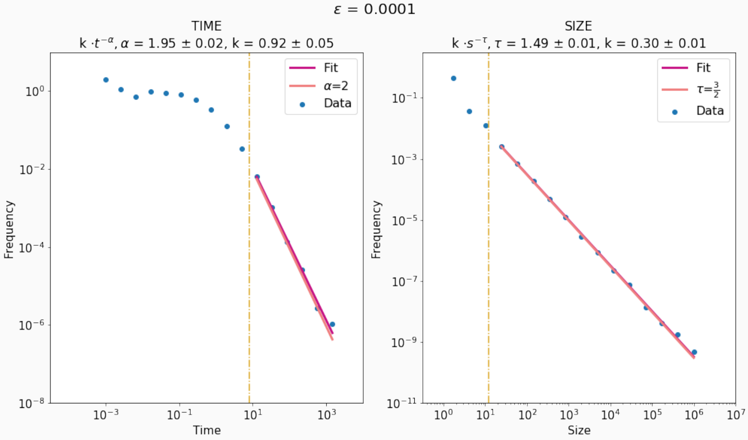


ε = 0.0100
TIME
k $\cdot t^{-\alpha}$, α = 1.94 ± 0.01, k = 0.77 ± 0.03
Frequency
Time
Fit
α=2
Data
SIZE
k $\cdot s^{-\tau}$, τ = 1.52 ± 0.03, k = 0.31 ± 0.03
Frequency
Size
Fit
$\tau=\frac{3}{2}$
Data

Appendix vii



Appendix viii



Appendix ix

Derivation of stationary solution for ρ (deterministic case):

$$\dot{\rho} = -\rho\mu + (1-\rho)\rho\lambda + (1-\rho)\epsilon \tag{6}$$

At stationarity $\dot{\rho} = 0$. Solutions are:

$$\rho* = \frac{\lambda - \mu - \epsilon \pm \sqrt{4\epsilon\lambda + (\lambda - \mu - \epsilon)^2}}{2\lambda} \tag{7}$$

Ignoring second order terms in ϵ and using $\sqrt{1+x} \sim 1 + \frac{x}{2}$ for small x one obtains:

$$\rho* = \frac{\lambda - \mu - \epsilon \pm (\lambda - \mu)\left(1 + \epsilon\frac{\mu+\lambda}{(\lambda-\mu)^2}\right)}{2\lambda} \tag{8}$$

Which leads to:

$$\frac{\epsilon}{\mu - \lambda} \quad \textit{for} \quad \mu > \lambda \tag{9a}$$

$$1 - \frac{\mu}{\lambda} + \epsilon \frac{\mu}{\lambda(\lambda - \mu)} \quad \textit{for} \quad \mu < \lambda \tag{9b}$$

Analytical derivation of the Langevin equation at stationarity.

Initial condition: $\rho_k = \frac{1}{N}$

Transition probabilities:

$$W^{+}(A_k) = \frac{A_k I}{N}\lambda \tag{10a}$$

$$W^{-}(A_k) = A_k \mu \tag{10b}$$

Master-equation:

$$\partial_t P(A_k, t) = -\Big[W^{+}(A_k) + W^{-}(A_k)\Big] P(A_k, t) + \\ + W^{+}(A_k - 1) P(A_k - 1, t) + W^{-}(A_k + 1) P(A_k + 1, t) \tag{11}$$

Continuous limit ($N \to \infty$) and Fokker-Planck equation:

$$\partial_t P(\rho, t) = \frac{\partial}{\partial \rho}\Big[A(\rho_k)P(\rho_k, t)\Big] + \frac{\partial^2}{\partial \rho^2}\Big[\frac{B(\rho_k)}{2N}P(\rho_k, t)\Big] \tag{12}$$

with:

$$A(\rho_k) = W^+(\rho_k) - W^-(\rho_k) = \Big[\lambda(1-\rho) - \mu\Big]\rho_k \tag{13a}$$

$$B(\rho_k) = \frac{W^+(\rho_k) + W^-(\rho_k)}{N} = \Big[\lambda(1-\rho) + \mu\Big]\rho_k \tag{13b}$$

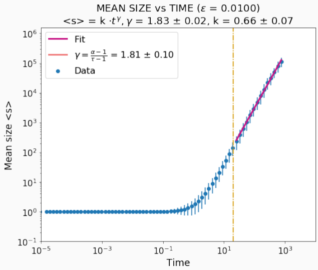
$$W^+(\rho_k) = \lambda\rho_k(1-\rho) \tag{13c}$$

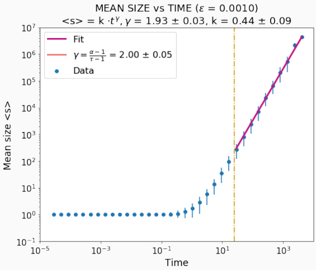
$$W^-(\rho_k) = \rho_k\mu \tag{13d}$$

Associated Langevin equation:

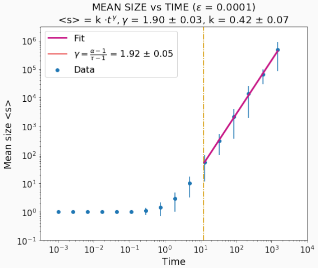
$$\dot{\rho}_k = A(\rho_k) + \sqrt{B(\rho_k)}\xi_k(t) =$$
$$= \rho_k\left[\lambda(1-\rho) - \mu\right] + \sqrt{\frac{\rho_k[\lambda(1-\rho)+\mu]}{N}}\xi_k(t) \quad (14)$$

Appendix xiv





Appendix xv



MEAN SIZE vs TIME (ε = 0.0001)
<s> = k ·t^γ, γ = 1.90 ± 0.03, k = 0.42 ± 0.07
Fit
γ = (α − 1)/(τ − 1) = 1.92 ± 0.05
Data
Mean size <s>
Time

References i