University of Padua

Department of Civil, Environmental and Architectural Engineering

M.Sc. in Mathematical Engineering

The Relationship between Multiscale Relevance and Network Properties in Simple Neural Networks of the Cortex

Supervisor
Prof. Samir Simon Suweis
Department of Physics and Astronomy
University of Padua

Co-Supervisor
Prof. Yasser Roudi
Kavli Institute for Systems Neuroscience
NTNU

Candidate
Simone Betteti

ID
1236340

To my family, who has faith in my abilities, vision, and has always supported me without compromises.

Preface

All sentient being are strongly bounded in the way in which they can efficiently and usefully manipulate their environment by their cortical circuits' proficiency at the meaningful encoding of incoming information into a compact and interpretable code, that ultimately results in an implicit (cognition) or explicit (motion) action. Cortical circuits are characterized by immensely complex networks of neurons and inter-neurons that continuously interact with each other by means of electrical currents and chemicals, with each of them propagating a non homogeneous series of bursts, called action potentials, that fully characterizes the neuron's code. It is still to be understood the extent to which this individual code ultimately impacts the organism's output, and whether and how the integration of many codes within and across cortical layers generates complex behaviour. Nonetheless, in order to gain some insight into how this organic maze works and how it is so efficient at doing what it does, it is necessary to start from an investigation of its building blocks, or its simplest units: neurons. Specifically, it would be interesting to understand whether, from the codes of neurons alone, it is possible to extract their overall encoding relevance for the task at hand, which in itself is a highly non-trivial goal given the simple nature of the signal. The problem arising from this type of quantification primarily resides in the lack of knowledge of the interactions in the network, the dynamics of the individual units in the network, and the time scale characterizing the generated code. A perhaps cumbersome approach at this quantification would first require the construction of an approximate model of the network under study, and only later an examination of the acquired codes; such method introduces a fair degree of arbitrariness in the whole process, thus potentially biasing the interpretation if not the results themselves. For this reason, it would be desirable to develop a coefficient that abstracts from any assumption on the underlying model and that at the same time is able to deliver a quantification of the relevance of the code generated by the unit under study.

The first chapter of this book will focus on the biological and mathematical description of the brain and its components, to the degree which I think is most suited for the reader to understand the rest of the thesis work. The second chapter will instead define Multiscale Relevance (MSR), which is, the coefficient for task relevance quantification mentioned above, along with a recent application to an ensemble of firing neurons. In the third chapter of the book, I will describe the theory behind a simple model of the cortex exhibiting sustained chaotic activity, to which the coefficient will be later applied and in the fourth chapter I will recall some basic concepts from graph theory. Finally, the remaining part of the thesis will be entirely devoted to the outline of the results obtained during my internship at the Kavli Institute for Systems Neuroscience at NTNU.

Contents

List of Figures

1

The Brain as a Complex System

1.1 Evolution of the Brain and its Anatomy

If we consider the life span of our planet, the Earth, which started its geological formation about 4.5 billion years ago, we immediately recognize how the dawn of intelligent life and, more specifically, the development of nervous systems, has only been a recent addition in history. Indeed, the first life forms presenting a rudimentary nervous system appeared on Earth about 700 million years ago [16], such as snails, and their system mainly consisted of collections of cells, known as ganglia, that were functional to the simplest motor and regulatory behaviour. It was only about 250 million years ago that life forms with the first brains developed and, given its position - embedded close to the major sensory organs - it provided a clear evolutionary advantage to those organisms, granting them rapid responses to hostile environments. From that moment on, those delicate yet extremely complex structures have constantly been under evolutionary pressure, with the human brain representing the pinna-

cle of this incredible process.

The human brain is an astonishingly complex organ constituted by approximately 100 billion neurons, each of which connects, on average, to other 10000 and enforce a small-world topology [26] on the network as whole. The average human brain has a volume of about 1400 cm^3, it is characterized by a heavily convoluted surface and is subdivided into four major cortices spanning two different interconnected hemispheres: the frontal cortex, the sensory-motor cortex, the parietal cortex and the occipital cortex. Yet, all of its processing units, the neurons characterizing the cortex, can be found in a strata of gray matter ranging from 1.5 mm to 4 mm in thickness and distributed at the surface; the bulk of the organ, the white matter, is characterized by nerve fibers connecting the different areas of the brain with each other as well as with the major pathways of the nervous system. The sheer size of the volume occupied by this connecting tissue should give the reader a vague idea of how complex the correlations in activity between different neurons and, more in general, different cortical areas are.

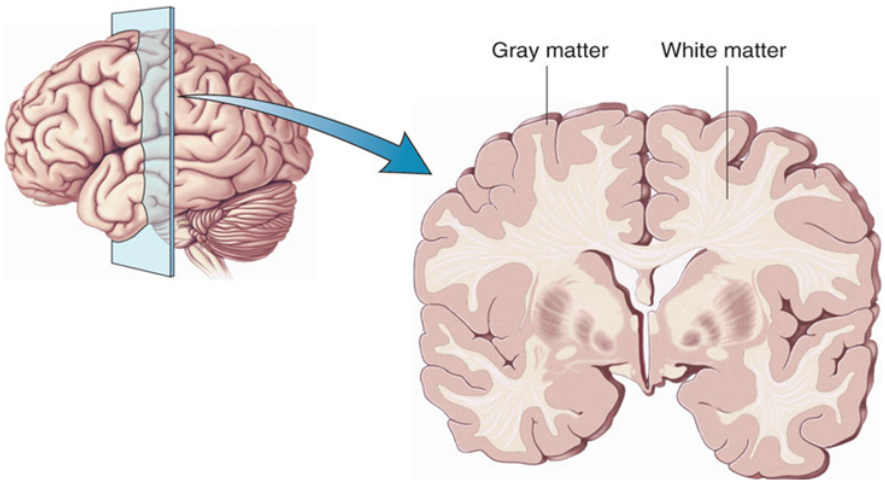


Figure 1.1: Visualization of the Difference between Gray and White matter (source:https://operativeneurosurgery.com/doku.php)

Focusing on the external layer of gray matter, and imagining to carve a perfect parallelepipedal column, we find six distinguishable layers characterized by different properties, which have

formed through a migration process during early fetal development. Namely

- Layer I: nerve fibers for connections with the other layers (upmost layer)
- Layer II: granule cells
- Layer III: pyramidal cells
- Layer IV: granule cells
- Layer V: pyramidal cells
- Layer VI: many types of neuron merging with the white matter below (down-most)

All these layer serve different functions and process different types of information, being it incoming from sensory modalities or from other layers. The coordination with which such information processing is carried out is the result of the functional evolution of the network topology, which undergoes a phase of increase in connection (synaptic) density, in tight relation with the exposition of the young organism to input coming from the external environment, and a subsequent phase of pruning, in which all the synapses that are not closely coupled with the activity of the neuron to which they relay information, die [23]. This topological evolutionary process, in humans, covers a period of about a year and a half, and should suggest to the reader how primary the tuning of an efficient network of synaptic connections is for the development of an information processing infrastructure that continuously deals with incoming complex external stimulation. It remains to be understood how this intricate network impacts the coding of its individual units.

1.2 The Neuron and the Action Potential

The cortex is characterized by two main types of cells, namely neurons (nerve cells) and glia (glia cells), and they serve different functions.
Glia cells greatly outnumber neurons [25], being 2 to 10 times more, and are not electrically excitable. Thereby, glia cells are not functional at the active transmission of information in the network, but rather have a support role, either through the provision of nutrients and

organic structure or through the engagement in immune system functions. Nonetheless, it would be imprudent to automatically discard glia cells from advanced modelling of information processing in biological neural networks (or neuronal networks), as their role in the generation of higher cognition may be as important as that of neurons. However, for the time being the latter will be the protagonists of our story.
Neurons are cells with a very peculiar structure, which in its most general classification can be schematized as follows

- **Soma**
  It is the main body of the nerve cell, and it hosts the nucleus with its genetic material as well as other cellular supportive structures that are functional at providing energy. It is also the part of the neuron where the electrical charge accumulates and, if not dissipated quickly enough, gives rise to the propagation of a signal.

- **Dendrites**
  Dendrites are long extension of the soma that collect incoming signals from other units in the network and propagate it according to the principles of temporal summation (many signal impinging on the same neuron in the interval $[t, t+\Delta t]$ are summed together) and spatial summation (signals arriving at the same dendritic branch are summed together).

- **Axon**
  The axon is an elongated branch of tissue that can span from $0.1\ mm$ to $2\ m$ and can be either not coated (slow signal propagation) or coated with a sheet of myelin (fast signal propagation). Each neuron has just one axon, which near its terminal part branches into many presynaptic terminals afferring to neighbouring neurons. The connecting region between the presynaptic terminal and the postsynaptic dendrite is called synapse, and can be either electrical (fast transmission) or chemical (slow transmission).

Aside from this very general descriptions, neurons vary greatly in their form, branching and functions, and this variability adds to the complexity of the information processing system as a whole, which is ultimately what allows the organism to respond quickly and selectively to changing external conditions.

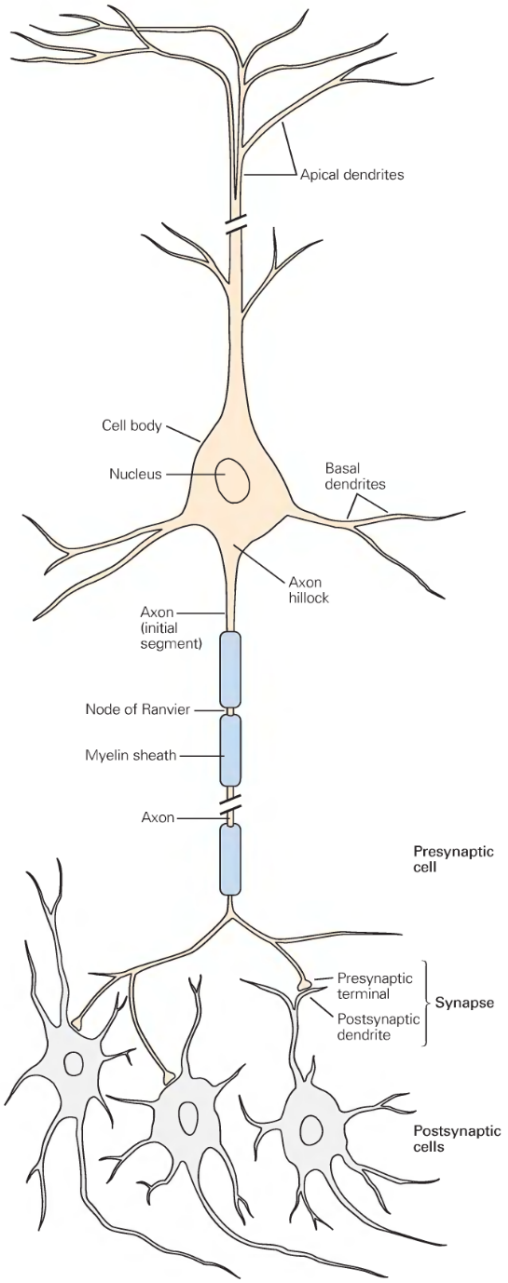


Figure 1.2: Structure of a neuron (source: Principles of Neural Science, *Kandel*[25])

The generation of the signal that conveys information to other neurons in the network takes place in the axon hillock (fig. 1.2), which is the initial segment of the axon. The gen-

erated signal, which then propagates across the entire length of the axon, is called action potential, and it is described as a all-or-none phenomenon. For the purpose of action potential generation, a particularly important component of the neuron is its cell membrane, the ionic conductance of which can change in response to different external stimuli [40], thus making neurons excitable cells. Indeed, before talking about the signal itself, we have to consider the membrane potential of the neuron, which at rest ranges from $-70\ mV$ to $-60\ mV$, and quantifies the difference in voltage between the intracellular and extracellular fluid, which is due to differences in ion concentration. As a matter of fact, the two fluids have a different ionic composition, with the former being rich in potassium cations (K^+) and macro-molecules P^-, while the latter is richer in sodium (Na^+), Calcium (Ca^{2+}), and chlorine (Cl^-). It is this difference in ion concentrations that ultimately makes the generation and propagation of the action potential possible, provided that the neuron is sufficiently excited.

When enough input currents, collected by the dendrites of the neuron, accumulate in the soma, depolirizing the cell to the point where the membrane potential is $-45mV$, an abrupt change in the ionic conductance of the cell membrane occurs. From this point on, the dynamic interplay between fluxes of Na^+ and K^+ continuously changes the ionic conductance of the cell membrane around the axon in both time and space, so that if $v(t, x) = 50\ mV$, then at time $t + dt$ the portion of the axon of coordinates $x + dx$ will be subjected to a significant Na^+ influx and a relatively smaller and late K^+ outflux, which then leads that portion of the membrane into a refractory state such that further excitation is impossible. Further actions by additional gates and sodium-potassium pumps will bring that axon segment back to its original equilibrium state, thus making the propagation of subsequent action potentials possible. In order to make the process clearer, consider the following image representing the propagation of an action potential in the giant axon of a squid, the same used by Hodgkin and Huxley for their revolutionary paper.

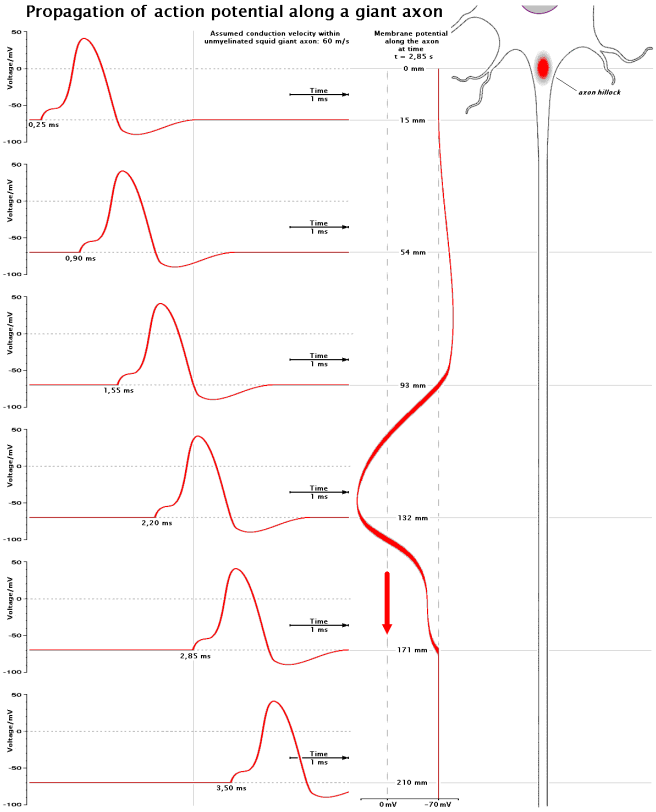


Figure 1.3: Propagation of an action potential in a giant axon of a squid (source: https://commons.wikimedia.org/wiki/File:Propagation_of_action_potential_along_a_giant_axon_en.png)

1.3 Formalization of Neurons as Point Processes

Consider a neuron n as a generic point in space, the coordinates of which are irrelevant, and which is subject to external inputs $\epsilon_{nj}(t)$ coming from a set of neighbouring neurons. Then we model the membrane potential of the chosen neuron, for a limited number of incoming inputs, as

$$u_n(t) = u_r + \sum_j \sum_{f \in F^j} \epsilon_{nj}(t - t_j^f) \tag{1.3.1}$$

where u_r is the resting value of the membrane potential and F^j denotes the set of firing times of neuron j [17]; instead, if the number of impinging signals is significantly high, then the interaction is non-linear. In both cases, the occurrence of a spike is modelled as an instan-

taneous and transient phenomenon, accurately defined by a Dirac's delta function (which in itself is a distribution rather than a function). Suppose then that neuron n fires at time $t_n^f \in [0, T]$; the event is captured by

$$\delta(t - t_n^f)$$

Thereby, the entire spike train associated to a neuron n in the interval $[0, T]$ is given by

$$S_n^T(t) = \sum_{f \in F(T)} \delta(t - t_n^f) \tag{1.3.2}$$

and here $F(T)$ denotes the firing time of the neuron up to time T. We will now exploit the definition of spike train to provide a more realistic model of neuronal firing, where we allow for non-linear interactions and we impose a reset condition, which brings the membrane potential instantaneously back to its resting value. Thus, for u_θ the membrane threshold for firing, we have

$$\frac{du_n}{dt}(t) = \frac{1}{\tau_m}\Big(- (u_n(t) - u_r) - RC(u_\theta - u_r)S_n(t) + RI(t)\Big) \tag{1.3.3}$$

where R and C are resistance and capacitance of the cell membrane, respectively, and $I(t)$ accounts for the non-linear input impinging on the cell. The term

$$C(u_\theta - u_r)S_n(t)$$

accounts for the instantaneous subtraction of charge every time that the threshold is reached. This model is known as Leaky-Integrate-and-Fire model (LIF) and provides a very simple yet powerful account of neuronal dynamics.

A more realistic model, which is also significantly harder to fully simulate, is the Hodgkin-Huxley model, formulated by Hodgkin and Huxley in the '50s, later granting them the No-

bel prize in medicine. The model starts from the same dynamic equation for the membrane potential

$$C\frac{du}{dt}(t) = I(t) - \sum_k I_k(t) \tag{1.3.4}$$

where the last term describes the contribution due to the fluxes of ionic currents. Specifically,

$$\sum_k I_k(t) = g_{Na}m^3h(u - E_{Na}) + g_K n^4 h(u - E_k) + g_L(u - E_L) \tag{1.3.5}$$

$$\begin{aligned} m &: Sodium\, activation\ \ (3\, sites) \\ n &: Potassium\, activation\ \ (4\, sites) \\ h &: Inactivation \end{aligned}$$

where the dynamics of the gates themselves are regulated by a set of O.D.E.s of the form

$$\frac{d\phi}{dt} = \frac{1}{\tau_\phi(u)}(\phi - \phi_0(u)) \qquad \phi \in \{m, n, h\}$$

This model, however, does not take into account the stochasticity inherent in channel opening; such phenomenon could be taken into account by gating dynamics

$$\frac{d\phi}{dt} = \alpha_\phi(u)(1 - \phi) + \beta_\phi(u)\phi$$

with

$$\begin{aligned} \alpha_\phi \Delta t &: p(\phi(t) = \ on \,|\phi(t - \Delta t) = \ off\ ) \\ \beta_\phi \Delta t &: p(\phi(t) = \ off \,|\phi(t - \Delta t) = \ on\ ) \end{aligned}$$

Above all, from any of the presented model we are able to extract a set of firing times for the neurons under study, where for example we could consider $\{t \in F(T)\}$ as the set of hitting

times in the interval $[0, T]$ satisfying the condition

$$t: \ u(t) = -45\, mV$$

that is, when a spike is generated. Thus, the information that we can extract from such models mainly consists of firing times, and we now venture into another avenue of the debate on how neurons encode information. There are two prevalent views on the issue

1. ***Rate (Frequency) View***
2. ***Spike-Times View***

1.3.1 Rate View

According to the rate view, most of the information carried by a spike-time series is contained in the cardinality of the set, or in the cardinality of different partitions of the set. Specifically, the rate of firing of a neuron when exposed to any given stimulus is in itself informative about the neuron selectivity and purpose, and it can be quantified in many ways. The most superficial coefficient that quantifies the firing rate of a neuron is

$$r_n(T) = \frac{1}{T} \int_0^T S_n^T(t)\, dt \tag{1.3.6}$$

This is a very coarse estimate of the firing rate, as we have no evidence that the spikes generated in the considered interval are archetypal of that neuron response to the environment (it is possible to record during an anomalous response). Instead, a finer measure of firing rate closes down on a smaller interval $[t, t + \Delta t]$ but, rather than taking the actual spike train response S_n^T, it considers an average, which could be over many trials or over the spike intervals distribution, if the latter is known.

$$r_n(t) = \frac{1}{\Delta t} \int_t^{t+\Delta t} \mathbb{E}[S_n^T](\tau)\, d\tau \tag{1.3.7}$$

Notice that, in this last case, taking the limit $\Delta t \to 0$ does not make much sense, as the interval needs to be big enough to actually contain a spike. Indeed, for very small time intervals the firing rate coincides with the neural response; suppose that $\mathbb{E}[S_n^T] \in L^1[0,T]$ and that its primitive is $\varphi : [0,T] \to \mathbb{R}^+$

$$\begin{aligned} r_n(t) &= \frac{\varphi(t+\Delta t) - \varphi(t)}{\Delta t} \\ &= \frac{\frac{d}{dt}\varphi(t)\Delta(t) + o(\Delta t^2)}{\Delta t} \\ &\approx \frac{d}{dt}\varphi(t) = \mathbb{E}[S_n^T](t) \end{aligned}$$

Thereby, it follows that any function of the neural response that results from the convolution of a property $h \in L^2[0,T]$ and the expected neuronal response can actually be written as

$$\begin{aligned} \mathcal{F}(T) &= \frac{1}{T}\int_0^T h(\tau)\mathbb{E}[S_n^T](t-\tau)\, d\tau \\ &= \frac{1}{T}\int_0^T h(\tau) r(t-\tau)\, d\tau \end{aligned}$$

where the latter should hold a great empirical advantage. As a matter of fact, in empirical contexts the firing rate is estimated through a binning procedure, so that if we want to know the quantity for a small interval $[t, t+\Delta t]$ we use

$$\omega(\tau) = \begin{cases} (\Delta t)^{-1} & \tau \in [t, t+\Delta t] \\ 0 \end{cases}$$

$$r_{appr}(t) = \int_{\mathbb{R}} \omega(\tau) S_n^T(t-\tau)\, d\tau \tag{1.3.8}$$

Since we are considering only first order moments of the spike train statistics, we can suppose subsequent spikes to be independent from one another, and subsequently model the timing for the entire series as a homogeneous Poisson process, or rather the probability to have k

spikes in a time window $[0, T]$.

$$p_T(n) = \frac{(rT)^k}{k!} e^{-rT} \tag{1.3.9}$$

1.3.2 Spike-Times View

The spike-times view, on the other hand, hypothesizes that most of the information in a neural code is contained in the timing of the spikes, and consequently also in the length of the interspikes intervals (ISI). Thereby, this stance holds that most of the informative content resides in higher order moments of the neural code, as the autocorrelation of the spike train response function, which is defined as

$$Q_{S_n S_n} = \int_{\mathbb{R}} \mathbb{E}[(S_n(t) - \mathbb{E}[r_n])(S_n(t+\tau) - \mathbb{E}[r_n])]\, dt \tag{1.3.10}$$

Under this assumption, we are expressing the fact that the timing of the different spikes, rather than their number, is what actually conveys information, and this allows to expand the framework to inter-neuron correlations, where we consider the correlation between the neural codes of two or more different neurons to gauge how much they depend on each other.

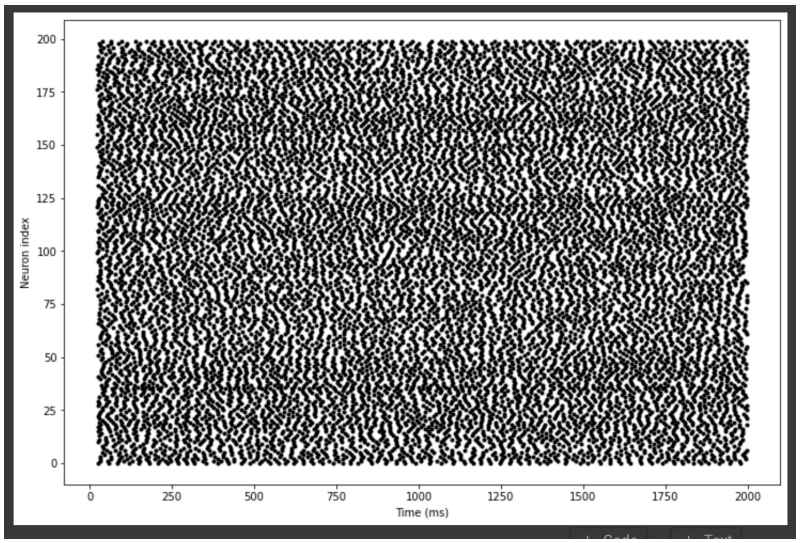
$$Q_{S_n S_m} = \int_{\mathbb{R}} \mathbb{E}[(S_n(t) - \mathbb{E}[r_n])(S_m(t) - \mathbb{E}[r_m])]\, dt \tag{1.3.11}$$

Consequently, since we are introducing higher order dependencies, we are in some sense considering the information content of the network as more sensible to state perturbations, and therefore as having higher variance across conditions. Furthermore, given the supposed dependencies and their importance for the framework, the neural code of a neuron can no longer be considered as generated by a homogeneous Poisson process. Instead, we must consider the firing rate as time dependent, and express

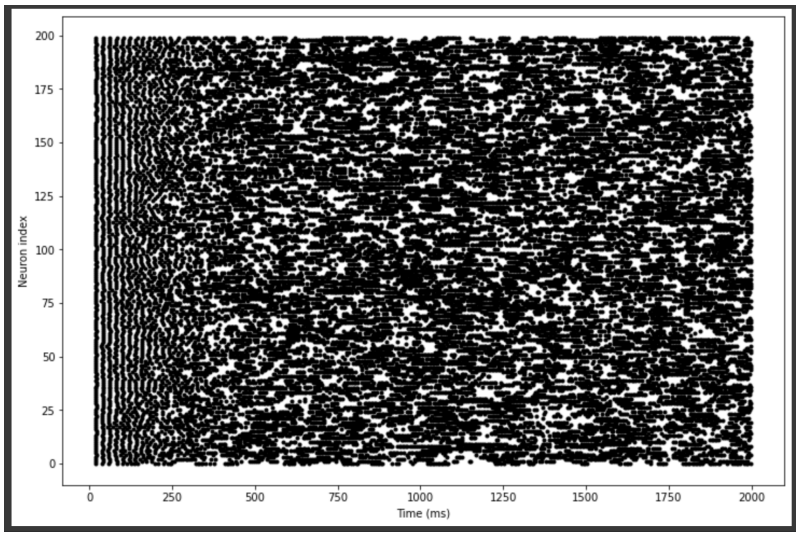
$$p_T(t_1, .., t_n) = e^{-\int_0^T r(t)\, dt} \prod_i^n r(t_i) \tag{1.3.12}$$

Notice that we still consider it a Poisson process, as the consideration of all the possible combination of inter-spike intervals yields

$$p_T(n) = \frac{p_T(t_1, .., t_n)}{n!}$$



(a) Slightly Coupled LIF neurons



(b) LIF neurons governed by double OU process

Figure 1.4: Example of correlated neural codes in a population - rasterplot visualization

To summarize, the reader can consider this principles as a modelling choice:

- $\left|\frac{dr}{dt}\right| < 1 \rightarrow$ rate code
- $\left|\frac{dr}{dt}\right| \gg 1 \rightarrow$ temporal code

Finally, we consider how entire networks of neurons are modelled, and how researchers try to understand their underlying mechanisms through the study of the spike statistics.

1.4 Statistical Mechanics and Neuronal Networks

The field of statistical mechanics deals with the study of large ensembles characterized by elementary units, the interaction of which results in the system exhibiting emergent properties - that is, properties that could not be studied or observed focusing on the individual unit. Since the systems taken into consideration are characterized by very large number of units (of the order of the Avogadro number $6.02 \cdot 10^{23}$), statistical mechanics cannot concern itself with the global deterministic dynamics of the system, but must instead focus on a probabilistic treatment of the same, narrowing on the statistical moments of the variable of interest. Throughout the years, it has provided invaluable insight into the behaviour of a wide breadth of physical systems, and specifically on some of their properties, such as the alignment of magnetic moments in lattices of interacting particles (Ising Model) or, more recently, in systems where the spatial structures among units is embodied in the randomness of the couplings (Spin Glasses). It has only recently started to gain traction on possible applications to neuroscience and artificial intelligence.

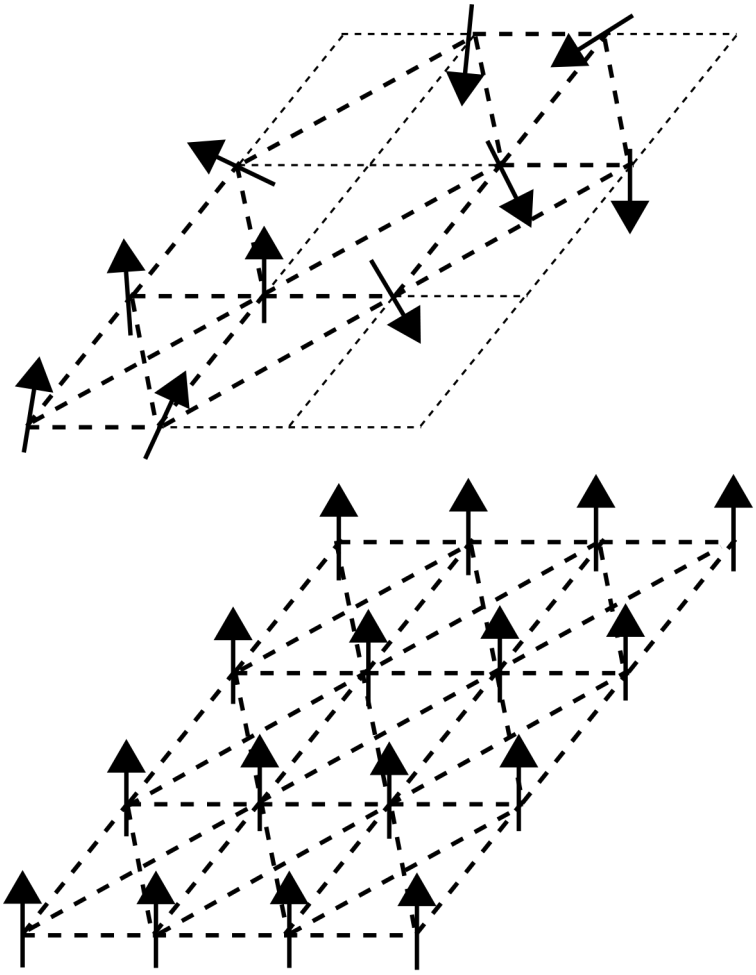


Figure 1.5: Visualization of a system in a Spin Glass State (above)

Consider now $\{\sigma_n\}$ as the set of neurons that have a synapse connected to neuron n's dendritic tree, and we are interested at understanding how the input incoming from all of these different sources affects the activity of this neuron. Suppose that it is possible to quantify the amount of neurotransmitters (chemicals) that are released at the synapse between any neuron in $\{\sigma_n\}$ and the receiving one, and that we proceed with such quantification for every neuron in the network. Then we can construct a coupling matrix J_{ij} such that $\forall i, j \in \{1, ..., N\},\ i \neq j$, the coefficient quantifies the amount of neurotransmitters that neuron j releases at the synapse afferring to neuron i, and finally evaluate the overall neuron-

dependent input as

$$I_n^{neu}(t) = \sum_{j \in \{\sigma_n\}} J_{ij} S_j^T(t)$$

By doing so, we obtain for the entire network of neurons a system of coupled differential equations (take n varing in {1,..,N}) of the following form

$$\frac{du_n}{dt} = \frac{1}{\tau_m}\Big(-(u_n(t) - u_r) - RC(u_\theta - u_r)S_n(t) + RI(t) + I_n^{neu}(t)\Big) \qquad (1.4.1)$$

This last set of equations gives a very detailed account of the dynamic state of each neuron in the network as time progresses. However, this level of details fails at providing relevant insight on how the activity of the network as a whole can result in information processing of some kind. In order to study the network response at a more fundamental level, it is necessary to take a step back and simplify the network model to the point where we consider neurons as units being either in an active or inactive state. For simplicity, and without notational overlap, consider the activity of neuron n as modelled by σ_n, with this being, in most of the cases, a variable obtaining two possible values. Furthermore, we completely neglect internal neuronal dynamics, as dictated by the preceding set of O.D.E.s, in favour of a magnetic field h_n which approximately quantifies internal contributions to the state; the meaning of the matrix J_{ij} remains the same, except that now we impose $J_{ij} = J_{ji}$ (symmetric matrix). Under these conditions, we can express the Hamiltonian of the network

$$H\{\sigma\} = -\frac{1}{2}\sum_{ij} J_{ij}\sigma_i\sigma_j - \sum_i h_i\sigma_i \qquad (1.4.2)$$

which quantifies the energy of the system when this is in the specific state σ. Notice that at any given time there are 2^N possible state configurations, and the update dynamics for the system can be either synchronous (the state value for all the units in the network is updated at the same time) or asynchronous (the state value is updated for one unit at a time). Then the probability of the system being in a specific state σ is given by the Boltzmann distribution

$$p(\{\sigma\}) = \frac{e^{-\beta H\{\sigma\}}}{Z} \tag{1.4.3}$$

where $\beta = (kT)^{-1}$ for T temperature parameter and k Boltzmann constant, and Z is the partition function

$$Z = \sum_{\{\sigma\}} e^{-\beta H\{\sigma\}} \tag{1.4.4}$$

which is not always tractable. This type of modelling has been widely exploited both for the study of simple cortical networks of associative memory and feature extraction, as well as, more in general, for other types of complex systems (e.g. magnets, financial networks, etc..).

1.4.1 The Hopfield Model

One of the most famous applications of statistical mechanics to neuroscience has been the so called Hopfield model, which formalizes how a system of simple units can, through systemic response, instantiate an associative memory mechanism. Specifically, the author observed how, by drastically simplifying the inherent dynamics of the units in the network, it was possible to observe the rise of memory as an emergent property of the system.
Maintaining the previous formalism, we will define a set of binary variables $\{\sigma_i\}_{i=1,..,N}$ which will be the units of our network. Now, suppose to have a set $M = \{X^1, .., X^n : X \in \{0,1\}^N\}$ of memories to be stored, and set the coupling matrix as [22]

$$J_{ij} = (1 - \delta_{ij}) \sum_{k=1}^{n} (2\sigma_i^k - 1)(2\sigma_j^k - 1) \tag{1.4.5}$$

where σ_i^k is the i^{th} component of the k^{th} memory in the set. The idea is then that, with this type of initialization, any instantiation of the network starting from a generic state σ will converge, according to the following asynchronous dynamics

$$\sigma_i = \sum_j J_{ij}\sigma_j \tag{1.4.6}$$

to the memory $X^h \in M$ such that $X^h = arg\, min_{x \in M} ||\sigma - x||$ at the start. If we consider the associated Hamiltonian

$$H\{\sigma\} = -\frac{1}{2} \sum_{ij} J_{ij} \sigma_i \sigma_j \tag{1.4.7}$$

and suppose that the following quantity is actually computable, we readily see how convergence gradually decreases the energy of the system, and leads it, from the way in which J_{ij} is initiated, to an energy well.

$$\frac{\partial E}{\partial \sigma_i} = -\sum_j J_{ij} \sigma_j < 0 \tag{1.4.8}$$

Indeed, the stationary states constitute local minima of the energy function $H\{\sigma\}$ with respect to the flipping of the variables $\{\sigma_i\}_{i=1,..,n}$[8]. As a matter of fact, it is the tuning of the couplings according to equation 1.4.5 that molds the energy surface $\tilde{H} = \{(\sigma, H\{\sigma\})\}$ so that the chosen configuration minimize H and, consequently, the obtained manifold instantiate a form of content-addressable memory. Each of the chosen configurations can later be correctly retrieved if the starting configuration of the system lies within its basin of attraction.

An ingenuous transformation of the system states $\sigma_i \to S_i : \quad S_i \in \{-1, +1\}$

$$S_i = 2\sigma_i - 1 \tag{1.4.9}$$

and of the Hamiltonian [4] as

$$H\{S\} = -\frac{1}{2} \sum_{ij} J_{ij} S_i S_j \tag{1.4.10}$$

where now the storing of specific patterns by means of the coupling matrix is given by

$$J_{ij} = (1 - \delta_{ij}) \frac{1}{n} \sum_{k=1}^{n} S_i^k S_j^k$$

allows us to draw more similarities with a classical spin-glass model, having states in $\{-1, 1\}$ and couplings that are quenched but of stochastic nature. Differently from the very popular SK model, the couplings in the Hopfield model are not drawn from a probability distribution (Gaussian, in this case), but are rather externally imposed on the base of the patterns that need to be stored.

Suppose now that we are interested at studying the thermodynamics of the model, and therefore that we must derive its free energy. Formally, it is defined as

$$-\beta\mathcal{F} = \mathbb{E}_\rho[log(Z)] \tag{1.4.11}$$

for a generic probability distribution ρ, and where we write, for notational simplicity

$$Z = \sum_\sigma e^{-\beta H\{\sigma\}} = Tr_\sigma\, e^{-\beta H\{\sigma\}} \tag{1.4.12}$$

The computation of the expected value of the logarithm of the partition function is oftentimes not possible, and for this reason the same functional must be evaluated by means of the replica method [31], which exploits the trick of "replicas", even though they loose physical meaning in the limit.

$$-\beta\mathcal{F} = \lim_{l\to 0} \frac{\mathbb{E}_\rho[Z^l] - 1}{l} \tag{1.4.13}$$

For this passage, we are implicitly supposing that the product $log(Z) \cdot \rho$ satisfies the hypothesis of the Lebesgue dominated convergence theorem, so that then the limit and expected value operators commute.

Following computations analogous to those carried out for the SK model, hence introducing additional Gaussian variables and the associated Gaussian densities, we get

$$\mathbb{E}[Z^l] = \prod_{\gamma<\delta} \int_{-i\infty}^{+i\infty} \left(\frac{N\beta^2\alpha}{2\pi}\right) dq^{\gamma\delta} \int_{-\infty}^{+\infty} dt^{\gamma\delta} \prod_{\gamma} \int_{-i\infty}^{+i\infty} \left(\frac{N\beta^2\alpha}{4\pi}\right) dx^{\gamma} \times \int_{-\infty}^{+\infty} dy^{\gamma} \prod_{\gamma, k\in\Gamma_\gamma} \int \frac{dm_\gamma^k}{(2\pi/N\beta)^{\frac{1}{2}}} e^{[nG(\{m_\gamma^k, x^\gamma, y^\gamma, t^{\gamma\delta}, q^{\gamma\delta}\})]} \tag{1.4.14}$$

with Γ_γ overlap between the (chosen) configuration k and the replica γ. The functional expression of $G(\{m_\gamma^k, x^\gamma, y^\gamma, t^{\gamma\delta}, q^{\gamma\delta}\})$ is not relevant for the current exposition (but can be found in [8]). However, the variables in G, and specifically their saddle point, identify the order parameters for the model, and under the assumption of replica symmetry they have the following physical interpretation

$$m_r = \mathbb{E}[(2\sigma_i^k - 1)\bar{\sigma}_i] \tag{1.4.15}$$

the overlap of the configuration k with the thermodynamic state,

$$q = \frac{1}{n}\sum_i \mathbb{E}[(\bar{\sigma}_i)^2] \tag{1.4.16}$$

$$x = \frac{1}{n}\sum_i \overline{\mathbb{E}[\sigma_i]} \tag{1.4.17}$$

are the thermal fluctuations and mean of the spin-glass state, and

$$t = \frac{1}{\alpha}\sum_{k\in\Gamma_\gamma} \left(\frac{1}{n}\sum_i (2\sigma_i^k - 1)\bar{\sigma}_i\right)^2 \tag{1.4.18}$$

$$y = \frac{1}{\alpha}\sum_{k\in\Gamma_\gamma} \overline{\left(\frac{1}{n}\sum_i (2\sigma_i^k - 1)\sigma_i\right)^2} \tag{1.4.19}$$

mean and thermal fluctuations of the overlap between the thermodynamic state and those nominated configurations with which its overlap is microscopic. In the given definition, the

bar above a certain quantity defines the thermal expectation, which is the expectation of the chosen quantity under the configurations of the system (instead, the expectation $\mathbb{E}$ is over the distribution of the couplings). The order parameters of the model are relevant for the understanding of how the system behaves in its different phases, and specifically how the system drastically alters some of its property during the phase transitions.

Studies of this type highlight the primary relevance of statistical mechanics in the investigation of large scale information processing in biological and artificial systems, by uncovering how the interaction among very simple units can result in the emergence of complex behaviour.

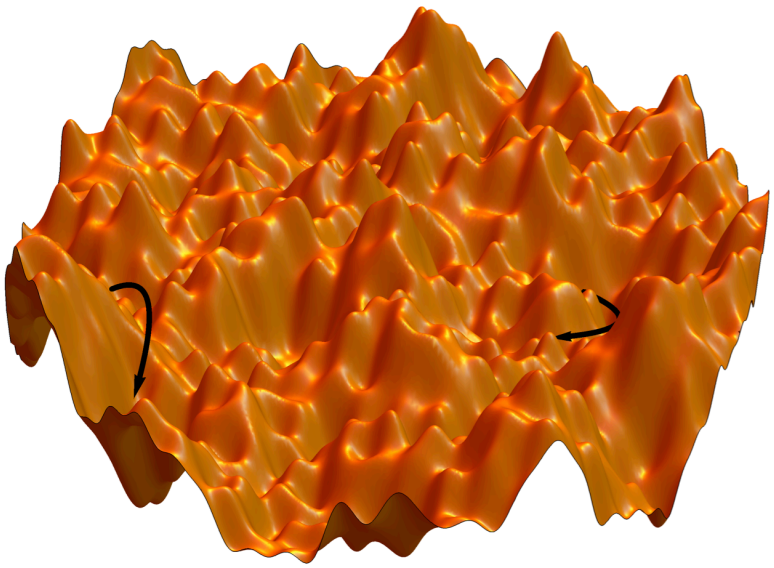


Figure 1.6: Visualization of the energy landscape for a spin glass model (source: https://scglass.uchicago.edu/wp-content/uploads/2017/11/Figure6.png)

2

Multiscale Relevance - MSR

Suppose to have a sample $\mathcal{S}$ drawn from an unknown probability distribution, and that we are interested at quantifying the informative content of $\mathcal{S}$ on the generative process. Most of the pre-existing theory bases this quantification on the a priori knowledge (or hypothesis) of the underlying model, or at least of the class of models to which the generative process belongs, thereby introducing a certain degree of arbitrariness in the assessment. Then given such approximate knowledge, it would be possible to assess how informative $\mathcal{S}$ is on the generative process, i.e. how probable the sample is under the unknown distribution. However, the introduction of extra assumptions can seriously bias the assessment, given that most of the times, for a generic sample, we have poor if any understanding of the underlying generative process, let alone the scale at which the generated events (elements of the sample) unfold. For this reason, it would be useful to devise a measure, or more loosely a sample metric, that exempts from any assumption on the underlying generative process, and only gauges the informative content of the sample based on its statistics.

Consider then a sample $\mathcal{S} = (s_1, .., s_N)$ of data-points drawn from a generic alphabet $\mathcal{A}$, and suppose that this events are sampled independently, which is formally equivalent to $\{s_i\}_{i=1,..,N}$ being i.i.d. and drawn from an unknown probability distribution; furthermore, suppose total absence of information on $\mathcal{A}$ to make the argument as general as possible. We are now interested at quantifying the informative content of our sample, expressed in the number of bits that are necessary to represent one of the outcomes [11]. Let

$$k(s) = \sum_{s_i \in \mathcal{S}} \delta(s - s_i) \quad s \in \mathcal{A} \tag{2.1.1}$$
$$N = \int_{s \in \mathcal{A}} k(s)\, ds$$

and broadly define the entropy (associated to the sample)

$$\hat{H}[s] = -\int_{s \in \mathcal{A}} \frac{k(s)}{N} \cdot log_2\left(\frac{k(s)}{N}\right) ds \tag{2.1.2}$$

Given our definition of $\hat{H}$ as the quantity of encoding bits necessary to represent $\mathcal{S}$, we will generally have that some of these bits will in fact be useful information, while others just noise, hence be random variables that, given the useful information, have a maximum entropy distribution [24] (statistical mechanics definition of noise). Given this differentiation, we are interested at better quantifying such useful information, which will be encoded by a set of parameters $\{\phi\}$, starting from the following hypothesis:

1. The data of the sample conditioned on $\{\phi\}$ is as random as possible.
2. $\{\phi\}$ provides the most succinct representation of the useful information
   $\rightarrow$ minimum description length.

A set of parameters $\{\phi\}$ satisfying the previous two hypothesis is defined as a *minimally sufficient representation*.

Without loss of generality, we can consider the alphabet $\mathcal{A}$ to be finite, and accordingly we will consider a set of features $\mathcal{O} = \{\phi_1, ..., \phi_N\}$ for $\phi_i \in \Phi$ finite set of values, and define the augmented set

$$\hat{d} = (s, \phi)$$

so that the total information content associated to $\hat{d}$ is

$$\begin{aligned}
\hat{H}[s, \phi] &= -\sum_{s \in \mathcal{S}} \sum_{\phi \in \mathcal{O}} \hat{p}_{s,\phi} \cdot log_2(\hat{p}_{s,\phi}) \\
&= -\sum_{s \in \mathcal{S}} \sum_{\phi \in \mathcal{O}} \hat{p}_{s,\phi} \cdot log_2(\hat{p}_s \cdot \hat{p}_{\phi|s}) \\
&= -\left[\sum_{s \in \mathcal{S}} \sum_{\phi \in \mathcal{O}} \hat{p}_{s,\phi} \cdot log_2(\hat{p}_s) \; + \; \sum_{s \in \mathcal{S}} \hat{p}_s \sum_{\phi \in \mathcal{O}} \hat{p}_{\phi|s} \cdot log_2(\hat{p}_{\phi|s}) \right] \\
&= \hat{H}[s] + \hat{H}[\phi|s] \qquad (2.1.3)
\end{aligned}$$

In order for the set of parameters $\mathcal{O}$ to be eligible for the definition of minimally sufficient representation, it must satisfy

1. Features $\phi \in \mathcal{O}$ do not introduce additional information on the generative process

$$\hat{H}[\phi|s] = 0$$

2. $\hat{H}[s|\phi]$ accounts only for noise in the sample

and we will now examine how this two assumptions interplay with the given definition of entropy. Specifically, observing that we can rewrite

$$\hat{H}[\phi|s] = -\sum_{s \in \mathcal{S}} \sum_{\phi \in \mathcal{O}} \hat{p}_{\phi} \hat{p}_{s|\phi} \cdot log_2(\hat{p}_{\phi|s})$$

then we can exploit the first requirement to obtain a new definition of the entropy; namely

$$\begin{aligned}
\hat{H}[s] &= -\sum_{s\in\mathcal{S}} \hat{p}_s \cdot log_2(\hat{p}_s) \\
&= -\sum_{\phi\in\mathcal{O}} \hat{p}_\phi \cdot log_2(\hat{p}_\phi) \; - \sum_{s\in\mathcal{S}} \hat{p}_s \cdot log_2(\hat{p}_s) \; + \sum_{\phi\in\mathcal{O}} \hat{p}_\phi \cdot log_2(\hat{p}_\phi) \\
&= \hat{H}[\phi] \; - \sum_{s\in\mathcal{S}}\sum_{\phi\in\mathcal{O}} \hat{p}_{s|\phi}\hat{p}_\phi \cdot log_2(\hat{p}_s) \\
&\quad + \sum_{\phi\in\mathcal{O}} \hat{p}_\phi \left(\sum_{s\in\mathcal{S}} \hat{p}_{s|\phi} \right) \cdot log_2(\hat{p}_\phi) \; + \hat{H}[\phi|s] \\
&= \hat{H}[\phi] \; - \sum_{s\in\mathcal{S}}\sum_{\phi\in\mathcal{O}} \hat{p}_\phi\hat{p}_{s|\phi} \cdot log_2\left(\frac{\hat{p}_s}{\hat{p}_\phi}\right) \; - \sum_{s\in\mathcal{S}}\sum_{\phi\in\mathcal{O}} \hat{p}_\phi\hat{p}_{s|\phi} \cdot log_2(\hat{p}_{\phi|s}) \\
&= \hat{H}[\phi] \; - \sum_{\phi\in\mathcal{O}} \hat{p}_\phi \sum_{s\in\mathcal{S}} \hat{p}_{s|\phi} \cdot log_2(\hat{p}_{s|\phi}) \\
&= \hat{H}[\phi] + \hat{H}[s|\phi]
\end{aligned} \tag{2.1.4}$$

Focusing now on the second requirement, we are imposing a state of maximal ignorance on the conditional $\hat{p}(s|\phi)$, hence that

$$\hat{p}(s|\phi) = \hat{p}(\bar{s}|\phi) \quad \forall s, \bar{s} : \phi(s) = \phi(\bar{s}) = \tilde{\phi}$$

and since we also have that

$$\begin{aligned}
\hat{p}(s) &= \hat{p}(s|\phi(s))\hat{p}(\phi(s)) \\
&= \hat{p}(\bar{s}|\phi(\bar{s}))\hat{p}(\phi(\bar{s})) \\
&= \hat{p}(\bar{s})
\end{aligned}$$

then by necessity, considering the finite alphabet we must have the equivalence for the associated frequencies

$$k(s) = k(\bar{s}) \tag{2.1.5}$$

so that, accordingly, $k(s) \neq k(\tilde{s}) \quad \forall s, \tilde{s} : \phi(s) \neq \phi(\tilde{s})$, which implies that we can express $k(s)$ as a function of $\phi(s)$ (in fact, we have a bijective map between the two).
In order to understand the implications of this last fact, we have to resort to tools specific of information theory and, specifically, to hinge on the data processing inequality. In plain terms, the data processing inequality asserts that no clever manipulation of the data can improve the inference that can be made from it [10].

Theorem 1 (Data Processing Inequality). *Let $X \to Y \to Z$ be a Markov Chain and $I(\ ;\ )$ the mutual information operator, defined as*

$$I(X;Y) = H[X] - H[X|Y]$$

Then

$$I(X;Y) \geq I(X;Z)$$

Proof.

$$\begin{aligned} I(X;Y,Z) &= H[X] - H[X|Y,Z] \\ &= I(X;Z) + I(X;Y|Z) \\ &= I(X;Y) + I(X;Z|Y) \end{aligned}$$

and since X and Z are independent given Y, then

$$I(X;Z|Y) = 0$$

$$I(X;Y|Z) > 0$$

□

where we have used, extending on the mutual information

$$\begin{aligned} I(X;Y,Z) &= \sum_{X,Y,Z} p(X,Y,Z) \cdot log_2\left(\frac{p(X,Y,Z)}{p(X)p(Y,Z)}\right) \\ &= \sum_{X,Y,Z} p(X,Y,Z) \cdot log_2\left(\frac{p(Z|X,Y)p(X,Y)}{p(X)p(Y)p(Z|Y)}\right) \\ &= \sum_{X,Y,Z} p(X,Y,Z) \cdot log_2\left(\frac{p(X,Y)}{p(X)p(Y)}\right) \\ &\quad + \sum_{X,Y,Z} p(X,Y,Z) \cdot_{log_2} \left(\frac{p(Z|X,Y)}{p(Z|Y)}\right) \\ &= I(X;Y) + \sum_{X,Y,Z} p(X,Y,Z) \cdot log_2\left(\frac{p(X,Z|Y)}{p(X|Y)p(Z|Y)}\right) \\ &= I(X;Y) + I(X;Z|Y) \end{aligned} \tag{2.1.6}$$

In addition, we will also need

Corollary 1. *In particular, if* $Z = g(Y)$ *we have*

$$I(X;Y) \geq I(X;g(Y)) \tag{2.1.7}$$

We can now apply this tools to the features ϕ and the frequency $k(s)$, which yields the following inequality

$$\begin{aligned} I(s;\phi(s)) &= H[s] - H[s|\phi(s)] \\ &= H[\phi(s)] + H[s|\phi(s)] - H[s|\phi(s)] \\ &\geq I(s;k(s)) = H[s] - H[s|k(s)] \\ &= H[k(s)] + H[s|k(s)] - H[s|k(s)] \end{aligned} \tag{2.1.8}$$

$$\Rightarrow \hat{H}[\phi(s)] \geq \hat{H}[k(s)] \tag{2.1.9}$$

Since for $\phi(s)$ to constitute a minimally sufficient representation we have that $H[s|\phi(s)]$ must be maximised over the set of all possible features Φ, and since $H[s]$ is constant for any given sample $\mathcal{S}$ with non-changing classification, then we have that

$$\phi(s) = arg\, min_{\hat{\phi}\in\Phi} H[\hat{\phi}]$$

At the same time, minimally sufficient representation must saturate the inequality 2.1.8. Consequently, we reach the following proposition [11][12]

Proposition 1. *The frequency $k(s)$ provides a minimally sufficient representation of the sample $\mathcal{S}$ in the sense that*

1. *The total content of the sample $\mathcal{S}$ can be divided as*

$$\hat{H}[s] = \hat{H}[k] + \hat{H}[s|k]$$

with

$$\hat{H}[k] = -\sum_k \frac{k \cdot m_k}{N} \cdot log_2\left(\frac{k \cdot m_k}{N}\right) \tag{2.1.10}$$

where m_k is the number of outcomes s for which $k(s) = k$, and

$$\hat{H}[s|k] = \sum_k \frac{k \cdot m_k}{N} \cdot log_2(m_k) \tag{2.1.11}$$

2. *In the absence of prior information, $\hat{H}[k]$ is the maximum number of bits (per data point) that can be used to estimate the underlying generative process, and $\hat{H}[s|k]$ is a measure of noise.*

Before diving further into the matter, an important remark on non-distinguishable samples is due. Suppose to have two outcomes $s, \tilde{s}$, each characterized by $l \in \mathbb{N}$ sample points and such that $k(s) = k(\tilde{s})$. The formal distinction between s and $\tilde{s}$ within $\mathcal{S}$ is the result of some arbitrary or a priori classification. Any alternative classification leaving $\mathcal{S}$ unvaried except for a reshuffling of the $2l$ data points into different outcomes of equal size l would finally

result in a sample yielding the same statistics, i.e. the data itself is void of any information on the classification characterizing the sample that would make it distinguishable, according to the defined statistics, from any other yielding the same class sizes. The degeneracy resulting from all this possible alternative classifications is captured by $\hat{H}[s|k]$. The arising question is how statistically equivalent classifications emerge, and how to formulate a metric that abstracts from all of them, thus providing a quantification that depends exclusively on the data points of the underlying generative process.

2.2 Resolution, Relevance, and MIS

Suppose now that we are interested in events that can be described, in their entire complexity, as vectors $\vec{x} \in \mathbb{R}^n$ for $n \gg 1$, and that we dispose of a submersion $s : \mathbb{R}^n \to \mathbb{R}^m$ for $m \ll n$ such that it encodes a description of the dataset $\mathcal{X} = \{\vec{x}_1, .., \vec{x}_N\}$ at a given level of details. This formalization allows us to represent rigorously the operation of classification for our objects. Suppose now, more loosely, that $s : \mathbb{R}^n \to \mathcal{H}$ where $\mathcal{H}$ is a countable set of labels. Then

- $\vec{x}$: Complete description of the object
- $\vec{s} = s(\vec{x}) \in \mathcal{S} \subset \mathcal{H}$ finite set of discrete labels

Suppose now to consider a family $\{s^i\}_i \;\; s^i : \mathbb{R}^n \to \mathcal{H}_i$ where each space $\mathcal{H}_i$ provides a description of the same dataset $\mathcal{X}$ at different levels of details. It does follow that the coding cost $\hat{H}[s]$ depends implicitly on $\mathcal{H}_i$ and this will result in different costs for different levels of details. Specifically, if we gauge the extremes of the frequency spectrum

1. Let $k(s) \in \{0, 1\} \;\; \forall s \in S \longrightarrow$ very fine level of details. Then

$$\begin{aligned} \hat{H}[s] &= -\sum_{s \in \mathcal{S}} \frac{k(s)}{N} \cdot log_2\left(\frac{k(s)}{N}\right) \\ &= -\frac{N}{N} \cdot log_2\left(\frac{1}{N}\right) \\ &= log_2(N) \end{aligned} \tag{2.2.1}$$

It thus follow that any potentially finer level of details would result in a mere relabelling of the objects, with the information content remaining the same.

2. Let $k(s_l) = N$ and $k(s_j) = 0 \ \forall j \neq l \rightarrow$ very coarse description. Then

$$\begin{aligned} \hat{H}[s] &= -\sum_{s \in \mathcal{S}} \frac{k(s)}{N} \cdot log_2\left(\frac{k(s)}{N}\right) \\ &= -\frac{N}{N} \cdot log_2\left(\frac{N}{N}\right) \\ &= 0 \end{aligned} \tag{2.2.2}$$

From the previous two point, we can formalize

Definition 1. *$\hat{H}[s]$ quantifies in bits the level of details of the chosen description $s(\vec{x})$ and it is called **resolution**.*

Now, the finest level of details represent the maximum resolution, while the coarsest the minimum, and consequently we have a range of intermediate values of $\hat{H}[s]$, each associated to a value $\hat{H}[k]$, which is an intrinsic property of the sample at that level of details [11].

Definition 2. *$\hat{H}[k]$ quantifies the amount of information that the sample of labels $\mathcal{S}$ contains on the generative process, and it is called **relevance**.*

Moreover, the relevance, which is a positive quantity, satisfies the following properties

1. $\hat{H}[s] = log_2(N) \rightarrow k(s) \leq 1$

$$\begin{aligned} \hat{H}[k] &= -\sum_{k} \frac{k \cdot m_k}{N} \cdot log_2\left(\frac{k \cdot m_k}{N}\right) \\ &= -\frac{1 \cdot N}{N} log_2\left(\frac{1 \cdot N}{N}\right) \\ &= 0 \end{aligned} \tag{2.2.3}$$

2. $\hat{H}[s] = 0 \rightarrow k(s) = N$

$$\begin{aligned}\hat{H}[k] &= -\sum_k \frac{k \cdot m_k}{N} \cdot log_2\left(\frac{k \cdot m_k}{N}\right) \\ &= -\frac{N \cdot 1}{N} log_2\left(\frac{N \cdot 1}{N}\right) \\ &= 0\end{aligned} \tag{2.2.4}$$

Given the bound $\hat{H}[k] \leq \hat{H}[s]$ and the previous property we have, under a continuum assumption, that the curve $(\hat{H}[s], \hat{H}[k])$ is parabolic and upper bounded by the line $\hat{H}[s] = \hat{H}[k]$.

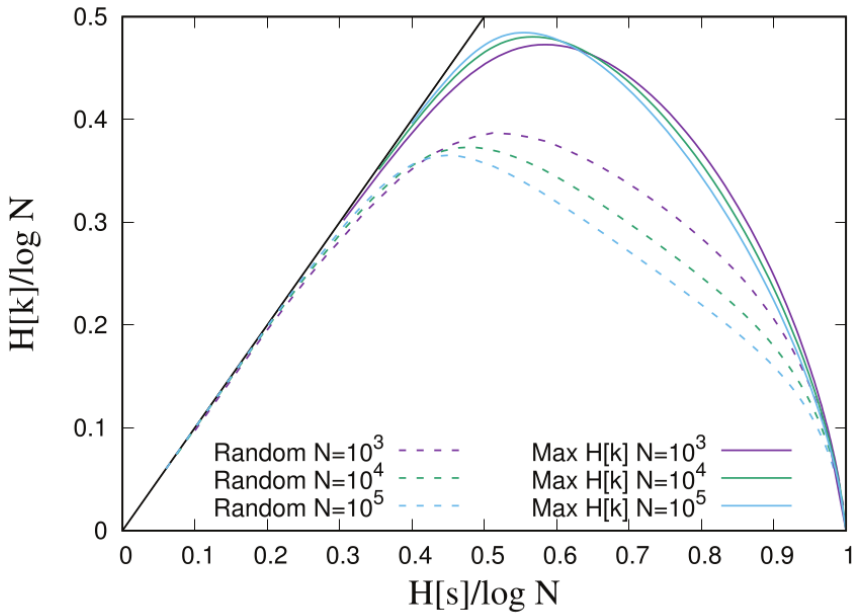


Figure 2.1: Visualization of possible curves in $\hat{H}[s] - \hat{H}[k]$ space (source: [13])

The region on the right of $max_{\hat{H}[s]} \hat{H}[k]$ is called under sampling regime, and any further increase in resolution will result in a loss of information on the generative process.
The question that now arises is *"Which are the samples that are the most informative on the*

generative process?". These samples can be found from the maximisation problem

$$max_{m_k} \ \hat{H}[k] + \mu \hat{H}[s] + \lambda \sum_k k \cdot m_k \tag{2.2.5}$$

where for convenience we rewrite

$$\hat{H}[s] = -\sum_k \frac{k \cdot m_k}{N} \cdot log_2\left(\frac{k}{N}\right) \tag{2.2.6}$$

The solution to this maximisation problem results in the fact that the maximally informative samples in the under sampling regime are those that have a characteristic power-law frequency distribution.

$$m_k \sim k^{-\mu-1} \tag{2.2.7}$$

which is, in itself, a signature of statistical criticality. From 2.2.7 we have that, in the under sampling regime, a decrease of Δ bits in resolution $\hat{H}[s]$ grants an increase of $\mu\Delta$ bits in relevance $\hat{H}[k]$. Consequently, we have that the parameter μ quantifies the trade off between resolution and relevance, and specifically:

- $\mu \geq 1 \rightarrow$ Lossless compression
- $\mu < 1 \rightarrow$ Lossy compression
- $\mu = 1 \rightarrow$ Zipf's law - optimal trade off between resolution and relevance

Now that we have defined both resolution and relevance, we can finally provide a formal and meaningful definition of the (statistical) metric Multiscale Relevance.

Definition 3. *Let $\hat{H}[s]$ denote the resolution associated to a sample $\mathcal{S}$ and $\hat{H}[k]$ the related relevance. Then we formally define the Multiscale Relevance (MSR) as*

$$\mathcal{R}_t = \int_{\hat{H}[s]} \hat{H}[k] \, d\hat{H}[s] \tag{2.2.8}$$

Now, it is called Multiscale Relevance because it evaluates the relevance of the considered objects across all possible classification, thus providing a metric which is independent on the choice of s^i. This comes particularly in hand, as knowledge of the resolution that best captures the unfolding of events is often unknown, and therefore it quantifies their information content in a way that abstracts from any external arbitrary choice.

2.3 MSR and Neural Spike Data

We now consider an application of *MSR* to neural spike data, provided in the form of time series $F(T)$ as defined for 1.3.2. In function of this definition, a spike train can be imagined as a graph being zero at any time except whenever the neuron fires, where instead it attains value 1.

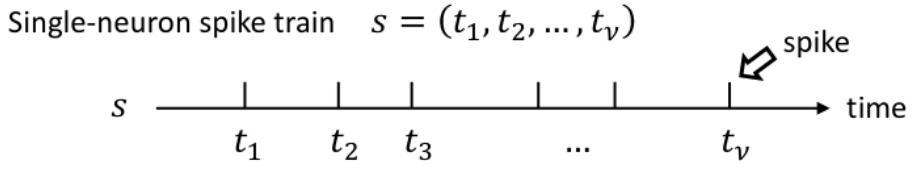


Figure 2.2: Exemplification of a Spike Train (source: [38])

Given this object, we will, in accordance with the preceding framework, impose $F(T) = \{t_1, ..., t_N\} := \mathcal{X}$ set of complex instances that admits many different classifications $\{s_i\}_i$, where T is the end time for the recording session. For simplicity, suppose that $t_1 < t_2 < .. < t_N$, where N is the cardinality of $F(T)$, and proceed with the following construction [13]

1. Define a time-step $\delta t \leq T$
2. Discretize the time interval $T = m \cdot \delta t$ into m equal size bins

3. Generate, using 1.3.2, a set of frequencies $\{k_1, .., k_m\}$:

$$k_j = \int_{j\cdot\delta t}^{(j+1)\cdot\delta t} S^T(t)\, dt \quad j \in \{1, .., m-1\} \tag{2.3.1}$$

Notice that the initial definition of δt allows us to fixate the time scale at which the spike sequence is analyzed, thus allowing us to formalize

$$s : \Delta \longrightarrow \tilde{\mathcal{S}}$$

such that

$$s^i = s(\delta t_i) = \{b_1, .., b_{m_i}\} \quad \delta t_i \in \Delta \tag{2.3.2}$$

where b_j denotes the time bin $[j \cdot \delta t_i, (j+1) \cdot \delta t_i], \quad j = 1, .., (m_i - 1)$. Adapting then the resolution metric to the present case, and normalizing in order to obtain $\hat{H}[s] \in [0, 1]$, we get

$$\begin{aligned} \hat{H}[s] &= -\frac{log(2)}{log(N)} \sum_{j=1}^{m} \frac{k_j}{N} \cdot log_2\left(\frac{k_j}{N}\right) \\ &= -\sum_{j=1}^{m} \frac{k_j}{N} \cdot log_N\left(\frac{k_j}{N}\right) \end{aligned} \tag{2.3.3}$$

and coherently with the previous theory,

- $\exists \delta t^- : \ \forall \delta t \leq \delta t^- \ \ k_j \in \{0, 1\} \Rightarrow \hat{H}[s] = 1$
- $\exists \delta t^+ : \ \forall \delta t \geq \delta t^+ \ \ k_l = N, k_j = 0 \ \forall j \neq l \Rightarrow \hat{H}[s] = 0$

For each resolution $\hat{H}[s]$ we are then interested at characterizing the dynamic response of the neuron [13]. Indeed, if for some reason $\exists l, h \in \{1, .., m\} \ \ k_l = k_h$, then the dynamic response of the neuron in the bins b_l and b_h cannot be tell apart. Consequently, one possible way to evaluate such response over a recording interval $[0, T]$ is to count the number of dynamic states at the given resolution. Specifically, defining $m_k \in \{1, .., m\}$ as the number

of bins characterized by a frequency $\frac{k}{N}$, then we define the relevance of the Spike train at the given resolution as

$$\begin{aligned}\hat{H}[k] &= -\frac{log(2)}{log(N)} \sum_{k=1}^{N} \frac{k \cdot m_k}{N} \cdot log_2\left(\frac{k \cdot m_k}{N}\right) \\ &= -\sum_{k=1}^{N} \frac{k \cdot m_k}{N} \cdot log_N\left(\frac{k \cdot m_k}{N}\right) \end{aligned} \tag{2.3.4}$$

and again, in accordance with the limit cases that we have generally discussed in the previous section, we have that

- $k = 1,\ m_k = N \Rightarrow \hat{H}[k] = 0$
- $k = N,\ m_k = 1 \Rightarrow \hat{H}[k] = 0$

Since *MSR*, in this specific case, is supposed to measure the richness in dynamic response of the neuron, it should correlate with other metrics previously defined for analogous purposes, such as the coefficient of local variation [34][33] and the burstiness and memory (statistical) metrics [19].

We now turn to an application of this adaptation of MSR to spike data, reporting results from [13] that were subsequently reproduced by myself as a proof of concept.

2.4 MSR and Applications to Real Spike Data recorded from Populations of Neurons

In the experiment run by Stensola et. al. [36], 65 functional neurons - not necessarily anatomical neurons, but point sources that emit a spike signal in a neuron like fashion (can be ensembles of neurons; recording resolution is not fine enough to discriminate) - were recorded from the medial Enthorinal cortex (mEC) of freely moving rats in a $150 \times 150\, cm^2$ box. The choice of investigating activity in the mEC has not been casual, but rather a task dependent choice, since the only activity performed by the rats in the given environment is navigation. Indeed, the medial Enthorinal Cortex serves the dynamic representation of self-location [36]

through the conglomerate working of grid-cells, which are neuron modules characterized by hexagonal neural fields that comb the navigated environment. Specifically, grid cells have several periodic firing fields arranged in an hexagonal lattice [20][29] and the coordinated activation of this cells provides the network with a spatial representation of the environment which is independent from feature extraction. The recorded network is comprehensive of several grid cells, as well as interneurons, which make up more than 99% of the nervous system and mainly serve an integration function, carrying sensory information and regulating motor activity [28]; overall, they provide relevant compounded information to modules carrying primary cognitive functions, such as memory, learning, cognition, and planning.
In their study, Cubero et al. investigated the MSR score for each of the 65 recorded neurons, generating a distribution of such values in the range $[0.25, 0.3]$, with the area difference between more and less informative neurons being clearly identifiable from the plot

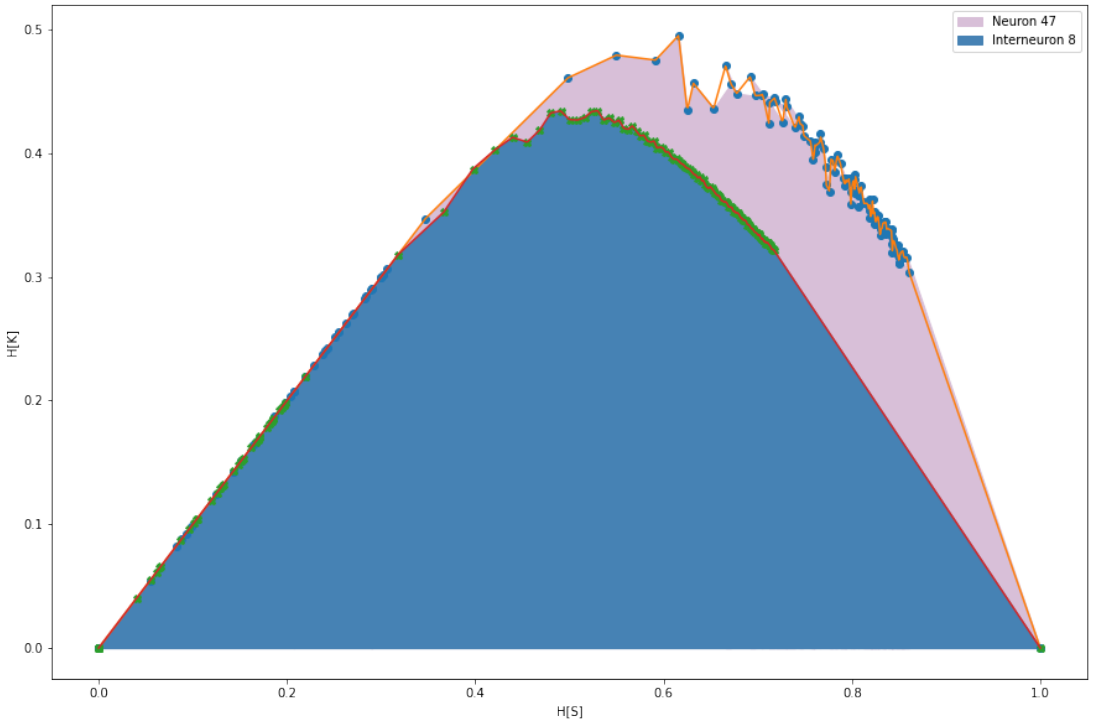


Figure 2.3: Difference in MSR area between the least and most informative neurons

It is now worth asking whether such scoring really reflects a difference in information coding relevance for the recorded neurons under the presented framework. Specifically, since the rats are performing a spatial task, and the investigated area is closely involved in spatial navigation, we would like to know whether neurons that are classified by MSR as highly informative (relative to the sample) are also those that provide information relevant for the task.

In order to answer this question, we will rely on a previously defined metric, which has general applications but, in origin, was tailored for the investigation of neural spatial relevance for signals generated by rats freely moving in an environment. The pseudo-measure in question is Skaggs spatial information[35], which is formally defined as

$$I = \int_{\Omega} \lambda(x) \cdot log_2\left(\frac{\lambda(x)}{\lambda}\right) \cdot p(x)\, dx \tag{2.4.1}$$

where $p : \Omega \longrightarrow [0, 1]$ is the probability measure over the state space Ω and $\lambda : \Omega \longrightarrow \mathbb{R}^+$ its the mean firing rate of the neuron when the agent is in state $x \in \Omega$, and

$$\lambda = \int_{\Omega} \lambda(x) \cdot p(x)\, dx \tag{2.4.2}$$

the overall mean firing rate. In the present case, x denotes the centre of each of a $7.5 \times 7.5\ \ cm^2$ bin, $\lambda(x)$ number of spikes of the considered neuron when the rat is located in the bin of centre $x \in \Omega$, and $p(x)$ the probability of being in such bin. Notice however that $x \in \Omega$ could represent any state of the agent, not necessarily position, as long as it is a stationary random variable, along with the associated spike train. Moreover, in order to make the transition of the considered quantities between different bins smoother, I have applied a Gaussian smoothing kernel to each of the bin centred state probabilities and mean firing rates. Formally, suppose that the binning procedure results in a continuum of centres - such that Ω and its binning are equal - and take a generic centre $\bar{x} \in \Omega$, and consider the generic

map ϕ (either state probability or mean firing rate); then the smoothed map will be

$$\begin{aligned}\tilde{\phi}(\bar{x}) &= \phi(\#) * \mathcal{G}(\#, \sigma)(\bar{x}) \\ &= \int_{\Omega} \phi(x) \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{||\bar{x}-x||}{2\sigma^2}} \, dx \end{aligned} \tag{2.4.3}$$

where, in the present case, $\sigma = \frac{1}{r_n^T(\bar{x})\frac{1}{12}}$ for $r_n^T(\bar{x})$ spatially dependent firing rate defined according to 1.3.7. Following this introduction on an auxiliary information pseudo-measure, I will present the receptive fields of two high and two low MSR scoring neurons.

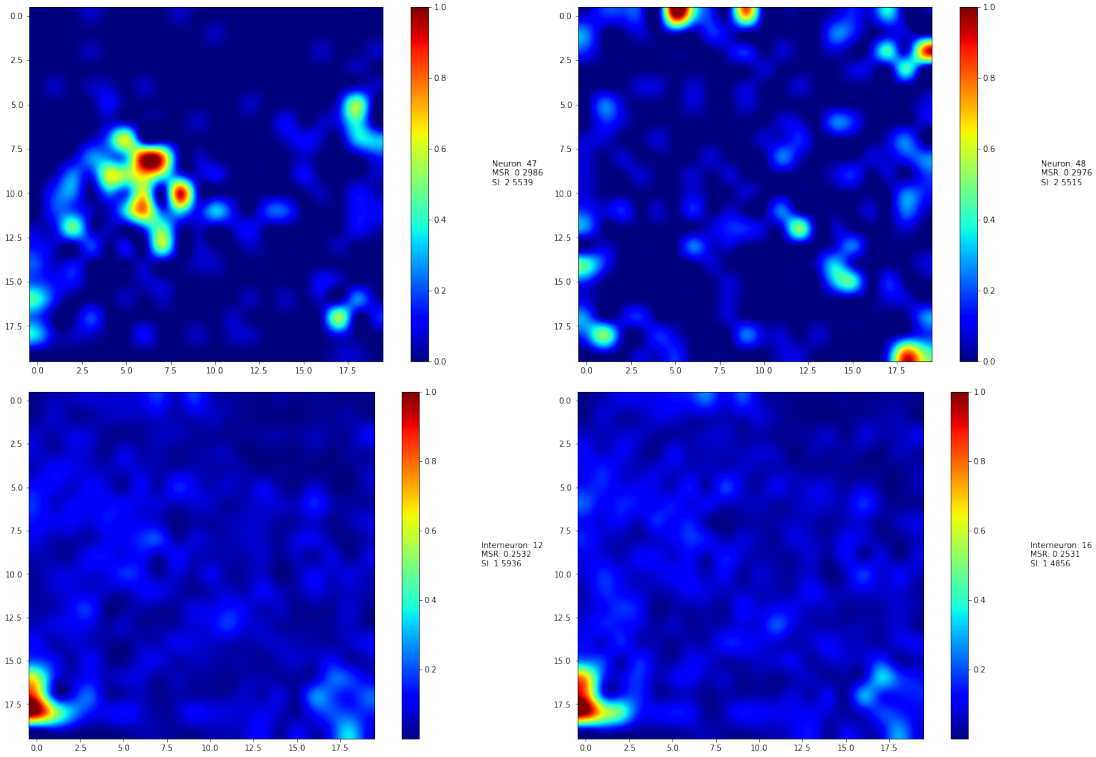


Figure 2.4: The top row presents the receptive fields of two high MSR scoring neurons, and as can be seen they display a strongly selective firing. The bottom row presents the receptive fields of two low MSr scoring neurons, and as can be seen, they are very noisy.

As can be readily seen from this visual representation and from the descriptive tables on the right of each figure, in this subset of neurons those associated with a relatively high MSR values are also associated with relatively high spatial information values. Furthermore, as

can be observed from 2.4, the receptive fields of high MSR neurons seem to be way more localized than those of low scoring ones, which presents substantial background noise. To further corroborate the hypothesis of the article by Cubero et al. [13], and matching their line of work, I have also computed the spatial information content associated to the head direction of the mouse, through a transposition of the space in spherical coordinates.

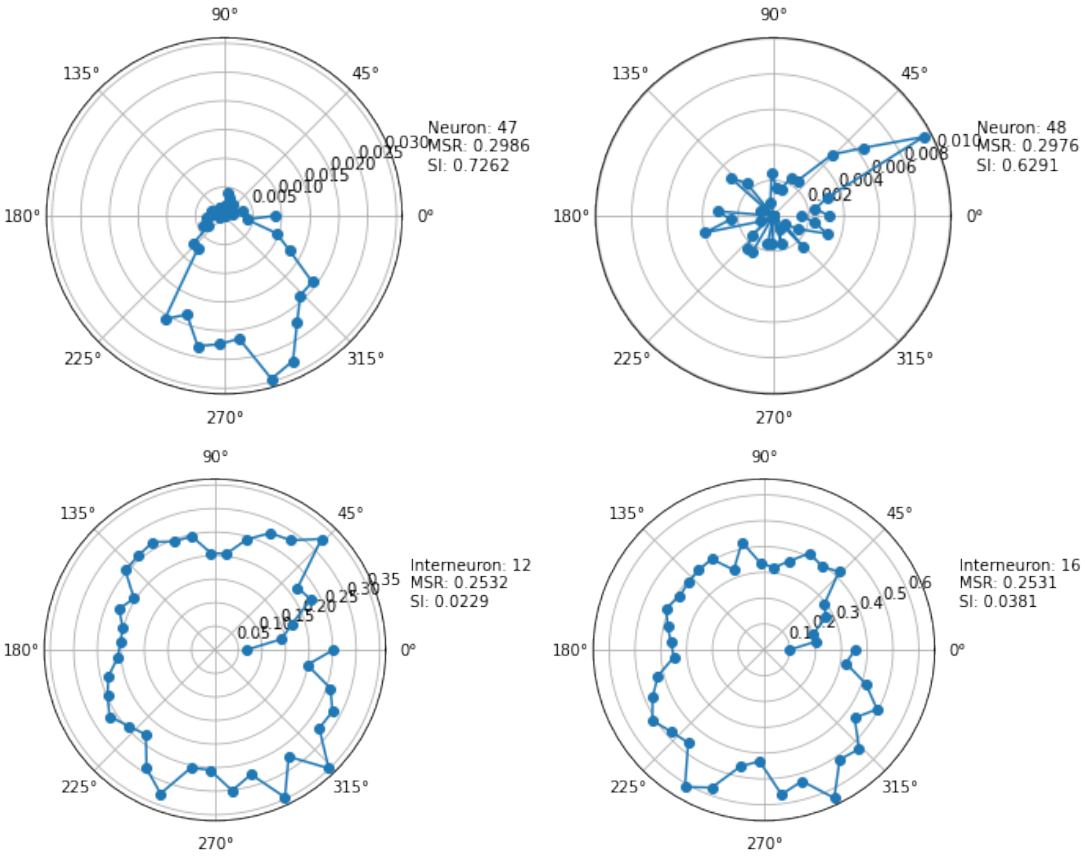


Figure 2.5: The top row presents the HD receptive fields of high MSR scoring neurons, which are constrained to a narrow interval of angular values. The bottom row presents the HD receptive fields of low MSR scoring neurons, which span almost the entire interval $[0, 360]^\circ$.

Again, we can see how the spatial information value associated to low MSR scoring neurons is significantly smaller (one order of magnitude smaller) than that of high MSR scoring neurons, and the associated receptive fields are much less localized, highlighting an almost invariance of those neurons selectivity with respect to head direction.

The question now remains of whether this difference is just a spurious correlation observed in an ad hoc sub-sample of neurons, or rather a manifestation of a more general correlation holding over the entire set of neurons. As visible from the subsequent graphs, we are in the latter case, with a significant correlation of MSR with both spatial and head direction Skaggs information.

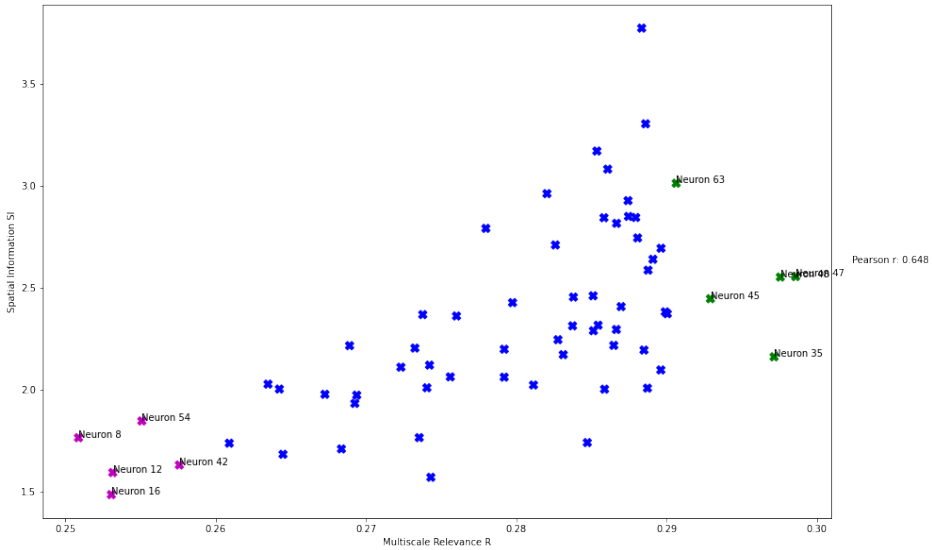


Figure 2.6: Correlation for MSR and Spatial Information

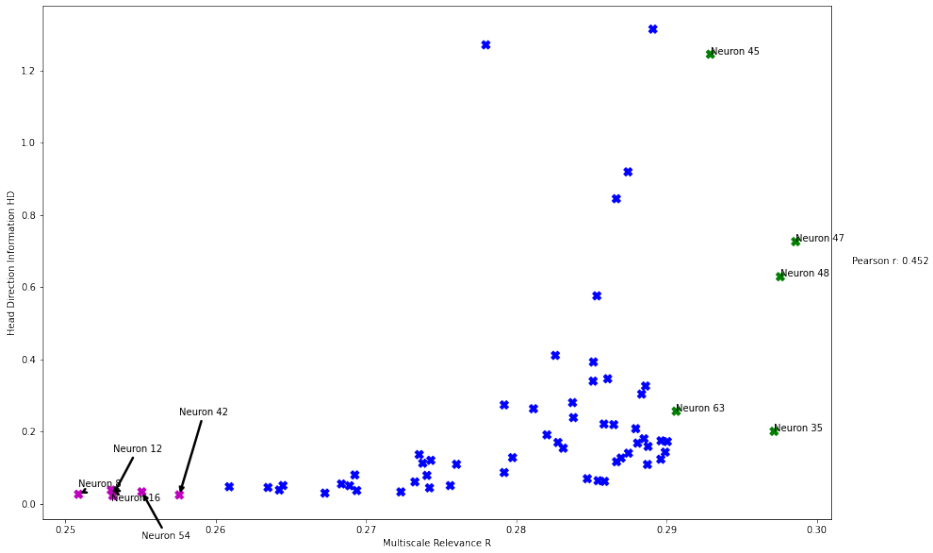


Figure 2.7: Correlation for MSR and Head Direction Information

2.5 Other Metrics of Variability in Spike Trains

Throughout the years, other metrics have been developed to characterize the richness of response of neuronal signals, or more in general, of signals from which inter-event times series can be extracted. Such metrics may provide a relevant starting point to understand which aspects of the signals the adapted MSR metric is able to capture, as well as if it characterizes degeneracy along some dimensions of variability.

Two such metrics are the burstiness and memory coefficients, first introduced by Goh and Barabasi [19], which jointly aim at a full definition of the irregularities observed in bursty patterns, as they intently target two complementary sources of randomness. The former, the burstiness coefficient, is presented as a distribution-based measure, as it directly addresses the variability of the inter-events time distribution, and it is formally defined as

$$B = \frac{\frac{\sigma_\tau}{m_\tau} - 1}{\frac{\sigma_\tau}{m_\tau} + 1} \equiv \frac{\sigma_\tau - m_\tau}{\sigma_\tau + m_\tau} \tag{2.5.1}$$

It is a bounded metric $B \in (-1, 1)$, and the two extremes provide a very insightful picture on the nature of the signal, as

- $B \to -1 \Rightarrow$ Regular Pattern ($\sigma_\tau \to 0$)
- $B \to 1 \Rightarrow$ Most random pattern ($\sigma_\tau \to +\infty$)

The latter, the memory coefficient, is presented as a correlation-based measure, as it directly addresses the degree of correlation in successive inter-event times, and is formally defined as

$$M = \frac{1}{n_\tau - 1} \sum_{i=1}^{n_\tau - 1} \frac{(\tau_i - m_1)(\tau_{i+1} - m_2)}{\sigma_1 \sigma_2} \tag{2.5.2}$$

where n_τ is the number of inter-event intervals extrapolated from the spike train, m_1 (m_2) and σ_1 (σ_2) are the mean and variance of the series $\{\tau_i\}$ ($\{\tau_{i+1}\}$) for $i = 1, .., n_\tau - 1$, respectively. As in the case of the B coefficient, we have $M \in (-1, 1)$, and of particular

relevance for the understanding of the properties of the signal is the meaning of the two range halves

- $M \ll 0 \Rightarrow$ Short (long) intervals likely to be followed by long (short) ones
- $M \gg 0 \Rightarrow$ Short (long) intervals likely to be followed by short (long) ones
- $M \approx 0 \Rightarrow$ No correlation

Yet another metric assessing the variability of inter-events times has been introduced by Shinomoto et. al. [34][33], and it is the coefficient of local variation, formally defined as

$$L_V = \frac{3}{n_\tau - 1} \sum_{i=1}^{n_\tau - 1} \frac{(\tau_{i+1} - \tau_i)^2}{(\tau_{i+1} + \tau_i)^2} \tag{2.5.3}$$

Differently from the global coefficient of variation C_V, which detects the global variability of the ISI sequence, the coefficient of local variation L_V detects step-wise variability in the firing of a neuron, or more generally the bursting of a process; nonetheless, and coherently with C_V, it is expected to take value 1 for a sufficiently long, Poisson distributed ISI.

These new metrics should allow us to thoroughly characterize the aspects of bursty signals variability that MSR is able to capture, as well as those that are not detected. Indeed, such comparison has been undertaken by Cubero et. al. on the data drawn from Stensola [36], providing initial evidence for MSR ranges where the metrics overlap or differ. Despite the encouraging results, studying experimentally the property of MSR in relation to cortical activity remains as much a challenge as an unexplored field, due to the present ethical and technological limitations in the acquisition of neural data. For this reason, it is necessary to spearhead the experimental inquiry of such properties through the consideration of simulated neuronal networks.

3

A Simple Model of the Cortex exhibiting Chaotic Activity

The importance of simulating large scale neuronal populations for the study of their macroscopic, emerging properties, requires a careful tailoring of the trade-off between the amount of details considered and the relevance of such computational effort for collective analysis. Therefore, it is necessary to define the minimum amount of details that our simulation should incorporate in order to produce results of interest. Specifically, we will need

- An excitatory population of neurons
- An inhibitory population of neurons
- A rule for the coupling of units
- An activation function determining the state of a unit in the network

3.1 Interaction of Homogeneous Populations

One of the most daunting problem in neural populations neuroscience has been to understand the nature of the strong temporal fluctuations characterizing neuronal spiking. Indeed, it has been observed, in animals, how the ISIs are characterized by broad distributions, resembling the one that could be generated by a Poisson process [1][9][14], and how the membrane potential itself displays strong temporal fluctuations. A first insight on the underlying mechanism comes from experimental studies of constant current injection in in-vitro neurons; in this case, we reliably observe a regular firing pattern. Thus, it has been accordingly hypothesized that the irregularity observed in real neuronal populations must be a consequence of the massive and heterogeneous synaptic input [27][21]. One of the leading hypothesis that aims at explaining such irregularity of firing is that the spatial and temporal integration of synaptic inputs coming from excitatory and inhibitory neurons tends to balance near the receiving neuron firing threshold, so that even slight fluctuations in such input prove sufficient to determine either the presence or absence of firing. This hypothesis implicitly implies that the inherent stochasticity in neuronal firing is a result of the deterministic summation of a large number of inputs. Furthermore, the high sensitivity to fluctuations should determine a general network state characterized by highly chaotic activity and weak cross-correlation between neurons' firing. We will now proceed with a complete description of the model that satisfies all of the above requirements.

Let P_E be the population of excitatory neurons, characterized by N_E units, and P_I the population of inhibitory neurons, with N_I units respectively, and define a population specific external input

$$I(t) = I_E(t)\mathbb{1}(P_E) + I_I(t)\mathbb{1}(P_I) \tag{3.1.1}$$

which, generally, may be either of deterministic or stochastic nature. Accordingly, define a population specific threshold for spike firing as

$$\theta(t) = \theta_E(t)\mathbb{1}(P_E) + \theta_I(t)\mathbb{1}(P_I) \tag{3.1.2}$$

where the time dependence considers the most general case of externally modulated threshold. Consider now between-populations specific coupling J_{kl} [42], for $k, l \in \{E, I\}$ that are constant in time, and let $K \ll min(N_E, N_I)$ be the average number of coupling afferring from the population l to each unit of the population k. Then the total synaptic strength of an active connection will be

$$J_{kl}^{ij} = \frac{J_{kl}}{\sqrt{K}} \tag{3.1.3}$$

and since we want our network to be sparse - biological networks are far from full connectivity - we additionally impose

$$P(J_{kl}^{ij} \neq 0) = \frac{K}{N_k} \tag{3.1.4}$$

Thereby, we can define the total synaptic input to unit i of population k as

$$u_k^i(t) = \sum_{l \in E,I} \sum_{j=1}^{N_l} J_{kl}^{ij} \sigma_l^j(t) + I_k(t) - \theta_k(t) \tag{3.1.5}$$

where σ is the state of a single unit, trivially defined as

$$\sigma_k^i(t) = \mathcal{H}(u_k^i(t)) \tag{3.1.6}$$

where $\mathcal{H} : \mathbb{R} \longrightarrow [0, 1]$ is the Heaviside step function. Of particular importance to the network dynamics is the asynchronous update of units, which is operated through the random choice of units from each population.

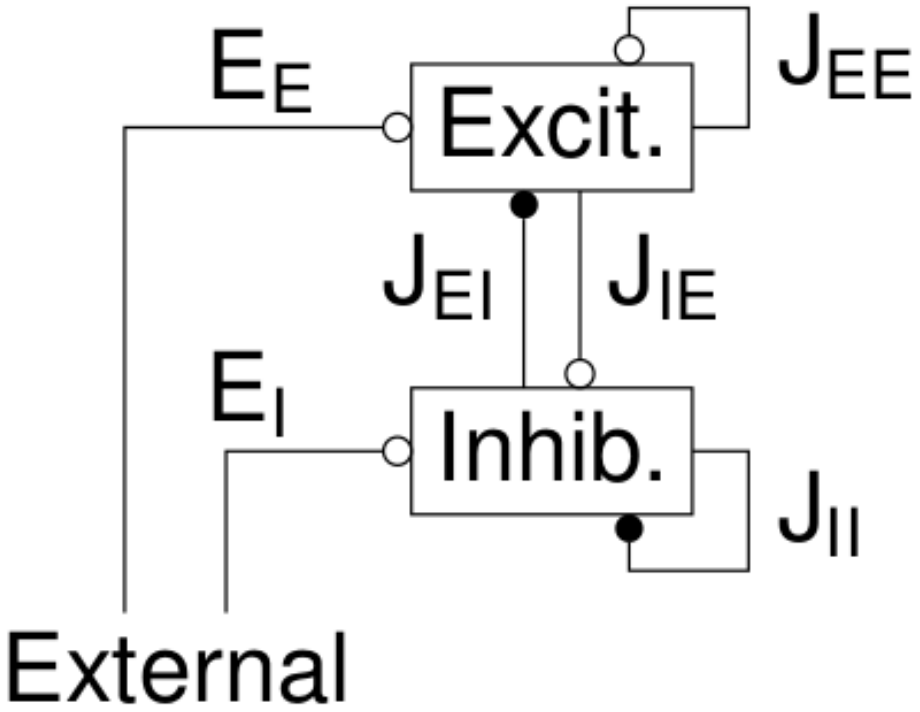


Figure 3.1: Visualization of heterogeneous population model (source: [42])

The central assumptions of our model are

1. The total excitatory and inhibitory inputs are large if compared the the unit threshold θ_k

2. The external input is large compared to the threshold θ_k and may be expressed as

$$I_k = \xi_k m_0 \sqrt{K} \quad \xi_k \sim \mathcal{O}(1),\ m_0 \in (0, 1) \tag{3.1.7}$$

In the presented case, the authors [42] have chosen to set both the external input $I_k(t) \equiv I_k$ and $\theta_k(t) \equiv \theta_k$ constant in time. Moreover, for simplicity they have also set

$$J_{EE} = J_{IE} = 1$$

In order to proceed with the mean field results for our network, we must first appropriately define the limit $N_E + N_I = N \longrightarrow +\infty$. Specifically, we require that $\frac{N_E}{N_I}$ remains fixed as $N \longrightarrow +\infty$, and that, in order to assess results at different degree of sparsity, we first take the limit $N \longrightarrow +\infty$ and only later $K \longrightarrow +\infty$; by doing so, we systematically avoid to incur in full connectivity. Thus, formally, for any network property $\gamma(N, K)$ we consider

$$\lim_{K \longrightarrow +\infty} \lim_{N \longrightarrow +\infty} \gamma(N, K)$$

From 3.1.6 we can obtain the population-averaged firing rate of a generic population k as

$$\begin{aligned} m_k(t) &= \frac{1}{N_k} \sum_{i=1}^{N_k} \sigma_k^i(t) \\ &= [\sigma_k^i(t)] \end{aligned} \tag{3.2.1}$$

which will be the starting point of our mean field approximation. With a slight change of perspective, consider now

$$m_k^i(t) = \langle \sigma_k^i(t) \rangle \tag{3.2.2}$$

which is the average of the activity of unit i in population k at time t across different initial conditions. Ginzburg and Sompolinsky (1994) [18] have proven that mean activity rates obey the following dynamics

$$\tau_k \frac{d}{dt} m_k^i(t) = -m_k^i(t) + \mathcal{H}(u_k^i(t)) \tag{3.2.3}$$

We now add a further simplification to our model, supposing that inputs from different units are uncorrelated for $K \ll log(N)$, by considering $n_E(t)$ and $n_I(t)$ the total excitatory and inhibitory inputs to a generic unit, so that we effectively write

$$u_k(t) = J_{k0} m_0 \sqrt{K} + \frac{J_{kE}}{\sqrt{K}} n_E(t) + \frac{J_{kI}}{\sqrt{K}} n_I(t) - \theta_k \tag{3.2.4}$$

and under

$$F_k(m_E(t), m_I(t)) = \sum_{n_E, n_I = 0}^{+\infty} p(n_E) p(n_I) \mathcal{H}(u_k(t)) \tag{3.2.5}$$

probability that the updating unit is activated at time t, with $p(n_l)$ probability of receiving n_l inputs from population l, we finally obtain the following dynamics

$$\begin{aligned} \tau_k \frac{d}{dt} m_k(t) &= [\tau_k \frac{d}{dt} m_k^i(t)] \\ &= -m_k(t) + F_k(m_E(t), m_I(t)) \end{aligned} \tag{3.2.6}$$

In the large N limit, it can be shown that

$$\begin{aligned} p(n_l) &= \sum_{s=n_l}^{+\infty} \frac{K^s \cdot e^{-K}}{s!} \binom{s}{n_l} r_l^{n_l} (1 - r_l)^{s - n_l} \\ &= \frac{(r_l K)^{n_l}}{n_l!} \cdot e^{-r_l K} \end{aligned} \tag{3.2.7}$$

where $r_l = \frac{K}{N_l}$ is the probability for a synapse to be active within population l. Most notably, since $p(n_l)$ is Poisson, we will have

$$\begin{aligned} \mathbb{E}[n_l] &= \mathbb{V}[n_l] \\ &= m_l K \end{aligned} \tag{3.2.8}$$

and consequently

$$\begin{aligned}\mathbb{V}[u_k] &= \mathbb{V}\left[\frac{J_{k0}m_0K + J_{kE}n_E + J_{kI}n_I}{\sqrt{K}} - \theta_k\right] \\ &= \frac{1}{K}\mathbb{V}[J_{kE}n_E + J_{kI}n_I] \\ &= \frac{1}{K}\Big(J_{kE}^2\mathbb{V}[n_E] + J_{kI}^2\mathbb{V}[n_I] + 2J_{kE}J_{kI}CoV(n_E, n_I)\Big) \\ &= \frac{1}{K}(J_{kE}^2 m_E K + J_{kI}^2 m_I K) \\ &= J_{kE}^2 m_E + J_{kI}^2 m_I \end{aligned} \tag{3.2.9}$$

where between the third and fourth equation we have exploited the uncorrelation of the inputs. Thus, in the limit $K \longrightarrow +\infty$

$$\begin{aligned}F_k(m_E, m_I) &= \int_{\mathbb{R}} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\mathcal{H}(u_k + \sqrt{\alpha_k}x)\, dx \\ &= H\Big(-\frac{u_k}{\sqrt{\alpha_k}}\Big)\end{aligned} \tag{3.2.10}$$

where $H()$ is the complementary error function. Thus, we have proven that under the mean field approximation

$$\tau_k \frac{d}{dt}m_k(t) = -m_k(t) + H\Big(-\frac{u_k}{\sqrt{\alpha_k}}\Big) \tag{3.2.11}$$

But if, as by assumption, the external input is constant, then so should be the average activity, and consequently

$$m_k = H\Big(-\frac{u_k}{\sqrt{\alpha_k}}\Big) \tag{3.2.12}$$

which reflects the fact that the input to each unit fluctuates across the entire population, with the fluctuation obeying a Normal statistics.

Using the defined parameters, we reach the following expression for the generic population

input

$$u_E = (Em_0 + m_E - J_E m_I)\sqrt{K} - \theta_E \tag{3.2.13}$$

$$u_I = (Im_0 + m_E - J_I m_I)\sqrt{K} - \theta_I \tag{3.2.14}$$

$$J_{EI} = -J_E$$

$$J_{II} = -J_I$$

where with the last two equations we explicit the inhibitory nature of the connections exiting a unit from the inhibitory population. Given this full characterization, we are now interested at understanding the network mean response in the balance state, which is, a state in which neither of the two populations has absent or saturated activity $\Rightarrow m_k \in (0, 1)$.
In a balanced state, the order of the temporal fluctuations is the same of the difference between the network mean response and the population threshold. Furthermore, if we assume that as $N \longrightarrow +\infty$ the generic population input remains finite - hence the balance - then from 3.2.13 and 3.2.14 we have that for $K \gg 1$

$$Em_0 + m_E - J_E m_I = \mathcal{O}\Big(\frac{1}{\sqrt{K}}\Big) \tag{3.2.15}$$

$$Im_0 + m_E - J_I m_I = \mathcal{O}\Big(\frac{1}{\sqrt{K}}\Big) \tag{3.2.16}$$

which in the large K limit reduces to

$$Em_0 + m_E - J_E m_I = 0 \tag{3.2.17}$$

$$Im_0 + m_E - J_I m_I = 0 \tag{3.2.18}$$

Solving the system for m_E and m_I leads to the following interesting results [42]

$$m_E = \frac{J_I E - J_E I}{J_E - J_I} m_0 \equiv \Phi_E m_0 \tag{3.2.19}$$

$$m_I = \frac{E - I}{J_E - J_I} m_0 \equiv \Phi_I m_0 \tag{3.2.20}$$

This result tells us that under a balanced chaotic regime the network responds linearly to its mean external input, albeit the activity of each individual unit is highly non-linear. This supports the original assumption that the input from the two populations balances out and only fluctuations contribute to determine the state of individual units. Expanding now on the requirement for this system to be well defined, we have $m_k \in (0,1) \Rightarrow \Phi_E, \Phi_I > 0$. Thus

$$\frac{E}{I} > \frac{J_E}{J_I} > 1 \tag{3.2.21}$$

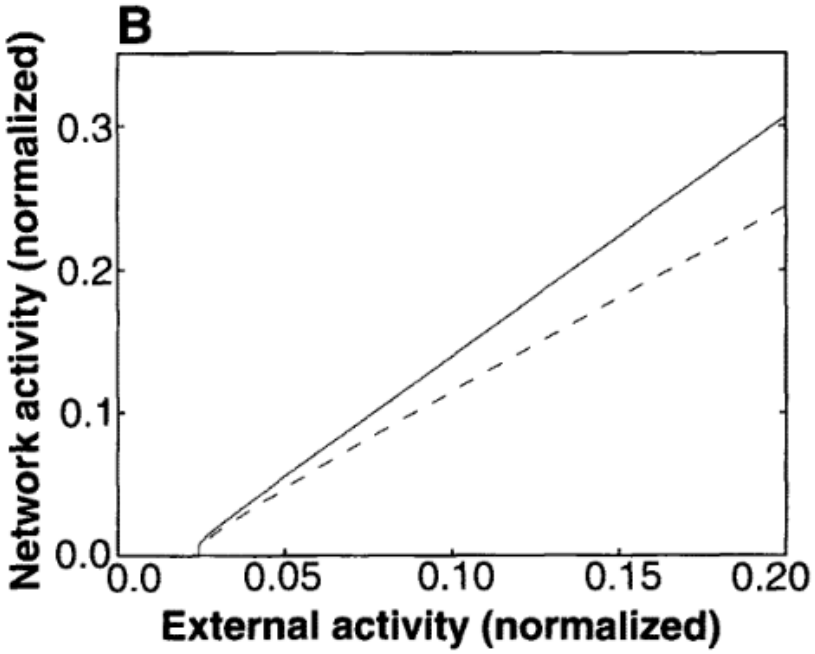
or

$$\frac{E}{I} < \frac{J_E}{J_I} < 1 \tag{3.2.22}$$

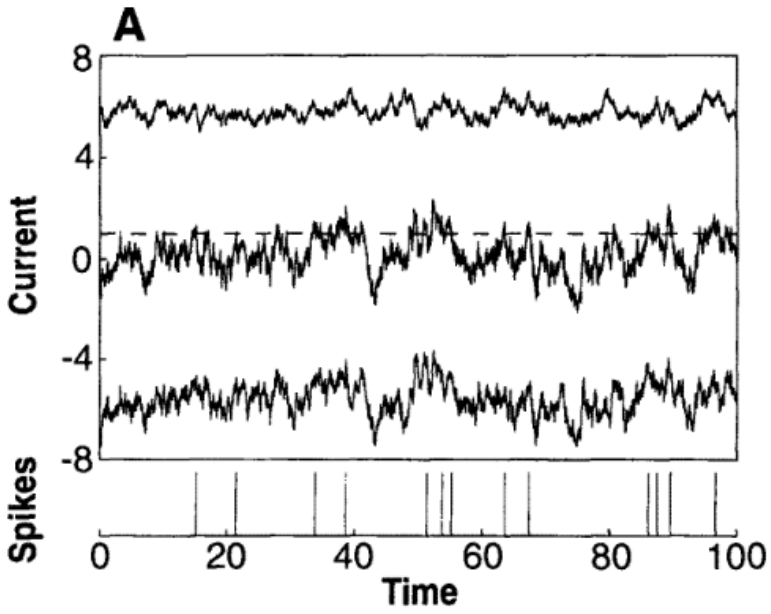
The further requirement of absence of stationary solutions $m_k = 0, 1$ excludes the second positivity condition and leaves us with the defining parameters

$$\boxed{\begin{array}{c} \frac{E}{I} > \frac{J_E}{J_I} > 1 \\ J_E > 1 \end{array}} \tag{3.2.23}$$

Notice however that for finite K the terms of order $\frac{1}{\sqrt{K}}$ are not negligible, and account for the deviations from the results predicted by 3.2.19 and 3.2.20.



(a) Mean Activity of the Network



(b) Total input to a single unit

Figure 3.2: Results for $E = 1, I = 0.8, J_E = 2, J_I = 1.8, \theta_E = 1$, and $\theta_I = 0.8$ (source: [41])

3.3 Higher Order Moments of Network Quantities

The study of the network mean response to the input provides only a partial description of the network dynamics, as it totally neglects the nature of the fluctuations that are characteristics of the balanced chaotic state. These fluctuations are necessary, since the only other way to obtain the linear response would be to have a "frozen" (fixed) subset of units continuously firing, which in turn will contradict the chaotic nature of the network states - that is, it would determine a stationary state.

Drawing from the theory of spin glasses, we then define another, second order, parameter for the network

$$q_k = \frac{1}{N_k} \sum_{i=1}^{N_k} (m_k^i)^2 \tag{3.3.1}$$

where, generalizing to the continuum, we set

$$m_k^i = \langle \sigma_k^i(t) \rangle_t = \frac{1}{T} \int_0^T \sigma_k^i(t)\, dt \tag{3.3.2}$$

3.3.1 Spatial Fluctuations

Consider now the spatial inhomogeneities in the inputs. Since K only defines the average connectivity of each unit in the network, different units may be characterized by significantly different number of connections; indeed, we have already seen that the "synaptic" distribution is Poisson distributed in the large N limit. Therefore, this variability in connectivity characterizes the first type of spatial inhomogenity, and can be quantified as

$$\bar{\delta} \langle u_k^i \rangle = \sum_{l \in E,I} \sum_{j=1}^{N_l} \delta J_{kl}^{ij} [m_l^j] \tag{3.3.3}$$

Additionally, even if we were to impose that $p(s) = \delta(K - s)$, the identities of the units projecting to any other given unit would be totally heterogeneous across the network, thus de-

termining self-consistently developed spatial inhomogeneities. These can be quantitatively defined through

$$\tilde{\delta}\langle u_k^i \rangle = \sum_{l \in E,I} \sum_{j=1}^{N_l} [J_{kl}^{ij}] \delta m_l^j \tag{3.3.4}$$

where we define, for a generic field ψ, $\delta\psi = \psi - [\psi]$. Adding both the components and considering the related variance estimator for the network configuration, we obtain

$$\begin{aligned} [(\delta\langle u_k^i \rangle)^2] &= \sum_{l \in E,I} J_{kl}^2 q_l = q_E + J_{kI}^2 q_I \\ &= \beta_k \end{aligned} \tag{3.3.5}$$

3.3.2 Temporal Fluctuations

Consider now the temporal deviation in the input to a given unit at a specific time t

$$\begin{aligned} \hat{\delta} u_k^i(t) &= u_k^i(t) - u_k^i \\ &= \sum_{l \in E,I} \sum_{j=1}^{N_l} J_{kl}^{ij} (\sigma_l^j(t) - m_l^j) \end{aligned} \tag{3.3.6}$$

Then the temporal fluctuation bearing on population k at time t can be fully expressed as

$$\begin{aligned} [(\hat{\delta} u_k^i)^2] &= \sum_{l \in E,I} J_{kl}^2 (m_l - q_l) \\ &= \alpha_k - \beta_k \end{aligned} \tag{3.3.7}$$

It then follows that for $x_i, y_i(t) \sim \mathcal{N}(0,1)$ independent random variables, we can express the state of unit i in population k as a function of the generic net input, the spatial and temporal fluctuations

$$\sigma_k^i(t) = \mathcal{H}(u_k + \sqrt{\beta_k} \cdot x_i + \sqrt{\alpha_k - \beta_k} \cdot y_i(t)) \tag{3.3.8}$$

Expanding now on the order parameter q_k, we first consider the rate distribution for population k, formally given by

$$p_k(m) = \frac{1}{N_K} \sum_{i=1}^{N_k} \delta(m - m_k^i) \tag{3.3.9}$$

and then exploit the bound $(m_k)^2 \leq q_k \leq m_k$. Qualitatively, the order parameter q_k quantifies the amount of coherence in the rate distribution, so that if

- $q_k \longrightarrow m_k$ then the system tends to be in a "frozen" state, and the homogeneity in the rate distribution is minimal.

- $q_k \longrightarrow (m_k)^2$ the system tends to be in a random chaotic state, and the homogeneity in the rate distribution is maximal.

We finally want to characterize the autocorrelation of the temporal fluctuations for each single unit. In order to do so, we define a time-dependent order parameter

$$q_k(\tau) = \frac{1}{N_k} \sum_{i=1}^{N_k} \langle \sigma_k^i(t) \sigma_k^i(t+\tau) \rangle_t \tag{3.3.10}$$

and specifically, we are interested at such order parameter in the two limiting cases, that is, when the interval shrinks to zero and when it extends to $+\infty$. For $\tau = 0$

$$\begin{aligned} q_k(0) &= \frac{1}{N_k} \sum_{i=1}^{N_k} \langle (\sigma_k^i(t))^2 \rangle_t \\ &\underset{\sigma \in 0,1}{=} \frac{1}{N_k} \sum_{i=1}^{N_k} m_k^i = m_k \end{aligned} \tag{3.3.11}$$

and for $\tau \longrightarrow +\infty$

$$\begin{aligned}\lim_{\tau\to+\infty} q_k(\tau) &= \lim_{\tau\to+\infty} \frac{1}{N_k}\sum_{i=1}^{N_k}\langle\sigma_k^i(t)\sigma_k^i(t+\tau)\rangle_t \\ &\underset{\text{ind.}}{=} \lim_{\tau\to+\infty} \frac{1}{N_k}\sum_{i}^{N_k}\langle\sigma_k^i(t)\rangle_t\langle\sigma_k^i(t+\tau)\rangle_t \\ &\approx \frac{1}{N_k}\sum_{i=1}^{N_k}(m_k^i)^2 = q_k \end{aligned} \tag{3.3.12}$$

Then if we characterize the autocorrelation of the temporal fluctuations for the population k as

$$\begin{aligned}\beta_k(\tau) &= [\langle\delta u_k^i(t)\delta u_k^i(t+\tau)\rangle_t] \\ &= q_E(\tau) + J_k^2 q_I(\tau)\end{aligned} \tag{3.3.13}$$

it is immediate to observe that

- $\beta_k(\tau) \xrightarrow[\tau\to 0]{} \alpha_k$
- $\beta_k(\tau) \xrightarrow[\tau\to+\infty]{} \beta_k$

and the new order parameter obeys the following dynamics [42]

$$\tau_k\frac{d}{d\tau}q_k(\tau) = -q_k(\tau) + \int_0^{+\infty}\frac{e^{-\frac{t}{\tau_k}}}{\tau_k}\int_{\mathbb{R}}\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\left[H\Big(\frac{-u_k - \sqrt{\beta_k(t+\tau)}\cdot x}{\sqrt{\alpha_k - \beta_k(t+\tau)}}\Big)\right]^2 dx\,dt$$

It is possible to prove that the balance state is stable and that any perturbation $\delta:\ \mathcal{O}(\delta) \ll \mathcal{O}(\frac{1}{\sqrt{K}})$ is smoothed out by the collective nature of the input. Notice however that for very large networks, even perturbations $\delta \propto \frac{1}{\sqrt{K}}$ are very small; consequently, since perturbations of this order drive the system towards a new mean state, we conclude that the network is extremely sensitive to small variations in its input. Thus, the system is well endowed to track subtle changes in its environment on a very small time scale.

A clear problem of the model outlined so far is the significant structural rigidity of the network, where the weights of each population are frozen both in time and in space. The introduction of this simplification has proved crucial for the derivation of the results on the mean linear response and the study of spatial and temporal fluctuations with respect to that base value. However, it is very unlikely for real neuronal networks to be characterized by population-frozen functional connectivity; instead, it is probable that the strength of the coupling between any two neurons is distributed along a continuum. What is the actual range of this continuum still remains, if addressable at all, an open question, since at the moment we still do not have a unique rigorous definition of functional connectivity. Indeed, different quantitative definitions, not necessarily in contradiction with one another, may provide different qualitative interpretations. Proceeding then one step at the time, it would be interesting to define a cortical model like the previous one but where couplings are drawn from a distribution that allows for the same type of population identification.
Consider then a generic probability distribution $\rho : I \to [0, 1]$ defined on a probability space $(I, \mathcal{F}, \rho)$, where $\mathcal{F}$ is the σ-algebra defined over I. Since we want to identify in a unique way excitatory and inhibitory units, we must impose $I \subseteq \mathbb{R}^+$ and define

- $J_{kE} = \mathcal{J}$ with $\mathbb{P}(J_{kE} = \mathcal{J}) = \rho(\mathcal{J})$
- $J_{kI} = -\mathcal{J}$ with $\mathbb{P}(J_{kE} = \mathcal{J}) = \rho(\mathcal{J})$

with $\mathcal{J} > 0$ for the given quantities to be well defined. In addition, since it would be ideal to be able to compute the order parameters defined for the previous network, we must require that

- $\exists \mathbb{E}_\rho[x] < +\infty \quad x \in I$
- $\exists \mathbb{E}_\rho[x^2] < +\infty \quad x \in I$

From Appendix B of [18] we see that the derivation of 3.2.3 does not depend on a particular choice of the connection architecture, so that the results seen so far, at least for the

linear response, should hold also in this new case. Thus, since we are directly introducing stochasticity in the network, when deriving the system associated to the linear response we need to take expected values with respect to the weights distribution. Thereby, without any specific assumption on the weights, we could write the generic population specific expected input, function of the expected weights, as

$$u_k = (\xi_k m_0 + \mathbb{E}_\rho[J_{kE}]m_E + \mathbb{E}_\rho[J_{kI}]m_I)\sqrt{K} - \theta_k \tag{3.4.1}$$

Taking then, in the following orders, the limits $N \to +\infty$ and $K \to +\infty$ we obtain a new system

$$\boxed{\begin{aligned} Em_0 + \mathbb{E}_\rho[J_{EE}]m_E - \mathbb{E}_\rho[J_{EI}]m_I = 0 \\ Im_0 + \mathbb{E}_\rho[J_{IE}]m_E - \mathbb{E}_\rho[J_{II}]m_I = 0 \end{aligned}} \tag{3.4.2}$$

Solving again for m_E and m_I, we obtain the following two expressions for the expected linear response of the network populations to an external input

$$\begin{aligned} m_E &= \frac{E\mathbb{E}_\rho[J_{II}] - I\mathbb{E}_\rho[J_{EI}]}{\mathbb{E}_\rho[J_{EI}]\mathbb{E}_\rho[J_{IE}] - \mathbb{E}_\rho[J_{EE}]\mathbb{E}_\rho[J_{II}]} \cdot m_0 \\ &= \Gamma_E m_0 \end{aligned} \tag{3.4.3}$$

and

$$\begin{aligned} m_I &= \frac{E\mathbb{E}_\rho[J_{IE}] - I\mathbb{E}_\rho[J_{EE}]}{\mathbb{E}_\rho[J_{EI}]\mathbb{E}_\rho[J_{IE}] - \mathbb{E}_\rho[J_{EE}]\mathbb{E}_\rho[J_{II}]} \cdot m_0 \\ &= \Gamma_I m_0 \end{aligned} \tag{3.4.4}$$

Notice that, analogously to the case with fixed weights, in order for the linear response to be well defined, we need to impose the following constraints: $\Gamma_E > 0$ and $\Gamma_I > 0$. This in

turn results in the following conditions

$$\frac{E}{I} < \frac{\mathbb{E}_\rho[J_{EI}]}{\mathbb{E}_\rho[J_{II}]} < \frac{\mathbb{E}_\rho[J_{EE}]}{\mathbb{E}_\rho[J_{IE}]} \tag{3.4.5}$$

or

$$\frac{E}{I} > \frac{\mathbb{E}_\rho[J_{EI}]}{\mathbb{E}_\rho[J_{II}]} > \frac{\mathbb{E}_\rho[J_{EE}]}{\mathbb{E}_\rho[J_{IE}]} \tag{3.4.6}$$

Further elaboration on which of the two conditions results, from a theoretical point of view, in a balanced chaotic state eludes the purpose of the present thesis and, for this reason, no further comments on how this couplings initialization affects dynamics will be made. Nonetheless, there are other network properties that could contribute to the creation of a balanced chaotic state and that provide a more realistic description of cortical networks.

As it stands, the networks considered so far rely on a very simplistic network topology - in fact, the simplest possible - since it instantaneously assign to each unit a fixed number of connections, on average, and the topological identity of the different units is not taken into consideration. Specifically, negligence towards topological identity means, in the present context, lack of proximity dependent wiring rules. This choice is in stark contrast with what seems to be the prevailing wiring principle in the brain. Indeed, neurons seem to form a tight bundle of connections with other neurons in a close neighbourhood, with a smaller fraction of longer ranging ones. This particular type of connectivity may incentivize the emergence of processing cliques, that is, clusters of highly dependent activity. Therefore, it would be advisable to study the impact that the choice of such a network topology will have on the network activity, and if and how this differ from the one generated by the original model.

Yet another factor that is not taken into account is that the formation of neuronal networks is not a stationary process, but rather a dynamic one characterized by generation of new units and rewiring of older ones according to both topological and chronological factors. Consequently, it would be interesting to study the mean response of a network that has undergone a dynamic generative process, to draw relevant conclusions on the main differences with net-

works generated through stationary processes. By doing so, we would be able to more fully characterize the principles underlying propagation of activity in real neuronal networks.

4

Graph Theory and its Importance for Neural Activity

Many real world phenomenons are characterized by the same underlying super-structure, where a collection of agents, or units, interact with one another on the basis of the existence or absence of a relation, the effectiveness of which may vary from one couple of interacting ends to another. One of the best exemplification of such phenomenons could be a social network, where people constitute the set of agents, and the relation may stand for friendship, political affiliation, business flow etc, with the effectiveness quantifying the quality of the friendship, the predisposition for a certain political orientation, or the quantity of goods and services exchanged. Within such a paradigm, a fundamental feature that strongly impacts how interactions propagate across agents, namely how the relation between person 1 and person N is affected by the relation of person 1 with other people, or by the relation of people external to 1 and N, is the global pattern of such relationships, or their topology.

The issue of modelling a network of the kind described above first occurred in the 18^{th} century, while Euler was seeking a way to allow people in Königsberg to walk around the city, and visit its central island, walking across each of its seven bridges only once. In order to solve this problem, Euler relied on an abstract object that has ever since been at the centre of discrete mathematics: the graph [39].

Definition 4. *A Graph G is identified by a tuple $G = (V, E)$, where $V = \{v_1, .., v_N\}$ denotes the set of nodes of G, and $E = \{(v_i, v_j)\}_{i,j=1,..,N}$ the set of its edges. In particular, if $(v_i, v_j) = (v_j, v_i) \quad \forall i, j = 1, .., N$, then G is undirected, otherwise it is directed.*

For computational simplicity, from here on we will define a coupling matrix $J \in \mathbb{R}^{N\times N}$ such that

$$J_{ij} = (v_i, v_j) \tag{4.1.1}$$

and that quantifies the strength of the connection from unit j to unit i. However, the existence of an edge between two nodes is not guaranteed, so that if $\nexists(v_\alpha, v_\beta) \in E \Rightarrow J_{\alpha\beta} = 0$.

Definition 5. *Let $G = (V, E)$ be a graph with nodes $V = \{v_1, .., v_N\}$. If $\forall i, j = 1, .., N$ $\exists(v_i, v_j) \in E$, then we say that the graph G is fully connected.*

Conversely, if $E = \emptyset$, then each node in the graph is isolated from the others, and any developing dynamics will be due to intrinsic properties of the node itself. The type of graphs that will be considered in this thesis lies somewhere in between these two extremes, and they are defined as having "sparse connectivity", meaning that, if we define $E_i = \{(v_i, v_j)\}_{j\in\{1,..,N\}}$, then $|E_i| = K \ll N$. But if $J_{\alpha\beta} = 0$, can we say that the state of node v_β has no effect on v_α? In explicit term, $J_{\alpha\beta} = 0$ means exactly this, but on an implicit level the statement is wrong, or at least partially incorrect; the state of v_β may indeed have an effect on the state of v_α if there exists a path between the two.

Definition 6. *Let $G = (V, E)$ be a graph, and consider the nodes $v_\alpha, v_\beta \in V$. We say that $\exists P_{\alpha\beta}$ path between the two nodes if it exists a set of consecutive edges*

$$E_{\beta\to\alpha} = \{(v_\beta, v_{i_1}), (v_{i_1}, v_{i_2}), ..., (v_{i_l}, v_\alpha)\} \tag{4.1.2}$$

and the length of such path is $L(P_{\alpha\beta}) = |E_{\beta\to\alpha}|$.

Since the existence of paths between nodes depends on the connectivity of the graph, or its set of edges, it follows that the more connected a graph is, the more likely the existence of a path between any two nodes is. To this aim, we define the degree of a node, and its variants, as

Definition 7. *Let $G = (V, E)$ be a directed graph; then we define*

- *In-degree of v_i: $d^{in}(v_i) = |\{(v_i, v_j)\}_{j\in\{1,..,N\}}|$*
- *Out-degree of v_i: $d^{out}(v_i) = |\{(v_j, v_i)\}_{j\in\{1,..,N\}}|$*

The total degree of the node v_i is then given by $d^{in}(v_i) + d^{out}(v_i) = d(v_i)$.

Notice that if the graph is undirected the two measures coincide, and in this case we simply speak of total degree of a node.

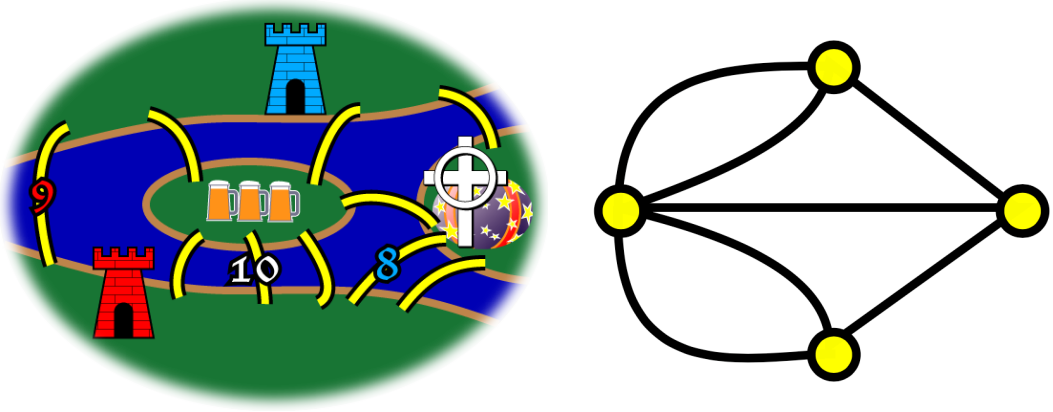


Figure 4.1: Visualization of the Königsberg bridge problem and of its graph transposition (source: https://www.wikipedia.org/)

Most of the work in graph theory up to the 1950s [2] has however focused on regular graphs, that is, graphs where each node has the same degree. It was only in the second part of the last century that graphs have been used to study large scale networks with no specific "wiring" rule for the edges, which are for this reason called random graphs. Indeed, most of the real world networks are generally random graphs, and their abstraction is first and foremost differentiated on the basis of whether it is generated by a static or dynamic algorithm.

4.1.1 The Erdős-Rényi Model

Consider a graph $G = (V, \#)$, where the $\#$ symbol signifies that the set of edges still needs to be constructed, and for simplicity assume that we are only considering presence or absence of a connection; hence

- $\exists (v_i, v_j) \in E \Rightarrow J_{ij} = 1$
- $\nexists (v_i, v_j) \in E \Rightarrow J_{ij} = 0$

Then the set of edges is statically constructed in the following way. Define $p \in [0, 1]$ [30], and set

$$\mathbb{P}(J_{ij} = 1) = p \tag{4.1.3}$$

$$\mathbb{P}(J_{ij} = 0) = 1 - p \tag{4.1.4}$$

Thereby, considering for the moment undirected graphs, we will have that nodes will be characterized by different degree, but that it is possible to characterize the expected degree of a generic node in the graph.

$$\begin{aligned} K &= \mathbb{E}[d(v_i)] \\ &= \sum_{ij} J_{ij}\mathbb{P}(J_{ij} = 1) + \sum_{ij} J_{ij}\mathbb{P}(J_{ij} = 0) \\ &= p\frac{N(N-1)}{2} \end{aligned} \tag{4.1.5}$$

It thus follow that there will be $\binom{\frac{N(N-1)}{2}}{K}$ equiprobable graphs, and we can define a probability space over such family as (Ω_G, F_G, ρ_G) such that $\forall G, \bar{G} \in \Omega_G \;\; \rho_G(G) = \rho_G(\bar{G})$. By doing so, we can formalize the notion of property of a graph as a function $Q : \Omega_G \to \mathcal{Y}$ such that

$$\rho_G(Q(G)) \xrightarrow[a.s.]{N \to +\infty} 1 \tag{4.1.6}$$

where the $a.s.$ (almost surely) signifies that the property Q holds for all the realizations of the graph in Ω_G except for at most a negligible set. Notice that the probability measure defined over Ω_G is dependent on the value of the parameter p, and as a consequence, so will be the observation of the property Q. Indeed, Erdõs and Rényi observed that properties of a graph arise quite suddenly [2] as p transitions over a range of values, and therefore it is possible to define a critical parameter $p_c : \quad \rho_G(Q, p) \to 1 \;\; \forall p \in [0, p_c[\vee]p_c, 1]$.

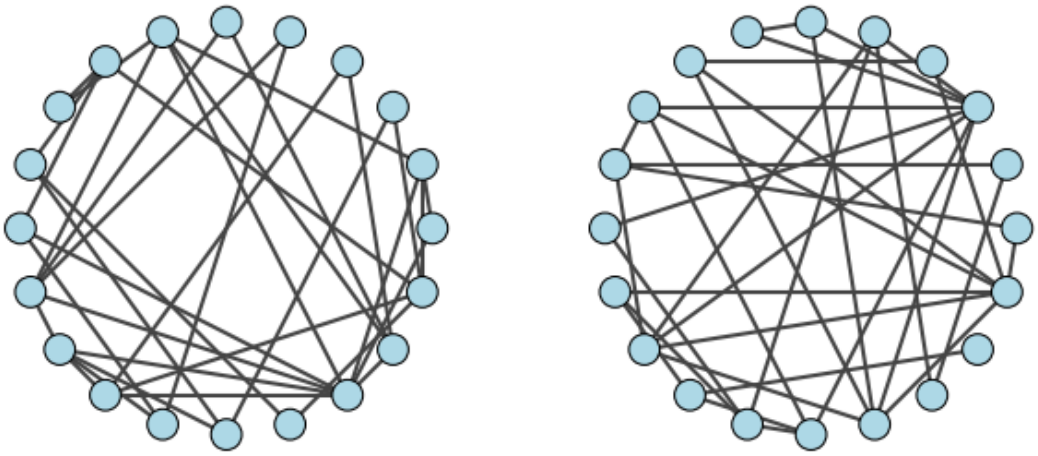


Figure 4.2: Visualization of two realizations of an Erdõs-Rényi random graph (source: https://igraph.org/python/tutorial/develop/tutorials/erdos_renyi/erdos_renyi.html)

Since all the nodes of the network are simultaneously instantiated, and therefore the emergence of edges does not depend on a specific ordering in which those are considered, then we are looking at a random graph generated by a stationary process.

4.1.2 The Watts-Strogatz Model

Let $G = (V, \#)$ be the usual graph without specified connectivity, and consider the set of nodes $V = \{v_1, .., v_N\}$ such that they are disposed around a ring and $v_i, v_{(i+1)},\ i = 1, .., N$ are nearest neighbours, for $N + 1 = 1$. Define $K \ll N$ as the initial number of edges that each node must have, and set

$$\mathbb{P}(J_{ij} = 1) = 1 \iff |i - j| < \frac{K}{2} \tag{4.1.7}$$

which means that each node is connected to its K nearest neighbours on a ring topology. Then we obtain a ring graph, which defines a state of maximum topological order.

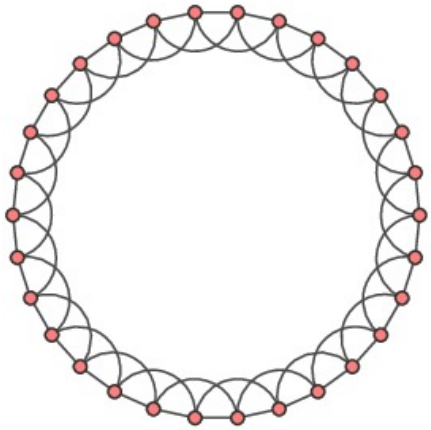


Figure 4.3: Visualization of a realization of a ring graph (source: http://www.ams.org/images/fcarc-august2012-ws.0.0.jpg)

Now, we want to introduce some disorder, or randomness, in our structure, and in order to do so, we define a parameter $q \in [0, 1]$ such that for each central node $v_i \in V$ we can define $\tilde{V}_i = \{v_j \in V : |i-j| < \frac{K}{2}\}$ set of K-nearest neighbours nodes and $\tilde{E}_i = \{(v_i, v_j) : |i - j| < \frac{K}{2}\}$ set of the related (undirected) edges. Then we define

$$\mathbb{P}(J_{ij} = 0 \ \wedge \ J_{ik} = 1) = q \quad v_j \in \tilde{V}_i, v_k \in \tilde{V}_i^c \tag{4.1.8}$$

which is essentially a rewiring process, which creates "long-range" connections with nodes outside the club of $\tilde{V}_i$. More concisely

$$\mathbb{P}(J_{ij} \neq 0) = (1-q)\frac{1}{\sqrt{K^2 2\pi}} e^{-\frac{|i-j|^2}{2K^2}} + q\frac{K}{N} \tag{4.1.9}$$

Notice that, given this parameter, we will have

- $q \to 0$: ordered (ring) network
- $q \to 1$: random network

Of particular interest is the range of q-values lying in between the two extremes, and specifically how the transition in such interval affects two fundamental graph measures, which are the average shortest path length and the average clustering coefficients.

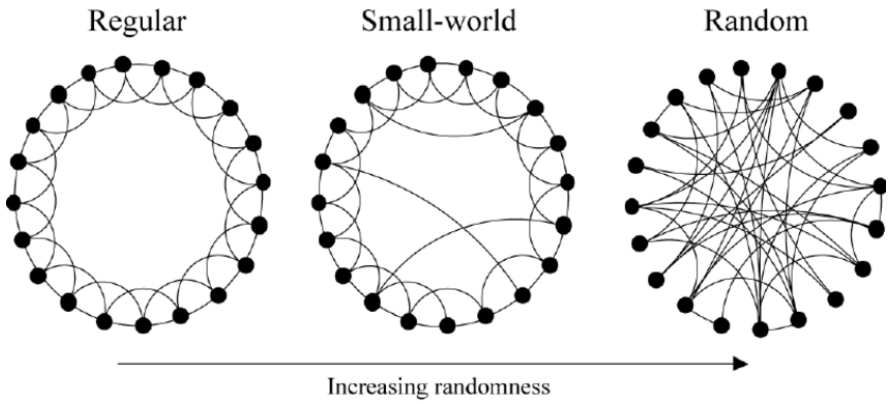


Figure 4.4: Visualization of the transition from a ring graph to a random graph (source: [37])

Formally, we define the average shortest path length for a directed network as

$$L(P)^{-1} = \frac{1}{N(N-1)} \sum_{ij} min(L(P_{ij}))^{-1} \tag{4.1.10}$$

also called harmonic mean since it handles the case where there does not exist a path between two nodes, and hence attributes $min(L) = +\infty$ to the shortest distance between the two.

Consider now a generic node $v_i \in V$, and for simplicity assume that $d(v_i) > 1$, and define the set of triangles connected to it as

$$\begin{aligned} T_i = \{(v_i, v_j, v_k) : \; & \exists P_{ij}, P_{jk}, P_{ki} \; \vee \; \exists P_{ik}, P_{kj}, P_{ji} \\ & \wedge \; min(L(P_{\alpha\beta})) = 1 \, , \forall \alpha, \beta \in \{i, j, k\}, \alpha \neq \beta\} \\ & v_j, v_k \in V_i \end{aligned}$$

and accordingly, define the set of triples that are connected to node $v_i \in V$ as

$$\begin{aligned} \bar{T}_i = \{(v_i, v_j, v_k) : \; & \exists P_{ij}, P_{jk} \; \vee \exists P_{ik}, P_{kj} \\ & \wedge \; min(L(P_{\alpha\beta})) = 1 \, , \forall \alpha, \beta \in \{i, j, k\}, \alpha \neq \beta\} \\ & v_j, v_k \in V \end{aligned}$$

Then, the clustering coefficient for an unweighted, undirected network associated to node v_i is

$$C_i = \frac{|T_i|}{|\bar{T}_i|} \tag{4.1.11}$$

and the average clustering coefficient

$$C = \frac{1}{N} \sum_{i=1}^{N} C_i \tag{4.1.12}$$

Now, we have that as $q \to 0$ both $L(P)$ and C tend to be relatively large, while for $q \to 1$ they both tend to be relatively small. However, there exists an intermediate range of q-values such that $L(P) \searrow$ while C remains relatively large; because of such properties, and drawing its etymology from social sciences, these networks are referred to as small-world. Indeed, the label refers to the way in which in social networks most people have a closed group of friends, with some sporadic acquaintances from the other part of the world.

In addition, notice that since all the nodes of the network are simultaneously instantiated,

the Watts-Strogatz model is an example of a network generated through a stationary process.

4.1.3 The Barabasi-Albert Model

Many real world networks are scale-free, meaning the the degree distribution $\mathbb{P}(d(v_i) = K)$ follows a power-law distribution that deviates significantly from a Poisson. The generative processes seen so far do not seem to be able to instantiate a distribution of this kind on the degree of the nodes of G. For this reason, a question that naturally arose was if there exists a generating process that produces this type of network, and whether such process is stationary or dynamic [2].

The first observation leading to the modelling of such network was that, in many real world contexts, networks do not emerge with a number of nodes fixed a priori, but rather undergo an evolutive process where new nodes are continuously added to an existing structure, and edges are created according to some well-defined dynamics. The second pivotal observation was that, differently from the networks studied so far, most real world networks exhibit preferential attachment, a phenomenon where the likelihood of a node to receive a new connection depends on its degree. Thereby, the algorithm for the creation of the new type of network, which will be called Barabasi-Albert network, is

1. Instantiate n nodes $\{v_1, .., v_n\} = V_0$ with arbitrary connectivity E_0.
2. Define a final evolution time T.
3. Add a new node: $V_1 = v_{n+1} \cup V_0$.
4. Define a fixed number of connections $n_0 < n$ that each new node must generate.
5. Let $d_j(\tau)$ denote the degree of a generic node $v_j \in V_\tau$ at time-step $\tau \leq T$. Define the probability for node $v_i \in V_{\tau-1}$ to receive an edge from the new node $v_{n+\tau}$ as

$$\Pi(d_i(\tau) = d_i(\tau - 1) + 1) = \frac{d_i(\tau - 1)}{\sum_j d_j(\tau - 1)} \tag{4.1.13}$$

6. At time T, we will have a graph $G = (V_T, E_T)$ such that

- $|V_T| = n + T$
- $|E_T| = |E_0| + n_0 T$

It has been shown that the resulting graph $G = (V_T, E_T)$ is such that the probability that a node has final degree $\mathbb{P}(d(v_i) = K)$ follows a power-law distribution with exponent $\mu = 3$, independently from n_0.

Interestingly, the average shortest path length is, for an equal number of nodes N, relevantly smaller in the Barabasi-Albert model than in a random network. Instead, the global clustering coefficient of the network seems to strictly depend on the exponent μ (supposing to consider different scale-free networks) [32].

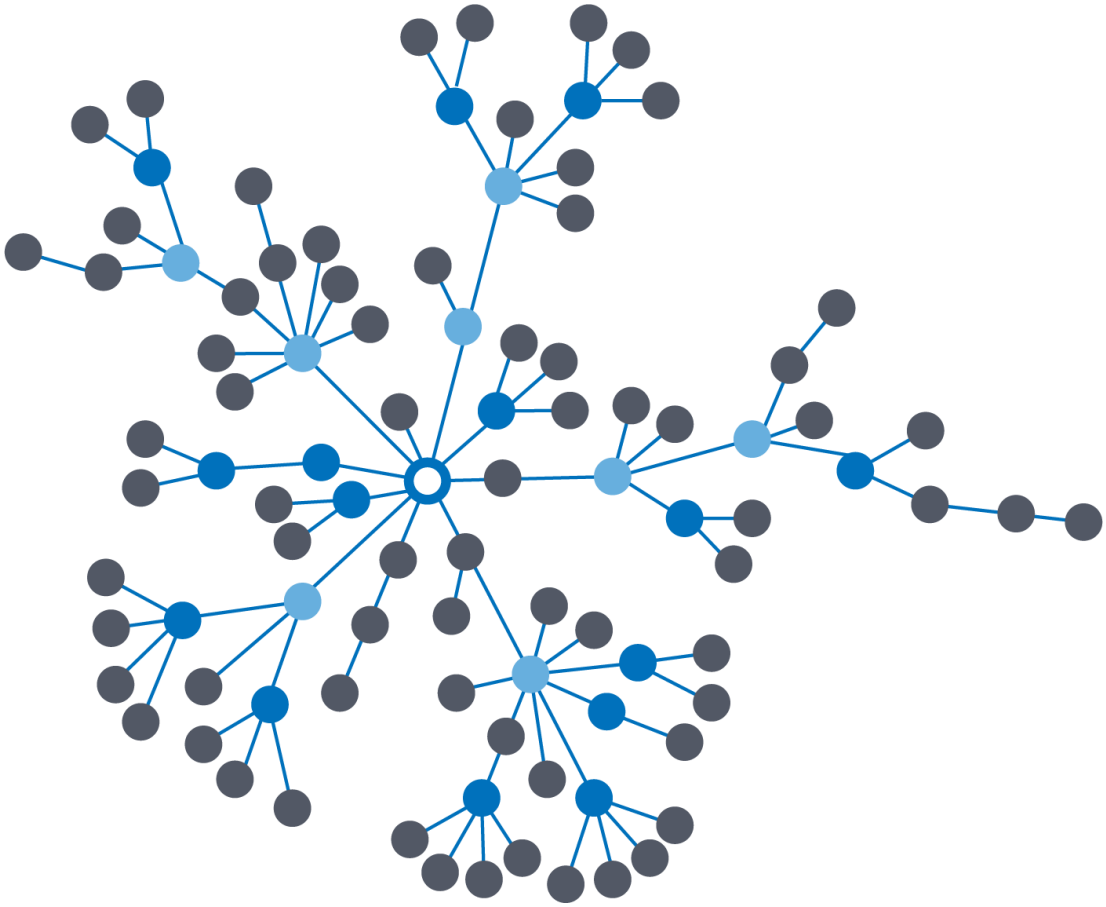


Figure 4.5: Visualization of a realization of a Barabasi-Albert Graph (source: https://www.oreilly.com/library/view/graph-algorithms/9781492047674/ch01.html)

In this section we proceed with a brief overview of some centrality measures, which quantify the importance of a node in the network according to some well-defined criteria. Each of this measures exploits different properties of the graph $G = (V, E)$, and for this reason there is none of them that unequivocally and absolutely ranks nodes, but rather each of them provides insight into a specific feature of the network [7]. Let $V = \{v_1, .., v_N\}$

Degree Centrality

As the name suggests, this centrality measure ranks the nodes of the network on the base of their degree, and is nothing more than the normalized degree.

$$c_{v_i}^{deg} = \frac{d(v_i)}{N-1} \tag{4.2.1}$$

Closeness Centrality

The measure of closeness centrality expands on degree centrality by considering the paths that stem from a generic node $v_i \in V$, considering neighbourhoods of all possible radii up to the diameter of the network.

$$c_{v_i}^{clos} = \frac{N-1}{\sum_{j \neq i} min(L(P_{ij}))} \tag{4.2.2}$$

Katz-Bonachic Centrality

The Katz-Bonachic centrality measure ranks the nodes in a graph on the basis of the number of walks radiating from it. It relies on the assumption that shorter walks should yield higher

prestige to the node, and for this reason it relies on a discount factor $\gamma \in (0, 1)$. Then

$$c_{v_i}^{ktz}(\gamma) = \sum_{l=1}^{+\infty} \gamma^l \sum_j \delta(l - L(P_{ij})) \tag{4.2.3}$$

where in this case we are not only considering the shortest path, but any possible walk between any two given nodes $v_i, v_j \in V$.

Betweenness Centrality

The betweenness centrality measure ranks the nodes in a graph on the basis of how important they are at connecting other nodes, hence their hubbiness. Formally, let $\nu_i(j, k)$ be the number of geodesic from node v_j to node v_k that pass through v_i, and $\nu(j, k)$ simply the number of geodesic between the two nodes. Then

$$c_{v_i}^{betw} = \frac{2}{(N-2)(N-3)} \sum_{v_j, v_k \in V \setminus \{v_i\}} \frac{\nu_i(j, k)}{\nu(j, k)} \tag{4.2.4}$$

Pagerank Centrality

The Pagerank centrality measure is a special centrality measure developed by Google's founders to rank web pages by popularity, and it relies on the following ideas (Zhang et al. [46]).

1. It must be proportional to the number of incoming edges
2. It is inversely proportional to the out-degree of the sending nodes
3. It is proportional to the pagerank centrality of the senders

Given this definitions, it is obvious that the formulation of pagerank centrality will be recursive

$$c_{v_i}^{pgrk} = \gamma \sum_j \left(\theta \frac{J_{ij}}{\sum_h J_{jh}} + (1 - \theta) \frac{A_{ij}}{\sum_h A_{jh}} \right) c_{v_j}^{pgrk} + \frac{(1 - \gamma)\beta_i}{\sum_j \beta_j} \tag{4.2.5}$$

If the coupling matrix $J \in \mathbb{R}^{\mathbb{N}\times\mathbb{N}}$ and the associated adjacency matrix $A \in \{0,1\}^{N\times N}$ coincide, we get the unweighted pagerank score. Finally, β is the non-uniform relative importance of each node.

4.3 Graph Structure and Firing Rate in a Network of LIF Neurons

Most of the examples permeating the literature of neural networks modelling up to the present moment instantiate different types of connectivity, primarily depending on the purpose of the modelling effort. Indeed, in the machine learning literature, where the goal is to build a network able to solve a specific task, most of the architectures are instantiated as fully connected, whereas in more realistic contexts, networks of neurons are connected by means of a random, sparse connectivity; however, as we have seen, different network topologies yields significantly different network properties. Consequently, it is sensible to ask ourselves whether neural networks characterized by paradigmatically different network topologies will also be characterized by significantly different activity, and how network features may be exploited to predict relevant aspects of such dynamics.

It has already been established that known examples of neuronal networks are organized into a small-world topology [43][45], which seems to promote efficient encoding of information and memory formation. Expanding on this knowledge, it would be interesting to understand if some of the measures detailed above are useful for the prediction of the associated neural activity. In addition, gauging how this measures correlate with the activity of networks with different topology, it may be possible to understand the network features that are most relevant for the propagation of activity. For this reason, Fletcher et. al. [15] have simulated a network of LIF neurons connected according to specific topological choices and examined how activity and measures correlate. Specifically, they have looked at

- Randomly Wired Excitatory Network
- Randomly Wired Excitatory and Inhibitory Network
- Randomly Wired Excitatory and Inhibitory Network satisfying Dale's principle

- Small-World Excitatory Network
- Small-World Excitatory and Inhibitory Network
- Small-World Excitatory and Inhibitory Network satisfying Dale's principle

where in the case of networks satisfying Dale's principle, we not only define excitatory and inhibitory connections, but inhibitory and excitatory units, possessing all the exiting connections belonging to the specified classification. Indeed, Dale's principle asserts that every mature neuron releases neurotransmitters of the same type at all its synaptic sites.
Fletcher et. al. [15] have found that the only measure that reliably and consistently correlates with the firing rate of the units is Katz centrality, whereas most of the other measures correlate only in the simplest case of random excitatory connectivity. The authors suggest that the reason why Katz centrality correlates so well with firing rate independently from the complexity of the network is because, as detailed by 4.2.3, the measure perfectly captures the influence that the activity of a unit has on another unit l-hops away, with such effect decaying with distance.

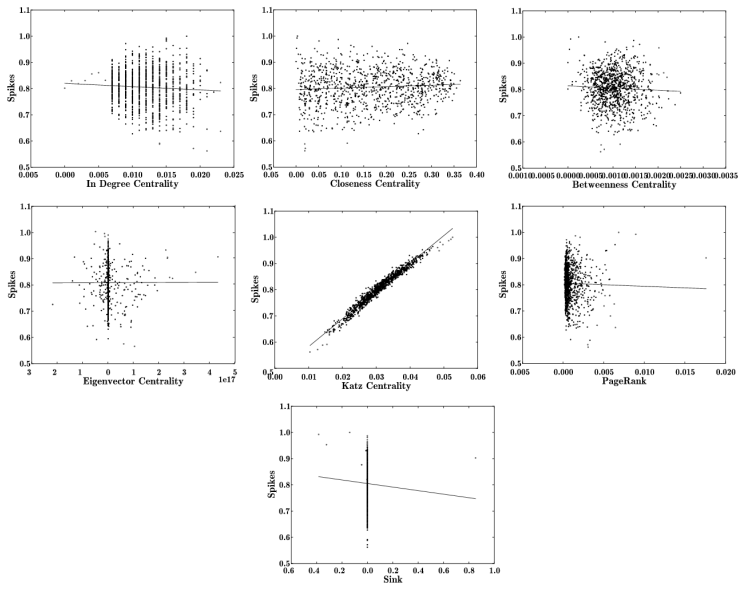


Figure 4.6: Plot for the normalized neural activity against the graph measure of interest. (source: [15])

It now remains to be understood whether such correlation holds independently of the specified network dynamics, that is, whether simulated cortical networks characterized by units with a different type of activity dynamics will display the same correlation.

4.4 Graph Structure and Memory in a Network of LIF Neurons

As we have already seen in subsection 4.1.2, the Watts-Strogatz model is endowed with a parameter q that allows for the transition from a state of maximum order to a state of pure randomness. The effects that the extremes of this parameter exercise on the network activity have been extensively studied, and it has been established that for $q = 0$ (ring topology) [6] the network exhibits bumps in activity coherent with self-sustained excitation of units in a localized neighbourhood. Instead, for $q = 1$ (random topology), the network exhibits distributed activity coherent with the emergence of autoassociative memory [3], where the short path length is functional at the network-wide propagation of activation. What remains to be established is whether W-S networks characterized by a q parameter within some intermediate interval $]q_1, q_2[$: $q_1 > 0 \wedge q_2 < 1$, hence small-world networks, are able to sustain both localized activity and retrieval of stored patterns in virtue of their high clustering coefficient coupled with a short path length. In principle, one would expect that a range of q-values providing the desired regime should exist, as the localized activity mainly depends on the existence of localized clusters of units, ensured by the high clustering coefficient [44], and the memory retrieval depends on the ability to propagate partial cues across the entire network, hence on short average path length.

In order to prove such hypothesis, Anishchenko and Trevez [5] have simulated a family of W-S networks of LIF neurons obeying dynamics akin to 1.4.1, and regulated by a varying parameter q. They have found that for low values of the q parameter ($q \in [0, 0.2]$), the network displays robust bumps in activity that are resilient to the introduction of full or partial cues, having the purpose of driving the system towards a specific state. Conversely, for $q \in [0.8, 1]$, hence for a condition where the model approaches a random network, we have

that the architecture is perfectly able to retrieve stored memories with partial cues, while displaying no localization of activity. Through a gradual increase of q-values, the authors have observed that the total loss of localization occurs for $q \geq 0.6$, while memory retrieval emerges for $q \geq 0.5$. Thereby, they have hypothesized that, if an interval of coexistence for localized activity and memory retrieval exists, it should be very narrow. Furthermore, through this study they have proven that the small-world regime is not functional at providing an understanding of the interplay between the two phenomena, which plays a major role in the activity of biological neuronal networks.

Following the same modus operandi, we now ask whether different degree of randomness in a network, as parametrized by q, generate significantly and statistically different MSR distributions and, if such change is observed, whether it reflects on other metrics of spiking variability, or more generally on other statistics of the network activity.

5 Methods

5.1 Hardware

Almost all of the simulations have been run on the IDUN cluster of NTNU, which is a high performance computer administered by NTNU and shared among the different departments, each owning a holding over the architecture. It comprises 8 possible login ports, each endowed with about 27 nodes with Intel Xeon E5 or Gold processors, 128 Gb of RAM per node, and eventually a GPU among different models (NVIDIA Tesla, NVIDIA Titan, NVIDIA A100). The simulations were run simulatneously on a number of nodes varying from $2-5$, each simulation lasting about 48 hours when not optimised, and 23 hours when optimised.

5.2 Software

The code for the simulations was written in Python 3.9, mainly exploiting the functionalities of the package NumPy, already provided in the Anaconda distribution of the product. The graph associated to the network was instantiated through the package Networkx, and all the graph measures were computed using the functionalities within this package. The plots for the results of the simulations were drawn with the support of the Matplotlib package.
Finally, the code was optimised by means of the Numba package, which provides a translation from Python to C++ for specified functions, with the best results on those that mainly exploit NumPy tools.

5.3 Network

The network was instantiated as an heterogeneous population of excitatory and inhibitory units, with outgoing connections satisfying Dale's principle. Specifically, the population was characterized by

- $N_E = 1000$ Excitatory units
- $N_I = 1000$ Inhibitory units
- $K = 0.1 \cdot min\{N_E, N_I\}$

and the associated coupling matrix had the following structure

$$J = \left(\begin{array}{c|c} J_{EE} & J_{EI} \\ \hline J_{IE} & J_{II} \end{array}\right)$$

and the topology of the network was not implemented on J as a whole, but rather on each block J_{kl}, so that by choosing a network topology $\mathcal{T}$, we would have that units in population k would be connected to units in population l according to $\mathcal{T}$, for $k, l \in \{E, I\}$. If, for example, we set $\mathcal{T}$ =W-S$(p = 0)$, we would have that excitatory units are organized according to a perfect ring topology of K-nearest neighbours, and so will be their connection to the inhibitory population, so that finally each unit $v_i \in E$ will have, on average, incoming connections from $2K$-nearest neighbours, half excitatory and half inhibitory.
The couplings were set, in the fixed weights condition, as

$$J_{EE} = 1$$
$$J_{IE} = 1$$
$$J_{EI} = -1.8$$
$$J_{II} = -2.0$$

The external currents were set in the range

$$I_E \in [1.35, 14]$$
$$I_I \in [1, 12.5]$$

The population thresholds were set, as in the original paper by Sompolinsky et. al. [41], as

$$\theta_E = 1.0$$
$$\theta_I = 0.8$$

Finally, the total simulation time is $T = 60000$, where the value has been appropriately chosen so that the maximum plausible number of spikes for the simulated networks is approximately the same as that registered by Stensola et al. [36] for the neurons in mEC.

5.4 Area and Time-Steps in MSR Computation

With reference to the MSR formulation 2.3.3, 2.3.4 adapted to a time series, we consider a discrete set of time-steps to discretize the spike series

$$\Delta T = \{\Delta t \in [2, 10000]\} : \quad |\Delta T| = 150$$

generated through the following two function

- np.linspace(start=2, stop=80, num=40, dtype=np.int64)
- np.linspace(start=100, stop=10000, num=110, dtype=np.int64)

which ensure the generation of a set of points $(\hat{H}[s], \hat{H}[k])$ informative enough to compute the area underlying the interpolating curve. The resulting set

$$\tilde{H} = \{(\hat{H}[s](\Delta t), \hat{H}[k](\Delta t)) : \quad \Delta t \in \Delta T\}$$

was linearly interpolated using the $scipy.interpolate.interp1d()$ function, and the area enclosed by the associated curve was computed by means of numerical quadrature through the use of the function $scipy.integrate.quad()$.
For computational purposes, units that spiked less than 5 times out of the whole simulation were discarded from the analysis, lest numerical errors.
Another important point of order is that, in most simulations, the highest MSR values needed, in theory, to be discarded, as sparsely spiking units generate very few $(\hat{H}[s], \hat{H}[k])$ pairs, and this in turn resulted in very coarse linear interpolation and subsequent numerical quadrature. For example, a unit generating only 4 of such pairs would produce an MSR graph that has a trapezoidal shape, and that potentially could include far more area than the expected parabola.

6

Results

The present chapter is entirely dedicated to the illustration of the experimental results obtained from the simulations of the network of chaotic activity while a host of parameters change gradually. Specifically, we will first verify that the simulated networks satisfy the theory outlined in chapter 3, and in particular the linear response, to make sure that there are no errors in the implementation. In addition, we will examine how, keeping the number of units, average number of connections per unit, population specific weights and thresholds fixed, the distributions of the different activity metrics change as we transition from a state of topological order ($q = 0$) to a state of topological randomness ($q = 1$) for the network, as detailed in chapter 4.

6.1 Fixed Population Specific Weights

Considering for each of the injected currents I the state matrix, which is a $S(I) \in \mathbb{R}^{(N_E+N_I)\times T}$ matrix having entry $S_{it} = 0$ if unit i is quiescent at time t, and $S_{it} = 1$ if it is active, we can get the network mean response as

$$\bar{m}(I) = \frac{1}{(N_E + N_I) \cdot T} \sum_{i=1}^{N_E+N_I} \sum_{t=1}^{T} S_{it}(I) \tag{6.1.1}$$

Starting from the case of the random topology, which is the same explored by Vreeswijk and Sompolinsky [41], we can indeed observe a linear response of both the excitatory and inhibitory populations to the provided input (suitably averaged between the two populations and subsequently normalized)

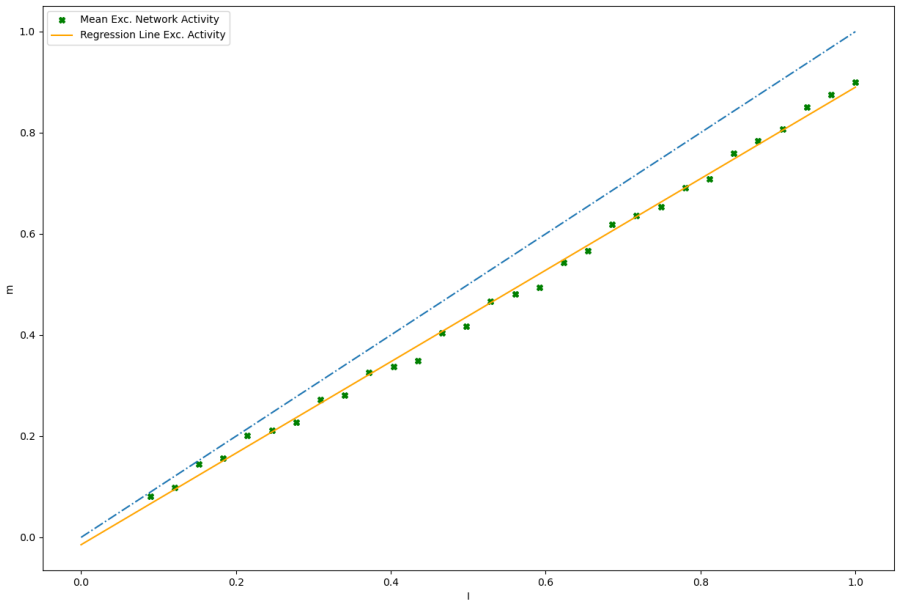


Figure 6.1: The excitatory population of the network characterized by random connections exhibits an almost perfect linear response to the averaged external input according to 3.2.19 with $\Phi_E < 1$.

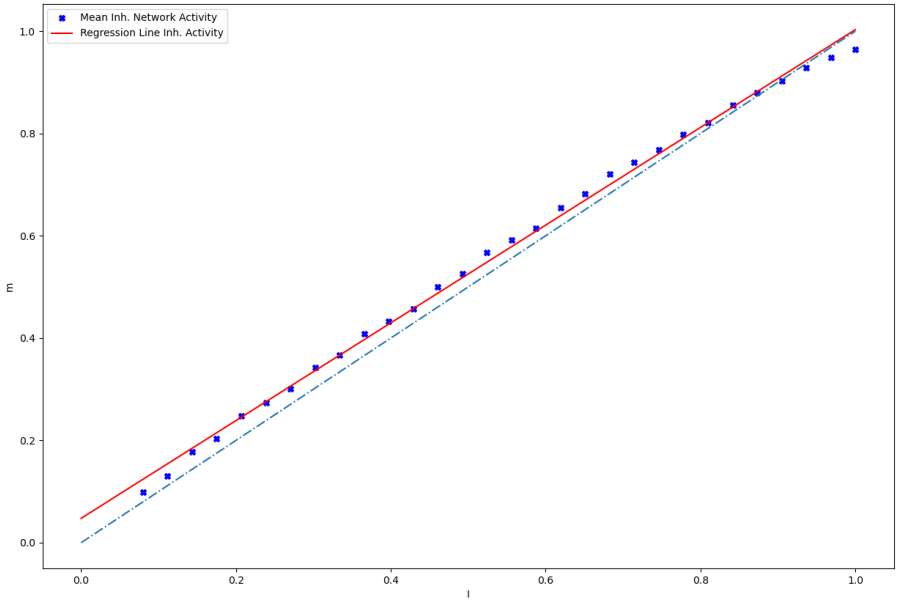
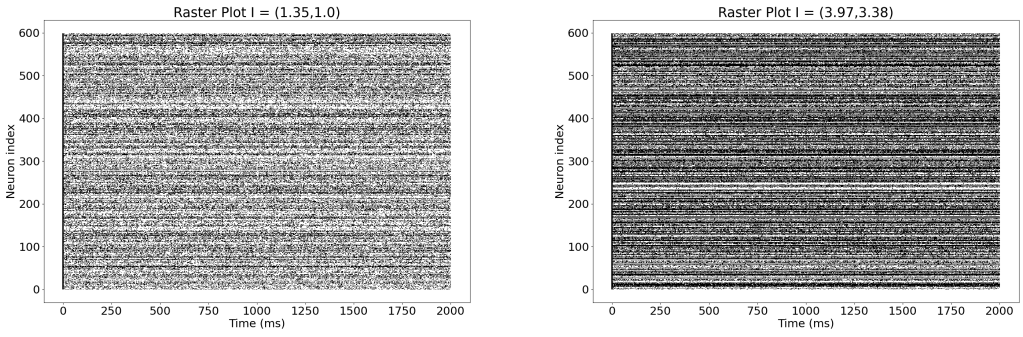


Figure 6.2: The inhibitory population of the network characterized by random connections exhibits an almost perfect linear response to the averaged external input according to 3.2.20 with $\Phi_I \approx 1$.

Consequently, the network described by the authors has been adequately simulated, and we are only left with the verification that this result is not related to the formation of frozen states, but rather to a time varying chaotic activity. Choosing 4 reference external inputs, in an increasing order, we indeed verify that the associated activity is chaotic.



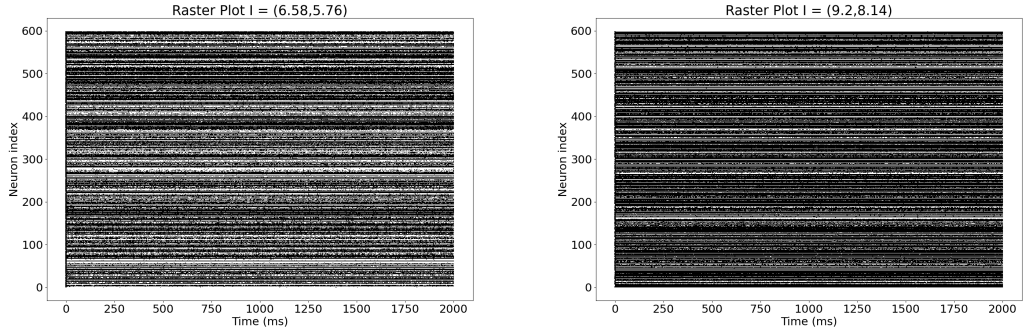


Figure 6.4: The four raster plots display different regimes of chaotic activity, with absence of frozen states, as the external current gradually increases. As can be observed, the increase in the external current is directly related to an increase in the network activity (black dot - spike; white dot - no spike).

Consequently, as can be readily observed from the raster plots above, there is no formation of frozen states ($m = 0 \vee m = 1$) for any of the presented external currents, and we can conclude that the archetypal network has been appropriately simulated. It now remains to gradually vary the q-parameter, decreasing it from 1 to 0, to investigate how different topological conditions affect the distributions of the metrics computed on the network activity and the correlation of the sames with relevant graph measures. Therefore, we begin by presenting the results on the MSR, Burstiness, Firing, and Memory distributions for the following configurations

- $q = 0.1$
- $q = 0.5$
- $q = 1.0$
- Appendix - $q = 0.3 \wedge q = 0.8$

and taking as reference the three following external current conditions

- $(i_E(t), i_I(t)) \equiv (1.35, 1.00)$
- $(i_E(t), i_I(t)) \equiv (6.15, 5.36)$
- $(i_E(t), i_I(t)) \equiv (12.26, 10.91)$

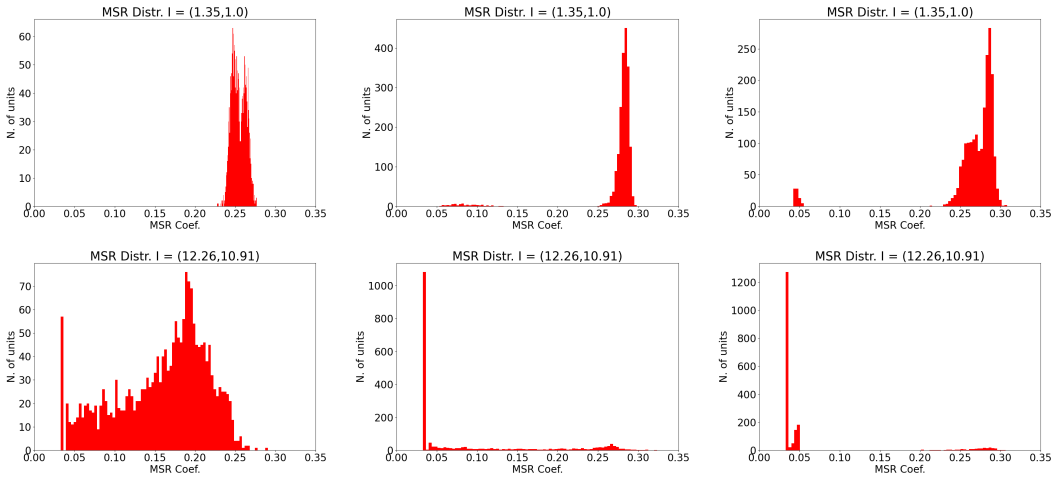


Figure 6.5: In the top panel, the MSR distributions for $(i_E, i_I) = (1.35, 1.00)$ and, from left to right $q \in \{0.1, 0.5, 1.0\}$ clearly show a narrowing and heightening of the peak for the small-world condition $q = 0.5$, while it tends to broaden for both high order and high randomness topological conditions. Instead, in the lower panel, for external conditions characterizing the saturation regime $(i_E, i_I) = (12.26, 10.91)$ and the same progression of q-values, we can observe how the high topological order associated to $q = 0.1$ and the consequent bump-like activity enables the network to sustain activity characterized by relatively high MSR, while for any other condition characterized by higher topological randomness the network saturates and produces meaningless (from the perspective of MSR) signals.

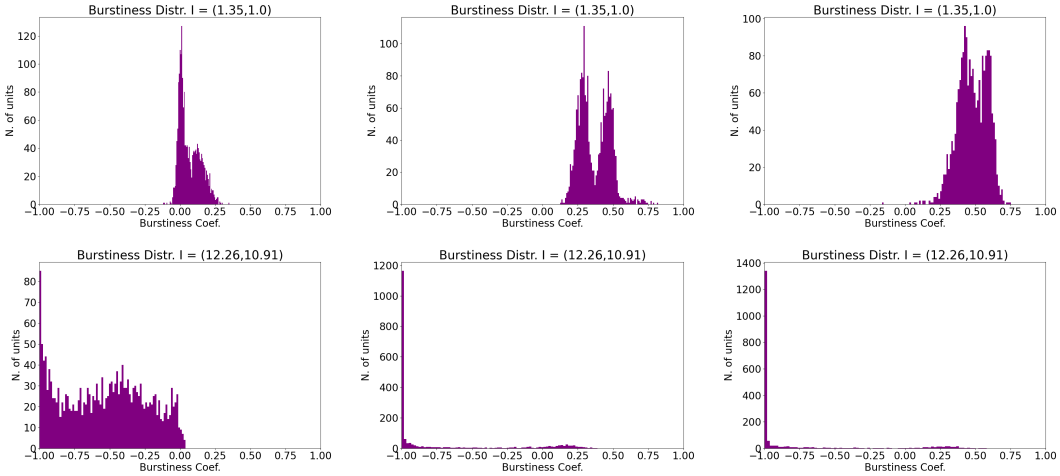


Figure 6.6: From the top panel of the Burstiness Coefficient distribution we can observe how for $(i_E, i_I) = (1.35, 1.00)$, its skewness towards high B-values becomes more and more evident as the q-parameter transitions from a state of high topological order ($q = 0.1$, top left figure) to one of topological randomness ($q = 1.0$, top right figure). Thus, for external conditions characterizing balanced mild network activity, the degree of randomness is directly proportional to the Burstiness Coefficient. In addition, from the bottom panel notice also how in the saturation regime $(i_E, i_I) = (12.26, 10.91)$ the higher the randomness in the network topology, the higher the variance in the inter-event times series (although for a small sample of units).

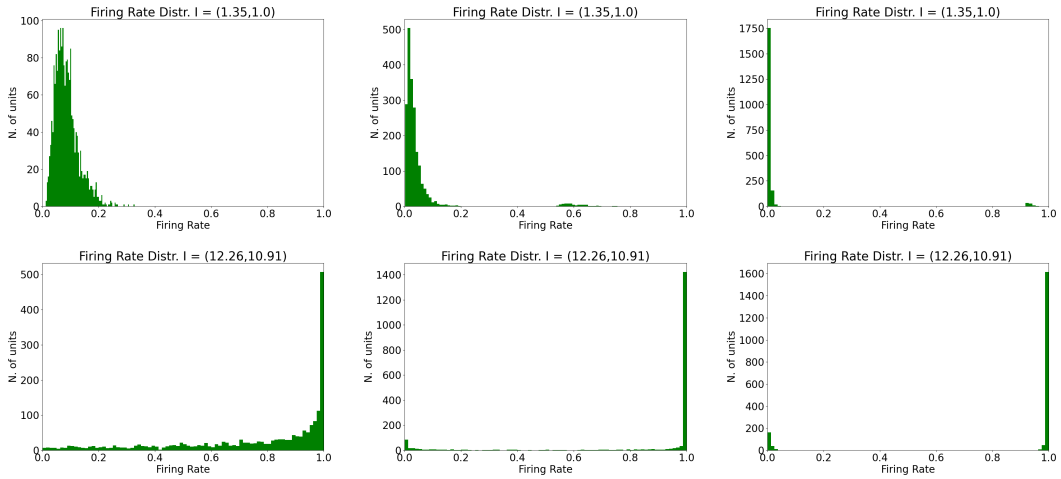


Figure 6.7: From both the top and bottom panels, displaying the firing rate distributions of networks characterized by $q \in \{0.1, 0.5, 1.0\}$ (left to right), we can observe how only networks characterized by high topological order present broad enough distributions, which in turn explains the much broader $\mathcal{R}_t, B$ distributions observed for the same network and for the corresponding external currents $(i_E, i_I) \in \{(1.35, 1, 00), (12.26, 10.91)\}$. Instead, for higher topological randomness ($q > 0.2$) the firing rate distributions are considerably peaked and characterized by sharp transitions, with a small sample of units which displays very high firing rates under low external currents, and another (not necessarily different) that displays very low firing rates under saturating external currents.

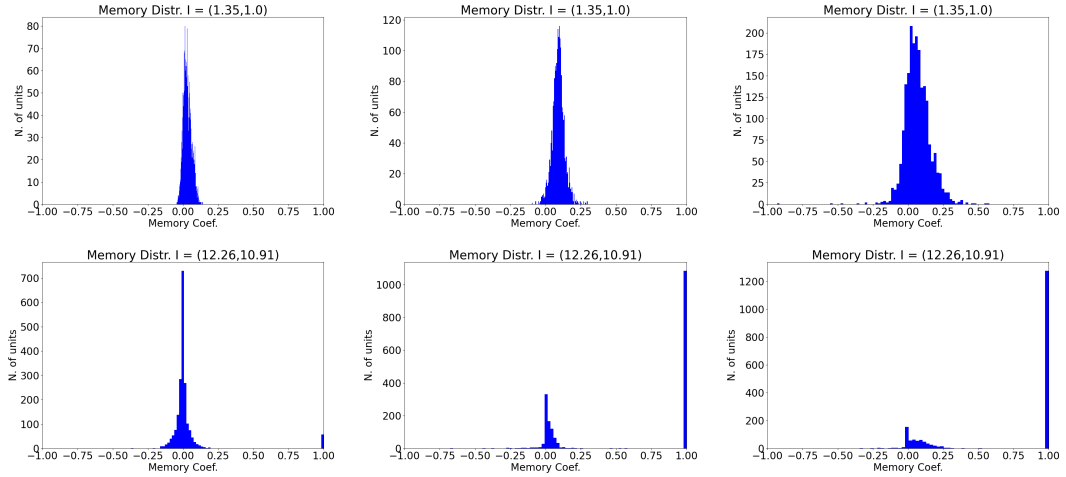


Figure 6.8: From the top panel, hence for low external currents $(i_E, i_I) = (1.35, 1.00)$, it can be observed how the Memory Coefficient distributions, which are characterized by an almost symmetric shape, display a drift of the mean towards the value $M = 0$ (no correlation between inter-events times) as we transition from states of high topological order ($q = 0.1$, top left figure) to states of high topological randomness ($q = 1.0$, top right panel). This is somehow to be expected, as networks characterized by higher topological order and displaying bump-like activity have the incoming input highly depending on their close neighbourhood. Results for very high currents are misleading, as the saturation resulting from the finite size of the network induces false correlations in the activity.

From a first inspection of the plots on the MSR distribution across q conditions, we can observe how both when the network tends to a state of maximum topological order (ring topology) and minimum topological order (random topology) the distribution tends to broaden, especially at low currents. However, as the current increases, the network with $q = 0.1$ is able to retain a fairly broad MSR distribution, probably due to the fact that the formed clusters are less sensible to external perturbations, and it is peaked towards the high MSR range, while the network with $q = 1.0$ becomes characterized by a MSR distribution sharply peaked around values in the low MSR range. Instead, it is interesting to see how networks characterized by $q \in [0.3, 0.8[$ present an MSR distribution (figures 8.5, 8.9) which, for low external currents, is visibly peaked in the high MSR range, and it presents an intermediate response to changes in external conditions, with the peak gradually shifting towards the low MSR range but not as drastically as for a completely random network. By virtue of what was presented in chapters 1 and 2, it might be argued that networks that are randomly connected may be somewhat unsuited for the propagation of information in a way that is compatible with the emergence of intelligent life, as when exposed to low excitation they present a fair number of units that are not optimally encoding, and all units become quickly useless for the purpose as excitation increases. To understand this heuristic argument, it suffices to think that we, as humans, are not exposed to homogeneous external inputs, but rather, to a heterogeneous variety continuosly varying in both relevance and intensity; thus, if the topology of the connections in our nervous system would have been completely random, we would be extremely sensible to external perturbations, which is somewhat in disagreement with the adaptability that we have displayed throughout evolution. On the other hand, a network organized according to a ring topology, while being resilient to external perturbation, would also be characterized by a large number of units that are not optimally encoding throughout all the possible external conditions, thus making the organism endowed with a neuronal topology somewhat unsuited to quick responses to the environment. Finally, networks that in virtue of their q-values might be characterized as small-world presents us with

results that are compatible with the evolution of an organism that is able to respond almost-optimally under many different external conditions, and that even in very high excitation regimes is able to retain some relevant and meaningful activity propagation capabilities. Of course, illustration of the results aside, the previous arguments are only qualitative and I dare say highly speculative.

An important note is in order, and it regards the range of external currents that a network is able to withstand without saturation in its activity. As can be proven experimentally (the reader is encouraged to try), the bigger is the size of the network, the bigger is the interval characterizing the balanced non-saturating network activity. Therefore, it must be kept in mind that any consideration made in this thesis with respect to regimes of balanced chaotic activity and saturated activity must be interpreted strictly within the adopted framework, hence the chosen number of units and the chosen weights (it has not been explored yet, but I suspect that the absolute value of the weights plays a relevant role in the determination of such range as well).

In order to assess whether there indeed exists a difference in the MSR evolution across external conditions for different network topologies, it is necessary to obtain an estimate that abstracts both from the identity of the single units and from the network initialization. For this reason, it is instructive to compute, for each current, a MSR coefficient that is the result of the average of the MSR coefficients of a single realization, averaged across a statistically significant number of realizations. Formally, let $G^1, ..., G^{\mathcal{L}}$ be the different realizations of the same network architecture at a fixed current (i_E, i_I); then we consider

$$\mathcal{R}_t((i_E, i_I)) = \frac{1}{\mathcal{L}(N_E + N_I)} \sum_{j=1}^{\mathcal{L}} \sum_{i=1}^{N_E+N_I} \mathcal{R}_t(v_i^j; (i_E, i_I)) \tag{6.1.2}$$

The results obtained for 5 realizations of each of the presented configurations are

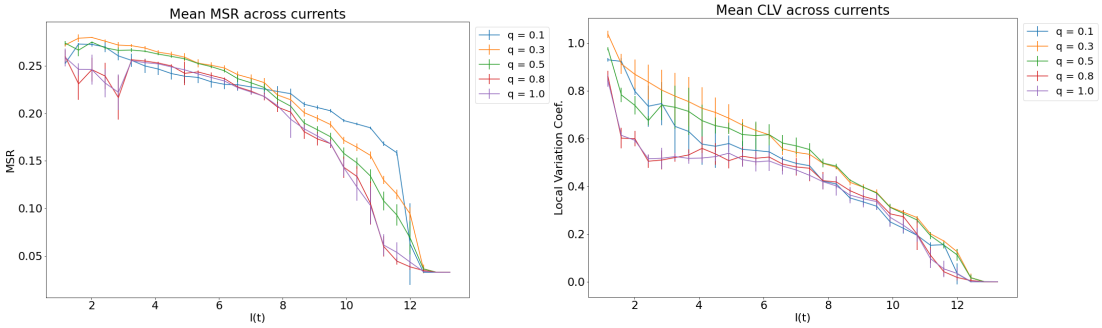


Figure 6.9: In the left panel, we see the MSR trajectory across external currents for networks characterized by different degrees of topological order, as quantified by the q-parameter. The average MSR coefficient for networks with q typical of the small-world range is systematically higher than the same for networks characterized by almost or complete topological randomness. Instead, for q-values typical of the high topological order condition, the same is true only up to $I(t) \approx 8$, where then the saturation of activity drives to zero the MSR values of the units of all the networks except those characterized by ring-topologies. In the right panel, the average L_V-coefficient remains systematically higher, across all currents, for networks characterized by q in the small-world range than for networks defined by either high topological order or randomness.

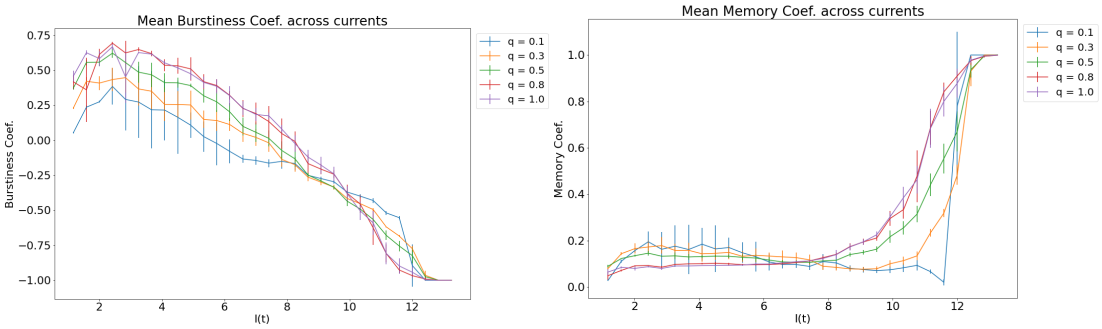


Figure 6.10: In the left panel, it is possible to observe the direct relationship existing between the degree of topological randomness of the network and the burstiness coefficient of its units: the more random is the network topology, the higher will the average burstiness coefficient be. The trend then reverses when external currents determine saturating regimes. In the right panel, we appreciate an opposite trend for the memory coefficient, as the higher the degree of randomness is, the lower the respective average memory coefficient. This reverses at $I(t) \approx 8$, again as a consequence of the networks being driven towards saturation and to spurious correlations appearing due to finite size effects.

As can be observed, the MSR values of networks characterized by q-parameters in the small-world range are statistically higher across most of the currents than those for q at the extremes of the interval (except of course for the high current regime when $q = 0.1$). This result supports the hypothesis that network wired according to a small-world topology produce, all else being equal, signals that are consistently more informative according to MSR than the same generated by random or highly regular topologies. We now want to under-

stand more thoroughly whether topology affects other metrics of signal variability in a way that is consistent with the results observed for MSR, and for the purpose we will focus on the burstiness coefficient, the coefficient of local variation, and the memory coefficient, suitably averaged in a way consistent with what was done for the MSR coefficient (see 6.1.2).

From the plot of the trajectory of the average burstiness coefficient across currents, it is clear how for most of the external conditions, the networks characterized by higher topological randomness preserve higher B-coefficients, and hence a higher inter-spikes time variability (of course, for very high currents networks characterized by greater topological order are the only one that do not saturate, and consequently preserve variance in the inter-spike times series).

The plot for the averaged coefficient of local variation across currents seems to depict a picture more similar to that portrayed by MSR, where the two networks characterized by q-parameters proper of a small-world regime obtain, throughout most of the external conditions, scores that are higher than those of networks defined by either high topological order or randomness. By means of this comparison, it may be safe to conclude that what MSR encodes in a signal is more closely related to the variability in the time elapsing between subsequent events, and how the entire inter-spike sequence may be discretized, which is in turn captured throughout the switches in temporal resolution defined by 2.3.2. It is worth noticing that the two measures that are the most in agreement are those that do not consider the inter-spike series as a whole, but rather those that systematically focus on chunks of the spike train and try to determine the amount of information that is possible to extract from that localized signal-stamp.

As could be expected, no significant trend is observable from the plot of the averaged memory coefficient across currents, as all the networks, by virtue of how they are modeled, are supposed to have complete or almost independently spiking units in the large N limit (but since our network is finite, obviously at low currents the ring topology yields higher averaged memory coefficients). The subsequent increase in the coefficient is mainly a numerical ar-

tifact as, given the finite size, the external currents increase drives all the networks towards saturation, and consequently spurious correlations between successive inter-spike times appear.

Finally, it is important to notice how, from $I(t) \approx 8$ on, the average MSR coefficient could enable the researcher, given appropriate information on the range of currents that the network under study can withstand, to determine the q-parameter characterizing the network, and therefore its topology. The same could potentially be extrapolated from the knowledge of the average memory coefficient; however, as visible from the plots 6.10, the distance in the average memory coefficient between different q-conditions at a fixed current is narrower, thus making discrimination of the topological signature somewhat harder. Specifically, in the graph for the average MSR coefficient there seems to be, for $I(t) > 8$, a clear differentiation of the ranges traversed by network characterized by ring, random, and small-world topologies. For this reason, we argue that given appropriate knowledge of the network sensitivity to external currents, experimental researchers could exploit the average MSR coefficient to determine, without any prior knowledge on connectivity, the topological structure of their network.

Now that we have established what is the topological dependence in the average trend of the considered spike metrics, it would be instructive to understand the dependence of the MSR coefficient on the firing rate of the units. This evaluation becomes necessary after considering that, for a network close to saturation (or quiescent, on the opposite end), the generated spike trains present little (high) variance in inter-event times, and consequently each time bin generated by 2.3.2 will have a similar (different) number of spikes. Consequently, we would expect MSR to be inversely proportional to the firing rate of each unit, and we want to understand if this hypothesis holds and whether there are any topological connotations in this functional dependence. In order to asses such hypothesis, we plot, for each unit in our network, its MSR value at a given current against the logarithm of the total number of spikes burst by the unit in the simulation. Following the conventions of 2.3.3, 2.3.4

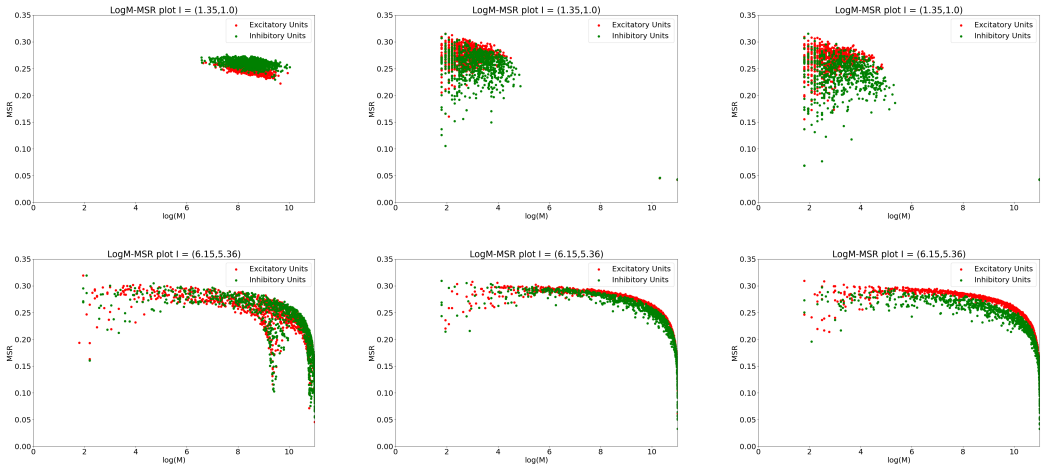


Figure 6.11: In the top panel, we have the $log(M) - \mathcal{R}_t$ plots for $(i_E, i_I) = (1.35, 1.00)$ with $q \in \{0.1, 0.5, 1.0\}$ (left to right), and we can appreciate two main differences. First, networks characterized by high topological order ($q < 0.2$, top left figure) display a considerably higher firing rate than either small-world or random networks, at low external currents. Second, networks within this high topological order range also display a much narrower MSR distribution for both the populations. In the bottom panel we consider instead the same plots but for $(i_E, i_I) = (6.15, 5.36)$; while the difference in the firing rate has disappeared, we can appreciate how, as q transitions from 0 to 1, the cluster of MSR values for the two populations seems to separate, thus hinting at the formation of two different population specific regimes of activity as the connectivity becomes random.

The first thing that we notice is that for any topological choice there is a clear relationship between the firing rate of each unit and its MSR value, with this relationship becoming increasingly clearer as we transition from low external currents to high external currents, with a drastic plummeting of MSR values as we approach the maximum possible firing rate. Notice however that, if we consider only the external current condition $(i_E, i_I) = (1.35, 1.00)$, then networks characterized by high topological order (fig. 8.21 and fig. 8.22) already display a quite high firing rate if compared to that of networks characterized by higher randomness and subjected to the same external condition. In addition, and keeping in mind that the images report results from a single realization of a specific network condition, it can be observed that as $q \to 1$, the two populations of units seem to separate with respect to the MSR values, with the excitatory population occupying the higher range. Furthermore, we can observe for $q \to 0$, compatibly with what can be seen in fig. 8.1 and 8.5, that the MSR distribution remains broad even for relatively high external conditions, while for $q \to 1$ the

entire $log(M) - \mathcal{R}_t$ plot tends to collapse on a few points occupying the right extreme of the $log(M)$ interval ($M \approx 60000$).

Once gained knowledge on the dependence of the MSR metric on the firing rate of the units, it would be interesting to understand the relationship between MSR and the other statistics that can be computed on the spike trains, to understand if there exist regions where the former and the latter provide the same quantification of the generated signal. Moreover, color-mapping the points in the plot according to the firing rate allows us to better evaluate how also the other statistics depend on the same. Finally, we extend such comparison to topologies characterized by different degree of randomness in connectivity, to assess the impact that this has on the correlation.

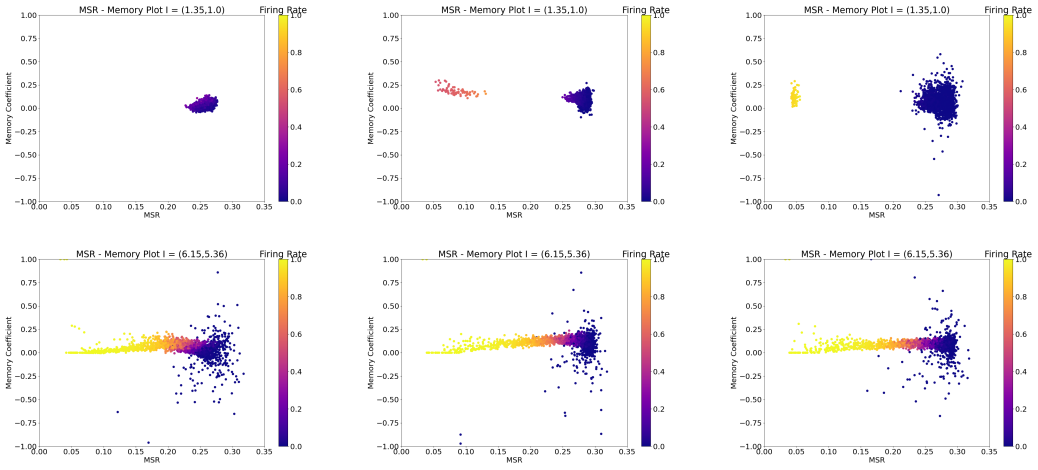


Figure 6.12: In the top panel, presenting results for the external condition $(i_E, i_I) = (1.35, 1.00)$, we see that for conditions of high topological randomness (top right panel), the memory M-values are much more spread around their central value $M = 0$. However, since from figure 6.10 we have seen that, for these conditions and low external currents, the average Memory coefficient is much closer to the central value, we can conclude that for most of the units characterized by memory coefficient $M = \tilde{m}$ there existing another unit in the network with $M \approx -\tilde{m}$; this does not appear to be the case when $q = 0.1$. In the bottom panel, for external current $(i_E, i_I) = (6.15, 5.36)$, the M-values are considerably more spread for $q = 0.1$ (bottom left panel), and the cluster of points characterizing the plots for $q = 0.5 \wedge q = 1.0$ (bottom middle and right panel) is gradually drifting towards higher M-ranges, thus informing us on the general rise of correlations between inter-events times as the external current increases.

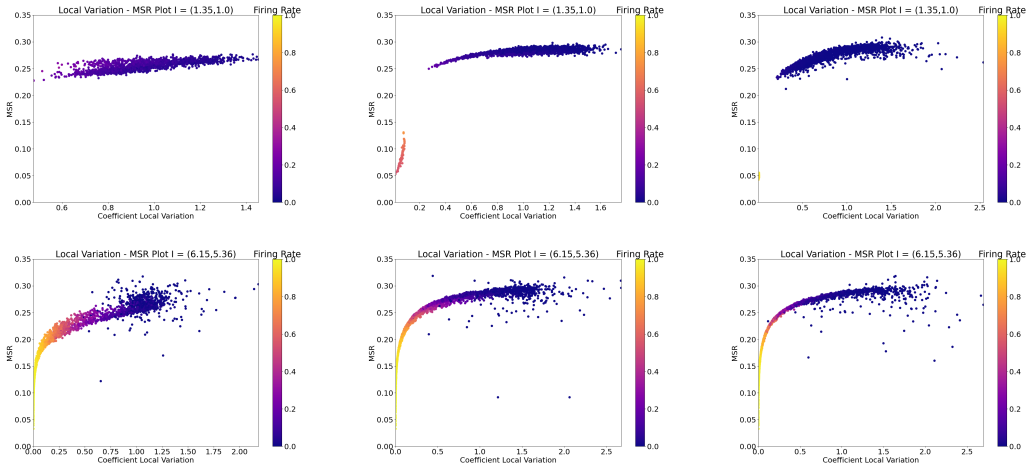


Figure 6.13: For relatively high topological order and low external current $(i_E, i_I) = (1.35, 1.00)$ ($q = 0.1$, top left figure) the relationship between the two metrics is approximately linear, and therefore the two encodes similar aspects of the units activity. As the topological randomness increases, we can still appreciate the presence of a general linear relationship, although as $q \to 1$ some non-linearity starts to appear (initial curvature). As the external current increases $(i_E, i_I) \geq (6.15, 5.36)$, a region of degeneracy appears (bottom panel); specifically, for all the topological conditions, we have that units with $\mathcal{R}_t \in [0, 0.15]$ span a very narrow L_V range. Furthermore, as q progresses towards values defining topological randomness, the degeneracy seems to become two-sided, with the relationship displaying an asymptotic trend.

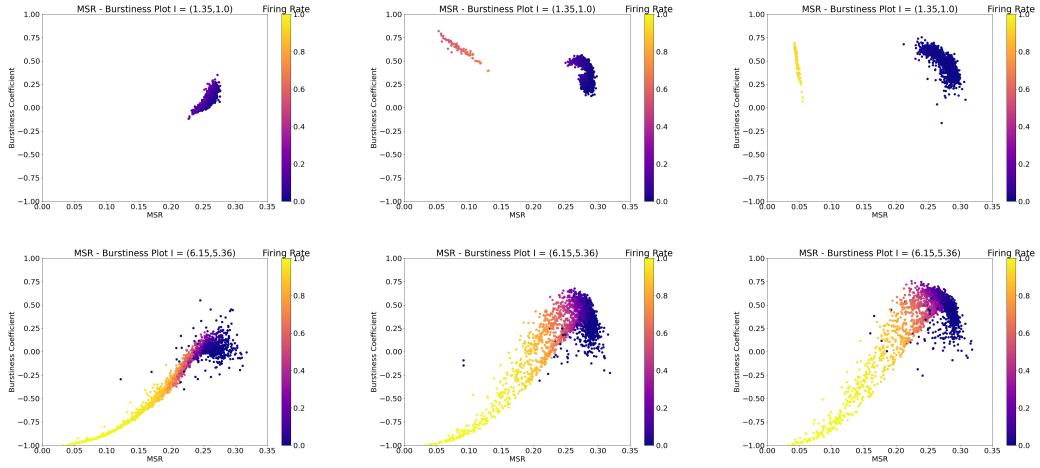


Figure 6.14: In the top panel, the relationship between MSR and the burstiness coefficient for $(i_E, i_I) = (1.35, 1.00)$ is almost linear. Instead, as q transitions towards values peculiar to higher topological randomness, the relation becomes a cloud of points with opposite slope, thus hinting at a possible switch in the correlation at a specific q-value. In the bottom panel, presenting results for the external current condition $(i_E, i_I) = (6.15, 5.36)$, we observe that, in general, the relationship between MSR and Burstiness remains linear across topological conditions. However, for $q = 0.5 \wedge q = 1.0$, the relation peaks and then decreases, thus determining an MSR range where units characterized by the same B-coefficient (hence same inter-events time variance) produce different $\mathcal{R}_t$ values.

Starting from the case $q = 0.1$, we can see how for the external condition $(i_E, i_I) = (1.35, 1.00)$ the relationship between MSR and the coefficient of local variation is almost linear and characterized by a low firing rate, thus hinting at the fact that in this regime the two metrics explains the same features of the spike train. Accordingly, although constrained to a much smaller range, we can observe the same kind of linear relationship for MSR and burstiness. Interestingly enough, both the coefficient of local variation and the burstiness coefficient hinge on the variance intrinsic in the inter-spikes time series, and for a mildly excited network this seems to coincide with what MSR considers relevant across multiple time-scales. As soon as we increase the external current, and subject the network to a strong-to-saturating current, the L_V maintains a linear relationship with MSR only in the interval $\mathcal{R}_t \in [0.15, 0.30]$, while most of the units with $\mathcal{R}_t \in [0.00, 0.15]$ are characterized by degeneracy in a very narrow strip of values of the comparison metric. Similarly, we can observe how the Burstiness coefficient B relates linearly to MSR in the interval $\mathcal{R}_t \in [0.05, 0.30]$, while a degeneracy over a small strip of B values appears for units with $\mathcal{R}_t \in [0, 0.05]$. By looking closely, it is possible to observe how the region of degeneracy is the same that is characterized by relatively high firing rates. Thus, it could be argue that, by virtue of this first observation, much of the statistical information contained in densely bursting signal is lost when encoded by both the coefficients of local variation and burstiness; instead, MSR with its spectrum-like time-scale approach, discerns variations that go beyond the simple differences in subsequent inter-spike times or overall variation of the series. Closing narrowly on the correlation between MSR and burstiness for $(i_E, i_I) = (6.15, 5.36)$, we can appreciate how degeneracy is two-faceted, with multiple values of the burstiness coefficient associated to each MSR value, and vice versa. This could be due to the fact that, for a moderately strongly excited network, the two metrics, despite capturing an overall trend (observable from the shape of the plot), encode different aspects of the richness of response coded in a bursty signal. Finally, the memory coefficient remains rather broadly distributed across all currents, as could be expected by a network characterized by localized bump activity where

there is no complete independence from other units' activity (in fact, most of them belong to the same neighbourhood).

Switching now to networks characterized by q-values ($q \in [0.3, 0.8[$) in the small-world range, we can start to observe some subtle and some more evident differences. First of all, we notice that for $(i_E, i_I) = (1.35, 1.00)$ there is still a linear relationship between MSR and both the coefficient of local variation and the burstiness coefficient, but the interval where such a relation holds shrinks to $\mathcal{R}_t \in [0.25, 0.30]$ (neglecting the few outliers characterized by a surprisingly high firing rate for the given external condition). Accordingly, for external currents in the strong-to-saturating regimes the MSR interval characterized by degeneracy over narrow strips widen visibly for the coefficient of local variation, with units characterized by $\mathcal{R}_t \in [0.00, 0.20]$ collapsing on a few L_V values, and imperceptibly for the burstiness coefficient, with all units characterized by $\mathcal{R}_t \in [0.00, 0.05]$ being associated to $B \approx -1$ - hence to the region of almost homogeneous bursty signals. Interestingly, by looking at the plot for the relation between the MSR and the Burstiness coefficient, we can observe how the region of highest MSR scoring units, which coincides with the one where units have the lowest firing rate, starts to invert the linear trend. Additionally, it can be observed that the profile of each the $L_V - \mathcal{R}_t$ plot is significantly "tighter", with only a few comparison metric values associated to each MSR point; consequently, we can assert that MSR and the coefficient of local variation encode, across most of their entire range, the same statistical information of the bursty signal. Finally, and coherently with the known theory, most of the memory values for small-world networks cluster more closely to $M = 0$, as could be expected from an ideal network having $N \to +\infty$ and fraction of independent inputs proportional to the q-parameter; indeed, if inputs were fully independent, there would be no correlation among subsequent spikes.

At last, by considering networks characterized by $q \in [0.8, 1.0]$, hence in a range where the network topology is almost or totally random, we see that for $(i_E, i_I) = (1.35, 1.00)$ the coefficient of local variation maintains an almost linear relationship with MSR, although

the positive slope seems to increase with respect to the same in the other q-conditions. Instead, while also the burstiness coefficient still maintains a linear relationship with MSR, the slope of the relationship has become negative, hence it has inverted with respect to the condition $q \in [0, 0.3[$. As the external current to the network increases, we observe from the plot of the correlation between the MSR and the coefficient of local variation that, while the interval of degeneracy for $\mathcal{R}_t$-values onto L_V-values remains unvaried with respect to the small-world regime, in the high MSR range the relation between the two coefficients seems to be almost bijective and linear, thus hinting on the fact that, for random networks and external currents in the strong-to-saturating regime, $\mathcal{R}_t$ and L_V encode the same signal features. Surprisingly, we loose almost all the degeneracy of $\mathcal{R}_t$ onto B, with the relationship being linear for $\mathcal{R}_t \in [0, 0.25]$. Importantly enough, for units characterized by $\mathcal{R}_t > 0.25$ the inversion in the linear trend is now evident, more pronounced than the same for small-world networks, and coherent with what was observed for the initial external current. For what concerns the shape of the trend itself, it is still fairly spread, so that arguments of bijectivity will not hold. To conclude, for most of the currents the memory coefficients is tightly clustered around $M = 0$, confirming once again the network's drift towards a couplings configuration characterized by independent inputs.

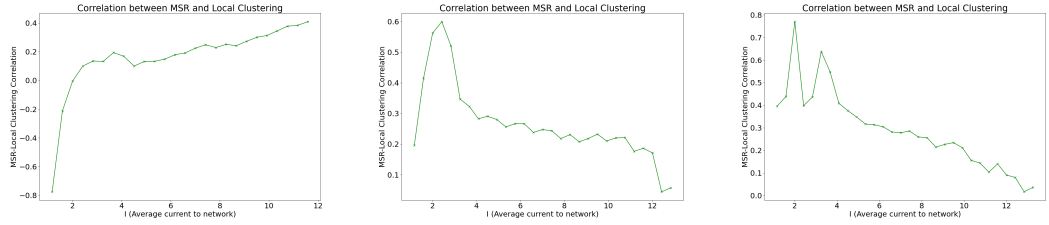
Given all these elements, a final remark on the relation between MSR and the burstiness coefficient is due. Indeed, it seems that for highly regular networks, the higher is the variance in the series of inter-event times, the higher will the respective MSR score be. Instead, as we increase the randomness of the network topology, for units with $\mathcal{R}_t > 0.25$ the trend inverts, and the higher is the variance in the series of inter-event times, the lower will the respective MSR scores be. A possible reason could be that, in topological states of high regularity where the associated activity can be expected to be to some degree correlated and localized, the variance in the inter-event times has an upper-bound coherent with the existence of the positive linear relationship with MSR. Conversely, in topological states of high randomness, given the absence of correlations and the naturally higher burstiness scores, it may be that a

maximality condition on B is reached

$$\tilde{B} = max_{\mathcal{R}_t}\, B(\mathcal{R}_t) \approx 0.7 \tag{6.1.3}$$

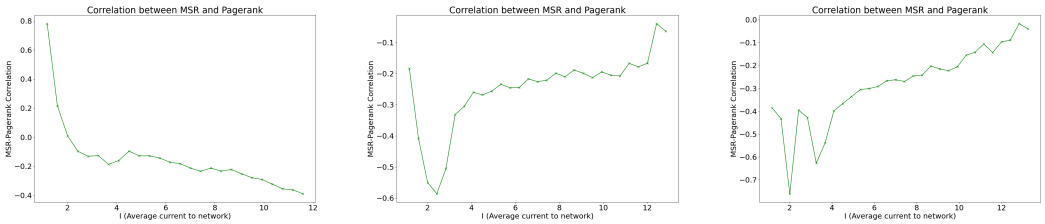
Thereby, we can conclude that spike trains presenting the highest richness of response according to the MSR metric are not those that also display the highest variance in inter-events times. The pressing question now is what determines differences in MSR score between units that are characterized by the same burstiness coefficient $B = \bar{B}$, where we are considering a small-world or random network and $\bar{B} = \tilde{B} - \delta$, for $0 < \delta << 1$.

Now that we have studied how MSR correlates with other statistics of the spike train, and since we have abundantly explored their dependence on the firing rate, it would be interesting to understand whether there are some graph theoretical measures that allow us to consistently predict the firing rate for each single unit, and by virtue of the MSR dependence on it, if there are some of them that consistently correlate with MSR, across currents. The results, for the same configurations of q-values, were the following (for a complete outline, see Appendix 8).



(a) $q = 0.1$ **(b)** $q = 0.5$ **(c)** $q = 1.0$

Figure 6.15: **(a)** In the condition $q = 0.1$, the correlation between the weighted, directed clustering coefficient and MSR increases with the external current, very rapidly for low ones and then gradually for higher ones. **(b)** In the condition $q = 0.5$, we have again a very sharp increase in the correlation for low external currents, followed by a peak and then a gradual decrease. **(c)** In the condition $q = 1.0$, the increase in correlation at low external currents is again abrupt, and differently from (b) the peak is followed by sharper decrease. It is worth noticing that the only condition where the correlation continues to increase is the one where the network is characterized by an almost ring topology; it may be possible that, given the high topological regularity of the network, subtle variations in the clustering coefficient significantly impact the MSR value of the units. Such advantage is lost as the topological randomness increases, and the perturbations are quickly relayed throughout the entire network.



(a) $q = 0.1$ **(b)** $q = 0.5$ **(c)** $q = 1.0$

Figure 6.16: **(a)** In the condition $q = 0.1$, the correlation between pagerank and MSR is characterized by a sudden decrease to $\rho_r = -0.2$ for relatively low currents, and promptly followed by a more gradual decrease as the external current continues to increase. **(b)** In the condition $q = 0.5$, the correlation is characterized by a sharp decrease for relatively low external currents, followed by a plateau and a subsequent gradual increase. **(c)** In the condition $q = 1.0$, the correlation is characterized by a steep but controlled decrease for low currents, followed by a plateau and a sharp increase. Given the recursiveness with which pagerank is defined, results (b) and (c) may highlight how nodes that are highly referred to by other pagerank relevant nodes in the network are particularly inefficient at the generation of meaningful (according to MSR) signals for relatively low currents.

As evident from the plots 8.46,8.49, 8.52, 8.55 and 8.58, there is no significant correlation between MSR and the graph measures of Betweenness, Closeness, In-degree, and Katz centrality, and this absence holds across all the external currents the network was subjected to. Now, given the absence of correlation between any of these measures and MSR, and since MSR depends on firing rate, we can infer that there will be no correlation between any of these measures and firing rate. This is somewhat surprising, as Fletcher et. al. [15] found that neural activity correlated strongly with Katz centrality, and consequently the measure could be used as a good predictor of the same. However, we have to consider the fact that they created a network of LIF neurons wired according to a random topology; this could have significantly contributed to the observed differences, as the dynamics for each of the involved units are significantly more sophisticated. Thereby, since in our network the activity of units depends in a step-like manner from internally incoming input, rather than differentially, the evaluation of the weighted impact of nodes k-hops away from the unit of interest could bear little or no relevance to the actual determination of its firing activity, since single units activation is highly stochastic and dependent on small fluctuations, as we have seen in chapter 3. On the other hand, the MSR of each unit is significantly correlated with its

weighted local clustering coefficient and anti-correlated with its pagerank coefficient, where the former is defined as

$$C_i^J = \frac{\sum_{j \in V_i} \sum_{k \in V_i \setminus \{j\}} J_{ij}^{\frac{1}{3}} J_{jk}^{\frac{1}{3}} J_{ki}^{\frac{1}{3}}}{d(v_i)(d(v_i)-1) \cdot max_{lh}\, J_{lh}} \tag{6.1.4}$$

Interestingly enough, these correlations behave in a drastically different way as they evolve across currents for networks with different degrees of randomness in their connectivity. Specifically, for networks defined by $q \in [0.0, 0.2[$, and therefore by considerably high clustering for all the units in the network (ring topologies with few random connections), the local clustering coefficient strongly and inversely correlates with MSR for low-currents, with a rapid correlation change driving it towards positive values. Indeed, as the current increases, so does the correlation, reaching a maximum of $\rho_r(\mathcal{R}_t; LC) \approx 0.4$. Conversely, the correlation between MSR and Pagerank starts at $\rho_r(\mathcal{R}_t; PGRK) = 0.8$ for low currents and rapidly decreases to a minimum of $\rho_r(\mathcal{R}_t; PGRK) \approx -0.4$. If we consider instead networks defined by $q \in [0.3, 0.8[$, hence small-world networks tending towards total random connectivity, we have that the local clustering coefficient significantly correlates with MSR - $\rho_r(\mathcal{R}_t, LC) \in [0.4, 0.7]$ - for low external currents, i.e. $(i_E, i_I) \in [1.00, 5.00]$, with a steep initial increase, and then gradually decreases to $\rho_r(\mathcal{R}_t, LC) = 0$ as the current increases. Conversely, the correlation between MSR and pagerank attains significant negative values for low external currents - $\rho_r(\mathcal{R}_t; PGRK) \in [-0.7, -0.4]$, and then gradually increases to $\rho_r(\mathcal{R}_t; PGRK) = 0$ as the current increases. Finally, for almost or totally random networks with $q \in [0.8, 1.0]$, we observe a steep initial increase to $\rho_r(\mathcal{R}_t, LC) \approx 0.8$, promptly followed by an equally steep decrease to $\rho_c(\mathcal{R}_t, LC) = 0$. In an almost symmetrical way, we can observe a steep decrease to $\rho_c(\mathcal{R}_t, PGRK) \approx -0.8$ for low external currents, and then a quick increase to $\rho_c(\mathcal{R}_t, PGRK) = 0$. In order to better understand the correlations, we focus on snapshots of the points relation between conditions of highest correlation change, and on a snapshot for the points relation in a strong-to-saturating current condition.

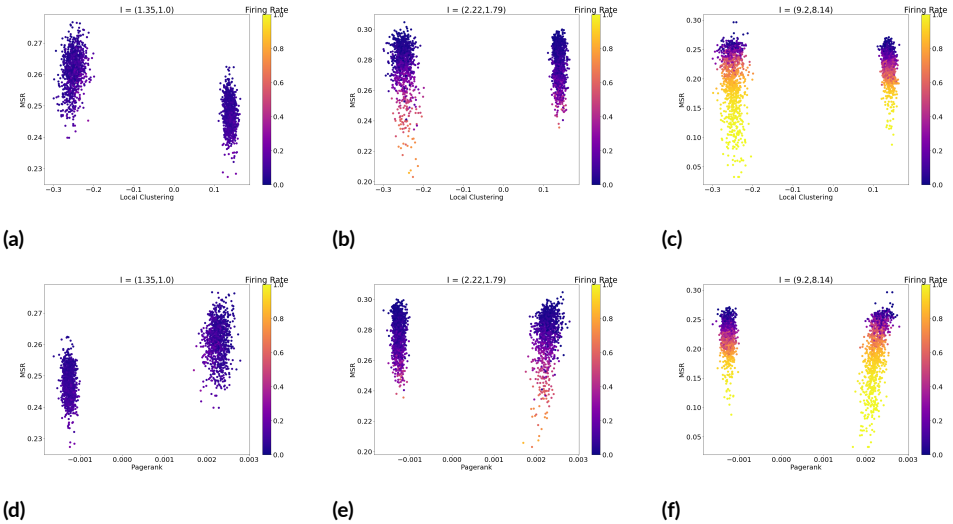


Figure 6.17: The top row (figures (a),(b),(c)) displays the point relation between Local Clustering and MSR for $q = 0.1$. From (a)$\longrightarrow$ (b) we can appreciate the steep increase in correlation, that leads it from strongly anti-correlated (a) to not correlated (b). Figure (c) instead presents a condition of positive correlation, with most of the high MSR preserving units localized on the cluster characterized by positive local clustering coefficient. The bottom row (figures (d),(e),(f)) displays the point relation between Pagerank and MSR, which follows an exact opposite evolution if compared to Local Clustering. From (d)$\longrightarrow$ (e) there is a steep change from strong positive correlation to no correlation, and in (f) we can appreciate how the highest MSR scoring units are localized to the left cluster (negative pagerank).

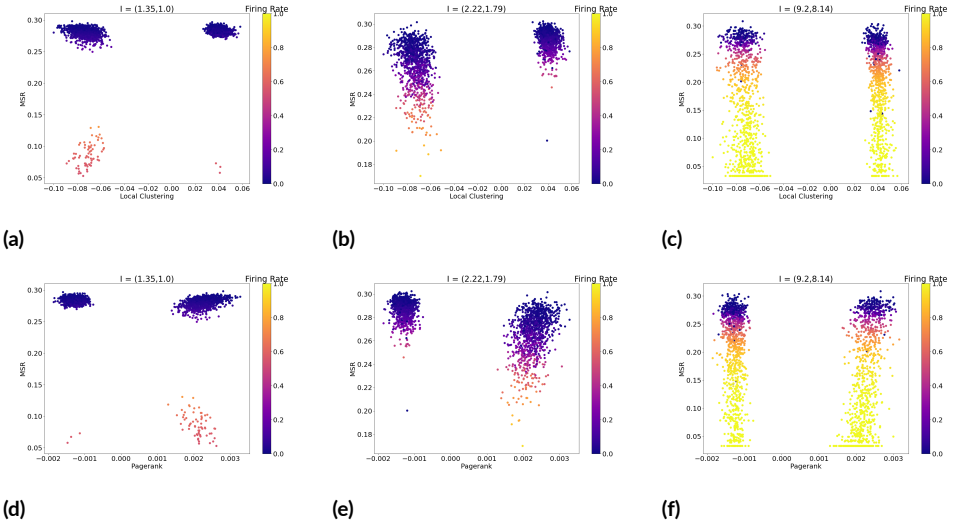


Figure 6.18: The top row (figures (a),(b),(c)) displays the point relation between Local Clustering and MSR for q=0.5. From (a)$\longrightarrow$ (b) a positive correlation change is displayed, with the population on the right (positive local clustering) presenting a fainter decay in MSR scores. Instead, figure (c) again presents a denser cluster of high MSR scoring units for positive clustering coefficients, despite the considerable tail of high firing rate, low MSR scoring units below it. Conversely, from (d)$\longrightarrow$ (e) we observe how the cluster of negatively pagerank scoring units is the one that displays slower MSR decay, while from (f) we see how the column of negative pagerank units is the one preserving the denser cluster of high MSR scores.

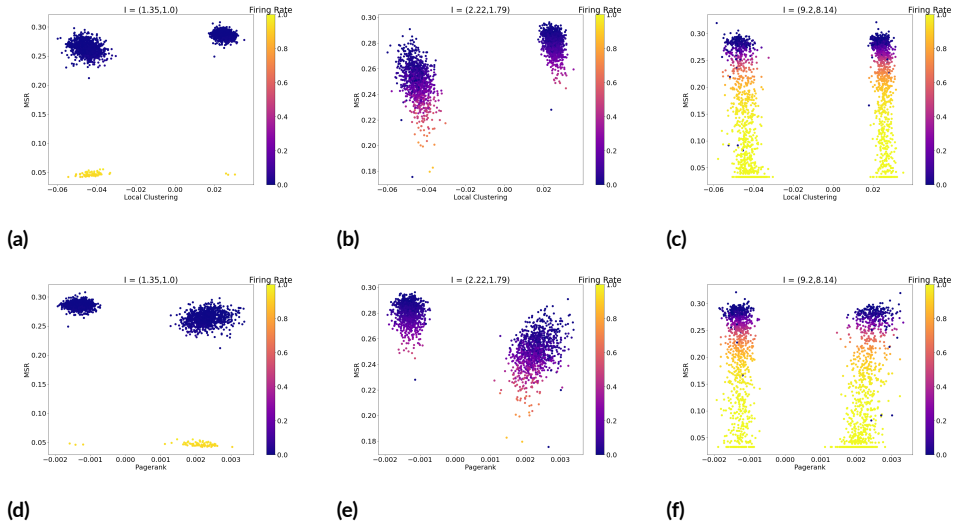


Figure 6.19: The top row (figures (a),(b),(c)) displays the point relation between Local Clustering and MSR for $q = 1.0$. From (a)$\rightarrow$ (b) we can appreciate the strong, positively oriented change in correlation, which has a peak at $\rho_c(\mathcal{R}_t, LC) \approx 0.8$. Figure (c) instead presents how most of the units escaping saturation and preserving high MSR scores are found in the cluster of positive local clustering units. The bottom row (figures (d),(e),(f)) displays the point relation between Pagerank and MSR. From (d)$\rightarrow$ (e) there is a steep negative change in correlation down to $\rho_c(\mathcal{R}_t, LC) \approx -0.8$, and in (f) we can appreciate how the most high MSR scoring, non-saturating units are localized to the left cluster (negative pagerank).

From the graphs of the point relation between MSR and Local Clustering and Pagerank, respectively, we can make some inferences on the identity of the units that preserves high MSR scores across currents. First of all, we notice that the units in the network are subdivided, by each of the fore-mentioned graph measures, into two populations that span intervals of the measure's co-domain characterized by opposite signs, but these intervals are not perfectly symmetric with respect to their zero. As could be expected from the definition of local clustering coefficient and from the way in which we have defined our family of networks, the modulus of the values in the spanned intervals shrinks as $q \rightarrow 1$. On the other hand, and quite interestingly, the intervals spanned by values of the pagerank coefficient remain, in magnitude, approximately unvaried across order conditions, except for the spread of the points around the central interval value, thus highlighting how more randomness in the network associates with more heterogeneity in the flow of activity through nodes. From a global consideration of figures 6.17, 6.18, and 6.19, and drawing a comparison with fig-

ures 6.15 and 6.16, we can observe how the point relation between MSR and the respective graph measures follows a population inverse pattern, thus providing us with a clue that the population that is characterized by positively (negatively) valued local clustering coefficients significantly overlaps with the one characterized by negatively (positively) valued pagerank coefficients. This clue seems to be further supported by an analysis of each population and the firing rate of the respective units, especially for regimes of high external currents. Indeed, by looking at plots (c), (f) of each of the considered figures, we can observe how the density of low-firing, high MSR scoring units is population-symmetric across graph measures - denser for positive local clustering and negative pagerank units - and this difference in density could be one of the driving correlation factors in conditions of high external currents. In addition, notice that this denser population achieves, across all the explored q-conditions, higher MSR values than the sparser one. In addition, from the plots (a), (b), (d), and (e) of each of the considered figures, which present the conditions of steepest correlation change in figures 6.15 and 6.16, we can see how such change always corresponds to an upward push for the population of positively (negatively) local clustering (pagerank) scoring units, thus highlighting once again, even in situations where an analysis of the firing rate is of no aid, how there might be a high degree of overlap between measure-symmetric populations. A quantification of the real degree of overlap between the considered populations would be advisable, but unfortunately unfeasible for the completion of the current thesis, as the cluster on which simulations are run is under maintenance.

On the basis of the definition of the weighted local clustering coefficient and of the weighted pagerank coefficient, and taking into consideration both the y-symmetric trajectory of their correlation with MSR and the hypothesized overlap in populations across graph measures, we propose the under-interacting hub hypothesis. Specifically, we hypothesize that in a network of excitatory and inhibitory units connected by population-specific weights, the units that interact mostly locally (high local clustering) but through which only a small fraction of the total network activity flows (low pagerank) are the ones capable of generating, across

the entire range of currents that the network can withstand without entering a saturation regime, high MSR spike trains.

6.1.1 The Barabasi-Albert Network

The defining feature of the Barabasi-Albert model, for what concerns the goals of the simulation, is the power-law distribution of the nodes' degree. As a result of preferential attachment, we will have nodes that display a broad distribution of degrees, and consequently we cannot expect, on average, the input to each unit to be the same. For this reason, the identity itself of the nodes within each population and across populations becomes relevant. In particular, the property of ergodicity, which is the property of a network where all nodes are connected to other nodes via an existing path, may play a fundamental role in shaping the network response. Indeed, since the Barabasi-Albert network, under the assumption of sparse connectivity, is not supposed to be ergodic, then each different realization of the same network architecture could potentially lead to drastically different results. To evaluate such difference, we will consider two reference realizations of a BA network, where all the parameters are kept the same except for the specific wiring between nodes. Starting from the mean network response to the external input

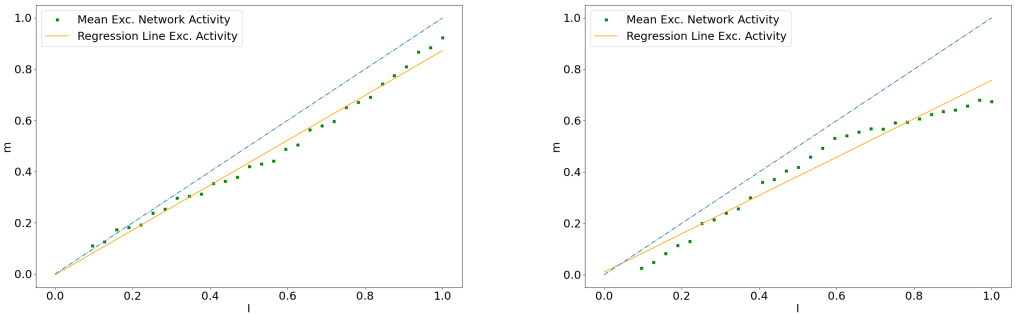


Figure 6.20: Mean response of the excitatory population for two different realizations of the same network

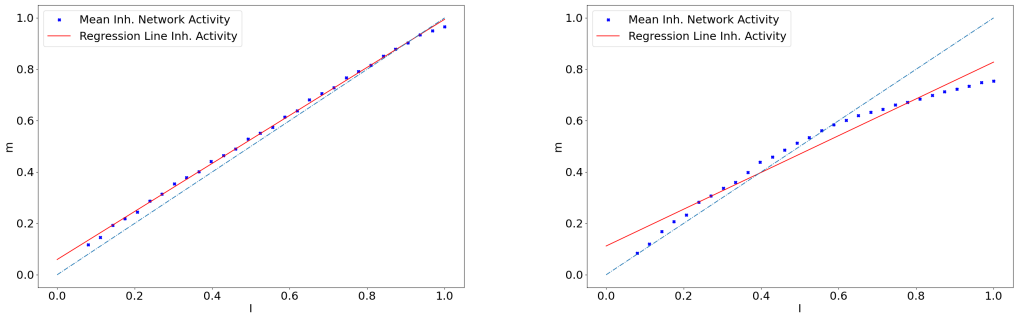


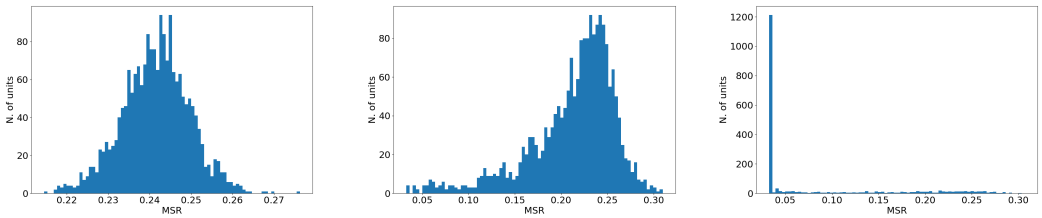
Figure 6.21: Mean response of the inhibitory population for two different realizations of the same network

This difference in mean response was not observed across different realization of the Watts-Strogatz model, for any of the chosen q-parameters, since the model satisfies ergodicity. Moreover, differences were negligible for any of the other considered distributions or correlations. Thus, this subsection will be mainly devoted to the presentation of how different realizations of the BA model may differ, and the possible solutions to the problem.

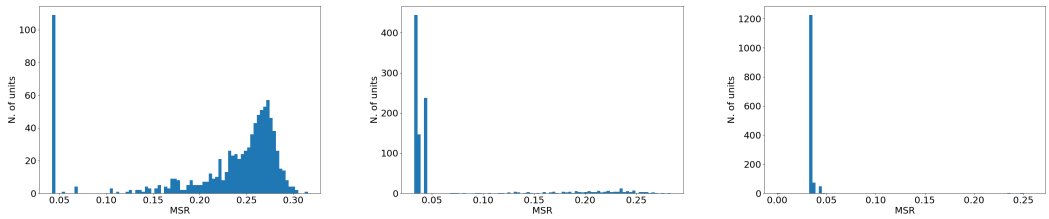
Starting from the MSR distribution for the external current conditions

$$(i_E, i_I) \in \{(1.35, 1.00), (6.15, 5.36), (12.26, 10.91)\}$$

- displayed in the given order - we have



MSR distribution for the first BA network realization.



MSR distribution for the second BA network realization.

and the $3D$-plots for $log(M) - \mathcal{R}_t$, displayed for any given current

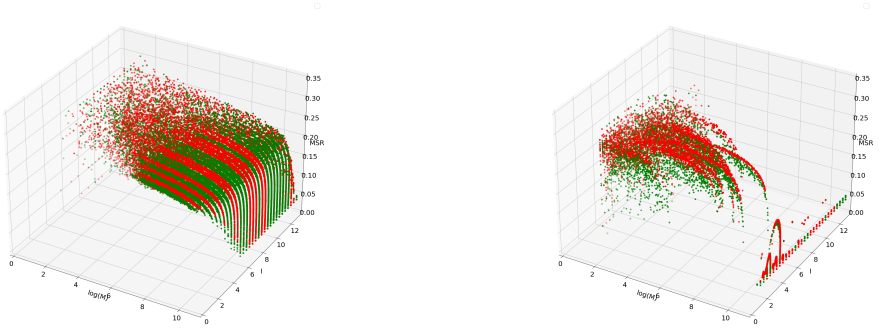


Figure 6.24: $log(M) - \mathcal{R}_t$ plots for any given current for the first BA realization (left) and the second (right).

Thus, it will closely follow that all the other results that were presented for the Watts-Strogatz model will vary considerably across different realizations. For this reason, in order to draw valuable conclusions on the MSR properties for signals generated by a Barabasi-Albert model, it would be advisable to consider average properties, where the average is taken across many realization of the same architecture (ensemble averaging).

..It was on Earth that the positronic brain was invented and
on Earth that robots had first been put to productive use...

Caves of Steel, Asimov

7

Conclusion

In this thesis, we have adequately simulated a family of sparse networks characterized by different degree of randomness in their connectivity, without having them to complete any task. Indeed, the only remarking feature of such networks was their ability to quickly track changes in their external input and have an average activity that linearly depended on them. The aim of the thesis was then to compute the metric Multiscale Relevance on the output of each single unit, and study its correlation with other metrics or graph measures, in order to gain a more thorough understanding of its interplay with different features of the network and of its activity.

The results that were found led us to conclude that the Multiscale Relevance metric depends indirectly on the topology of the network and that it encodes features of the signal similar to the ones considered by the Coefficient of Local Variation and the Burstiness Coefficient. The presence of intervals of linear and non-linear relationship between such metrics urges us to draw the reader's attention towards specific ranges of sensitivity. Indeed, in inter-

vals where the relation between MSR and such metrics is linear, and keeping in mind that such relation does depend on the external input, we are led to believe that the two are essentially equally valuing the properties of the considered signal. Instead, in intervals where such relation is non-linear, we are prompted to consider either one of the two metric (MSR and either the coefficient of local variation or the burstiness coefficient) as providing more information on the properties underlying the considered signals, although what this information is and what is its importance still remains an open problem.

In addition, we found that among a set of graph measures MSR correlates significantly, and depending on the external current, with the local clustering coefficient and the pagerank coefficient. While with the former the correlation is mostly positive, with the latter is mostly negative and the correlation evolution across external currents for the two measures follows a y-symmetric trajectory. This led us to formulate the under-interactive hub hypothesis, according to which, for our family of networks, the units that seem best positioned to produce signals resulting in high MSR coefficients, across currents, are those characterized by relatively high fractions of local interactions (high clustering coefficients) but relatively low flow of total network activity through them (low pagerank). This result seems to hold consistently for our cortical model, which is very simple and excludes from any internal dynamic of the unit. Consequently, it would be interesting to understand whether different types of model cortical networks, such as networks of LIF neurons, preserve this same property or whether it is peculiar to ours'.

Much work still needs to be done to understand Multiscale Relevance more thoroughly, even while constraining ourselves to the same family of networks. For example, this thesis has not addressed the same problem but with weights that are not population specific, but rather drawn from a probability distribution that still allow us to unequivocally distinguish an excitatory and inhibitory population, as outlined in the paragraph 3.4. Furthermore, we could also consider the same type of network, either with population-specific or random weights, but with a time-varying external current, both of deterministic and stochastic nature. Ex-

panding then to other types of network, the same kind of analysis could be run on a network of LIF neurons or, if the reader is interested in even more biologically plausible models, to small networks of Hodgkin-Huxley neurons. All these experiments are considering a purposeless network that simply simulates activity on the basis of some specified dynamics and connectivity. Going a step further, the reader may apply the same analysis on networks of spiking units that perform a relevant task, such as the storing and retrieval of memories, the classification of objects, etc. Experimenting with all this different, well-understood types of network will allow the research community to gain a solid understanding of which are the general properties that relate Multiscale Relevance to network activity, while analysing results that are specific to a particular network will in turn, hopefully, shed more light on the dynamics and properties underlying its global and unit-specific activity.

8 Appendix

8.1 Fixed Weights Network

In this section of the Appendix I will provide all the relevant, generated plots for the distributions and the correlations pertaining to the fixed weights condition of the network. Starting from results on the distribution of the considered spike statistics, and closely following with the correlation between the log of the total number of spikes and MSR, between the different spike statistics, and finally the correlation of MSR with the graph measures detailed in chapter 4

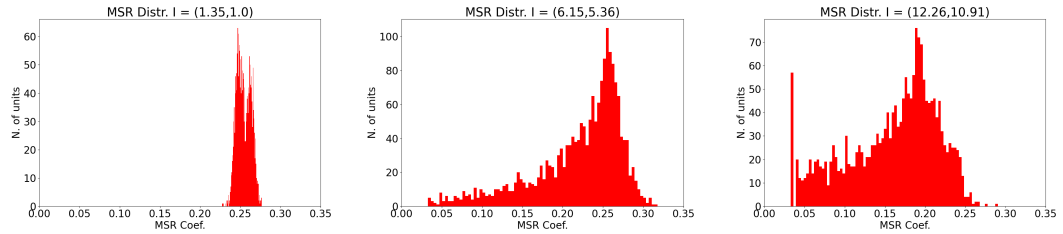


Figure 8.1: MSR Distribution for $q = 0.1$

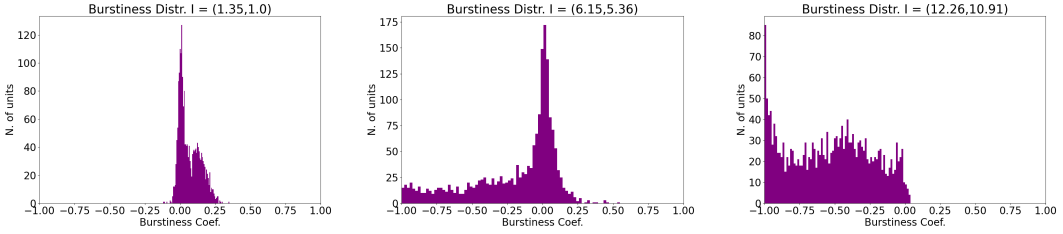


Figure 8.2: Burstiness Distribution for $q = 0.1$

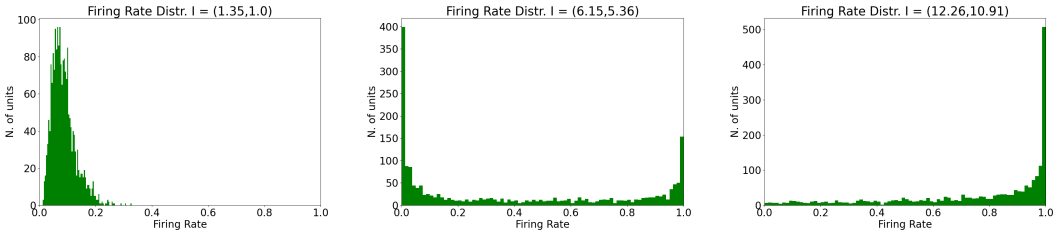


Figure 8.3: Firing Rate Distribution for $q = 0.1$

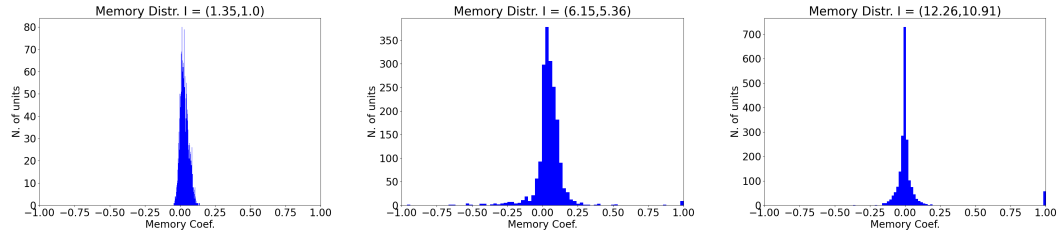


Figure 8.4: Memory Distribution for $q = 0.1$

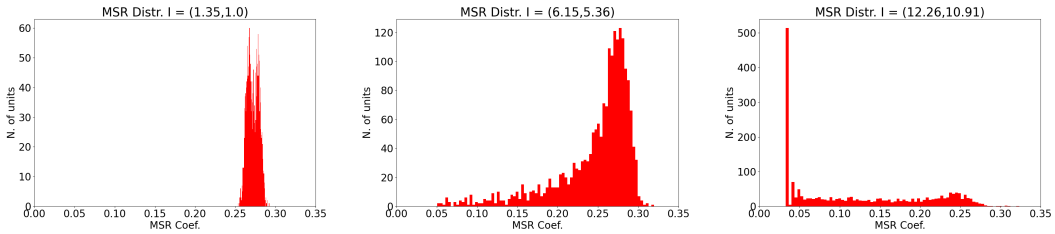


Figure 8.5: MSR Distribution for $q = 0.3$

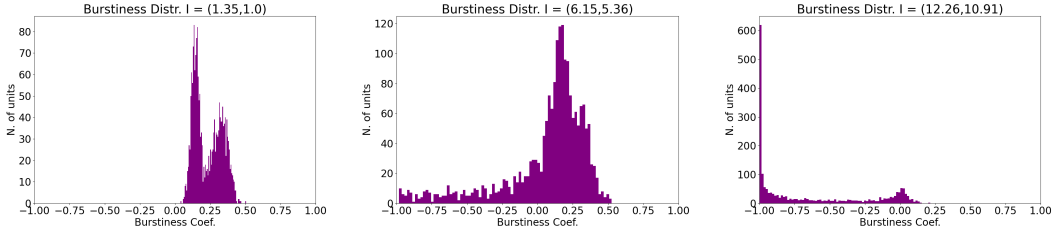


Figure 8.6: Burstiness Distribution for $q = 0.3$

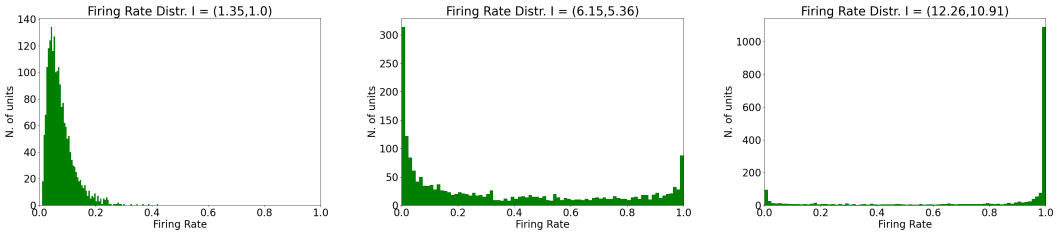


Figure 8.7: Firing Rate Distribution for $q = 0.3$

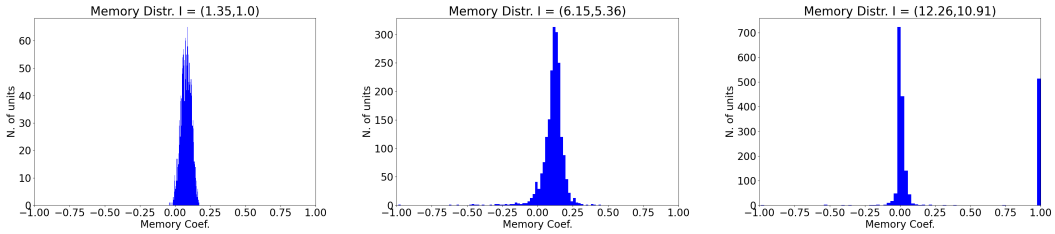


Figure 8.8: Memory Distribution for $q = 0.3$

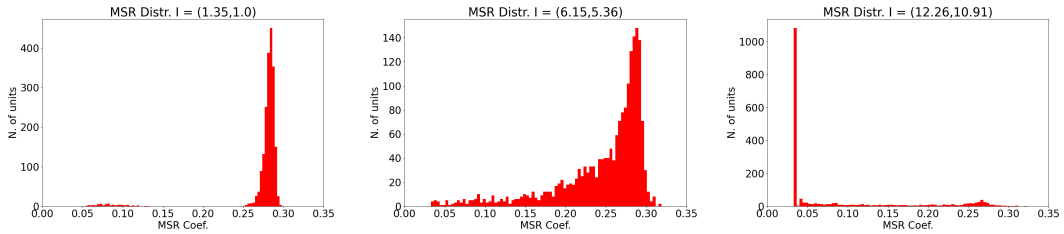


Figure 8.9: MSR Distribution for $q = 0.5$

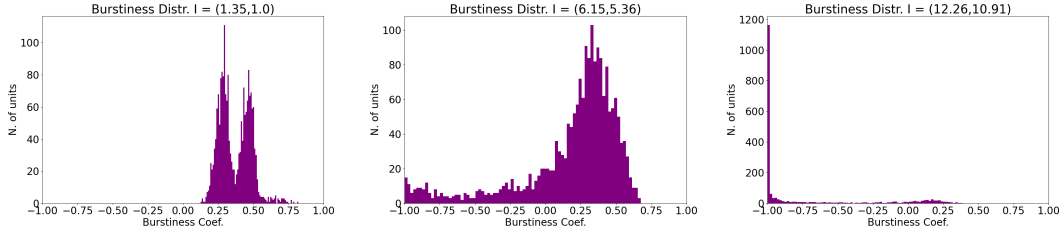


Figure 8.10: Burstiness Distribution for $q = 0.5$

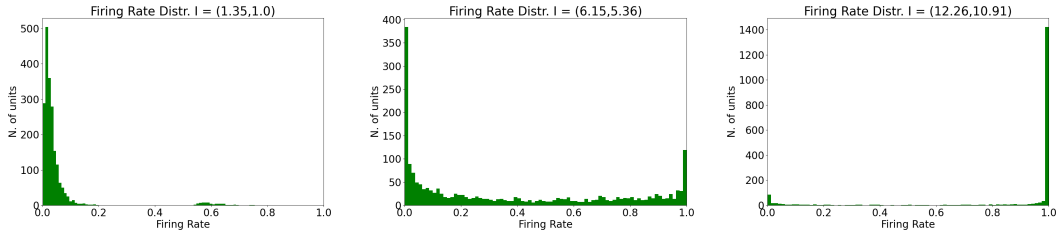


Figure 8.11: Firing Rate Distribution for $q = 0.5$

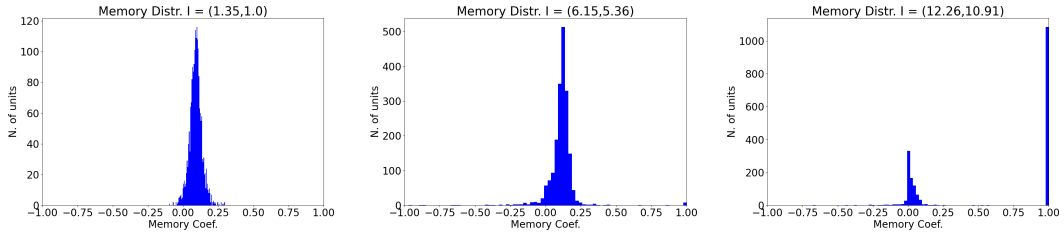


Figure 8.12: Memory Distribution for $q = 0.5$

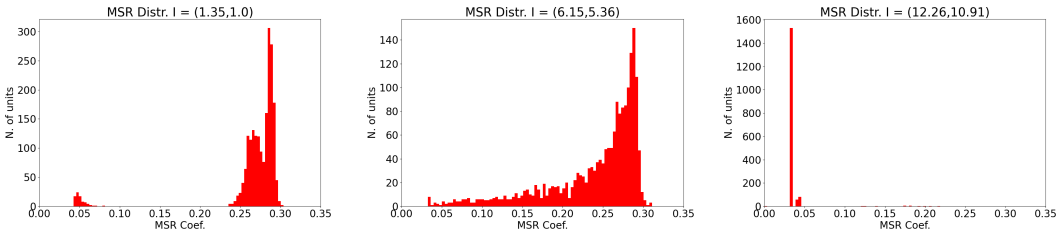


Figure 8.13: MSR Distribution for $q = 0.8$

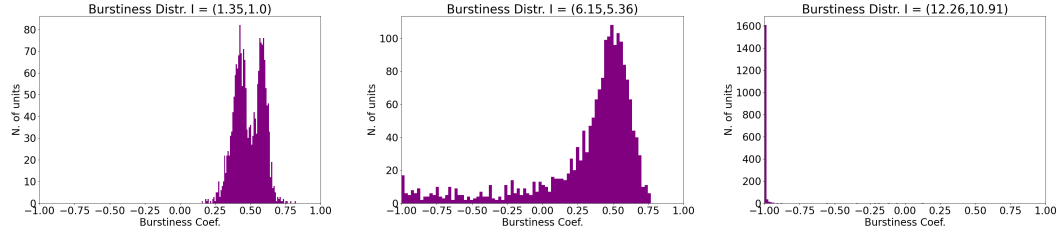


Figure 8.14: Burstiness Distribution for $q = 0.8$

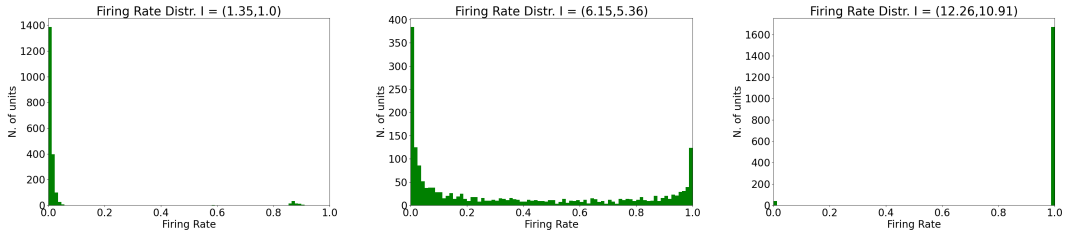


Figure 8.15: Firing Rate Distribution for $q = 0.8$

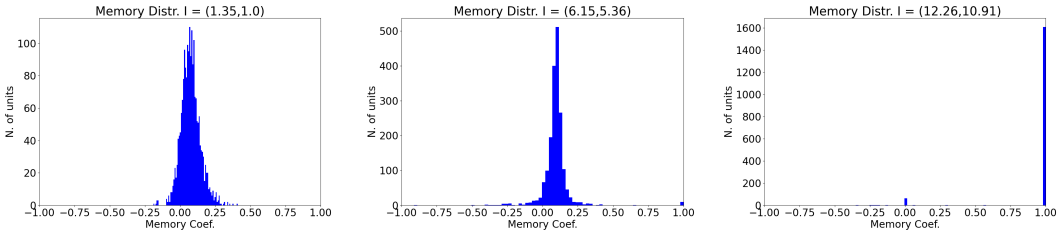


Figure 8.16: Memory Distribution for $q = 0.8$

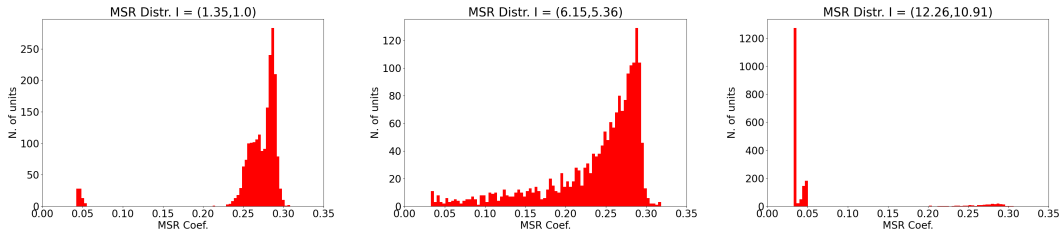


Figure 8.17: MSR Distribution for $q = 1.0$

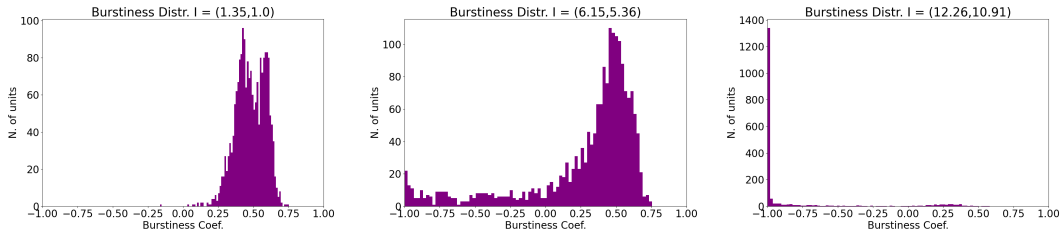


Figure 8.18: Burstiness Distribution for $q = 1.0$

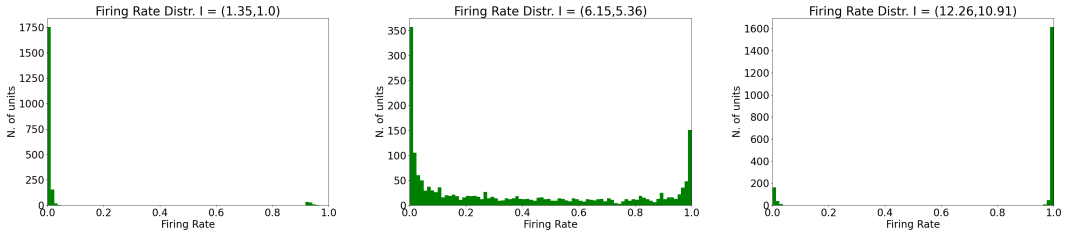


Figure 8.19: Firing Rate Distribution for $q = 1.0$

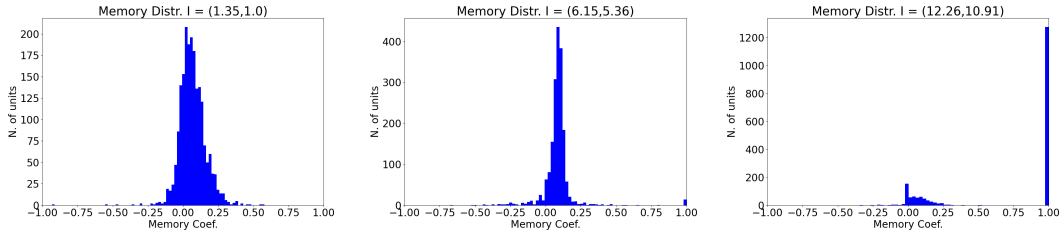


Figure 8.20: Memory Distribution for $q = 1.0$

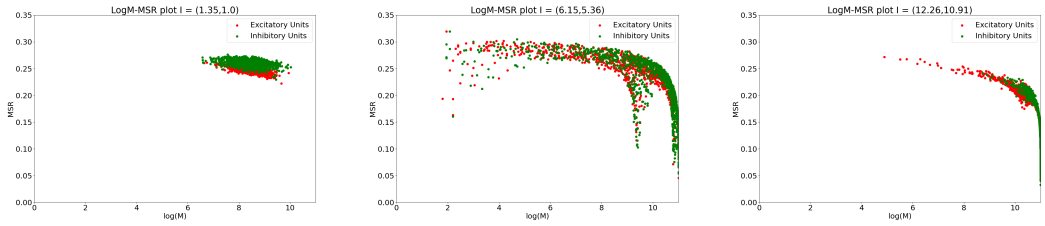


Figure 8.21: log(M)-MSR for $q = 0.1$

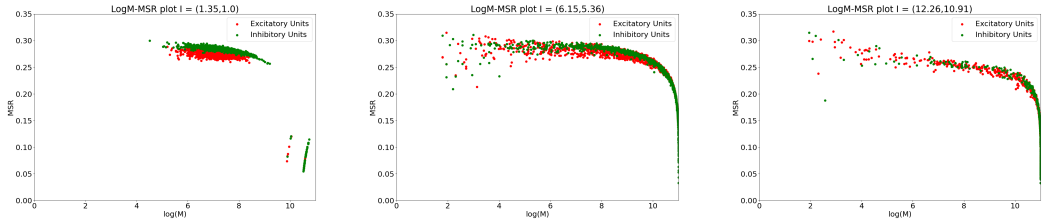


Figure 8.22: log(M)-MSR for $q = 0.3$

Figure 8.23: log(M)-MSR for $q = 0.5$

Figure 8.24: log(M)-MSR for $q = 0.8$

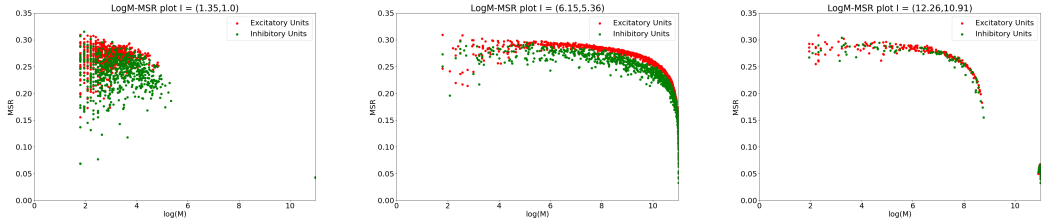


Figure 8.25: log(M)-MSR for $q = 1.0$

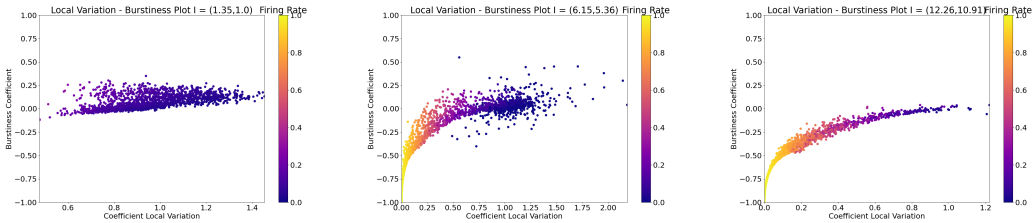


Figure 8.26: Local Variation - Burstiness correlation for $q = 0.1$

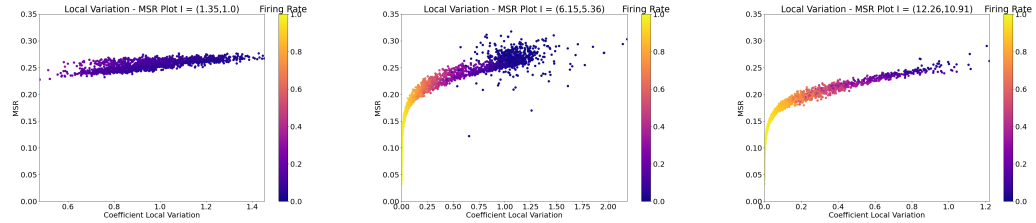


Figure 8.27: Local Variation - MSR correlation for $q = 0.1$

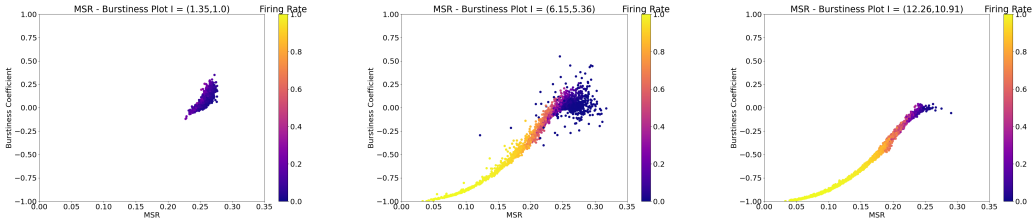


Figure 8.28: MSR - Burstiness for $q = 0.1$

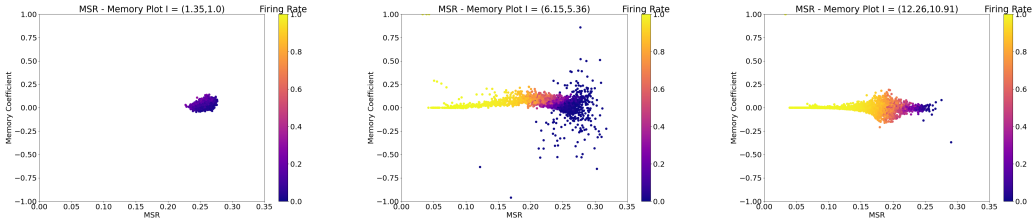


Figure 8.29: MSR - Memory for $q = 0.1$

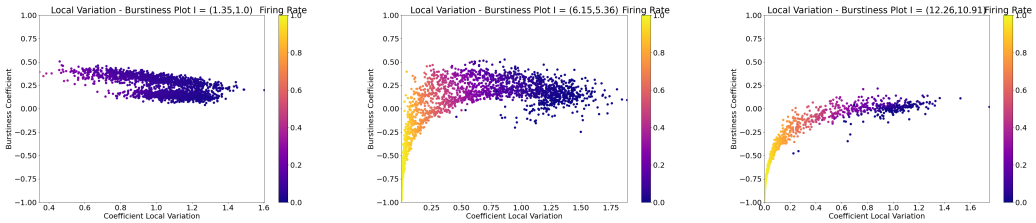


Figure 8.30: Local Variation - Burstiness correlation for $q = 0.3$

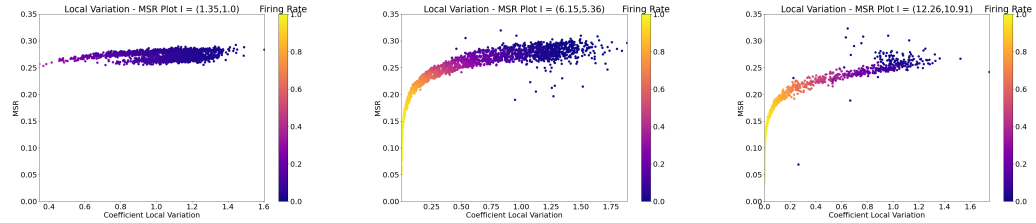


Figure 8.31: Local Variation - MSR correlation for $q = 0.3$

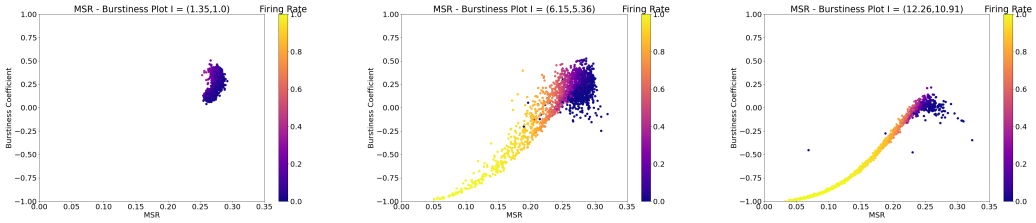


Figure 8.32: MSR - Burstiness for $q = 0.3$

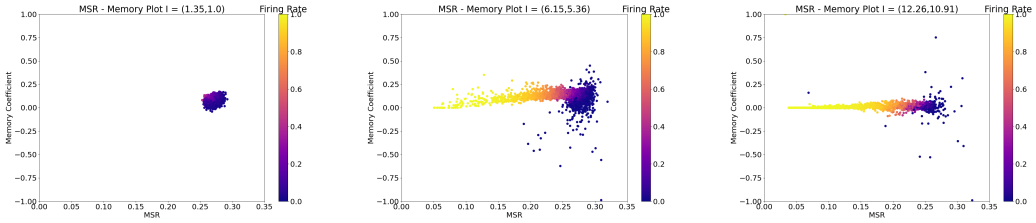


Figure 8.33: MSR - Memory for $q = 0.3$

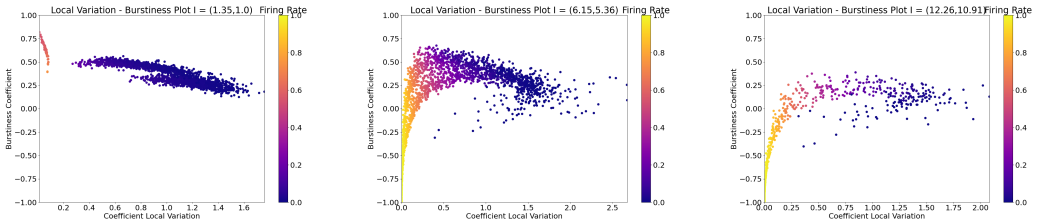


Figure 8.34: Local Variation - Burstiness correlation for $q = 0.5$

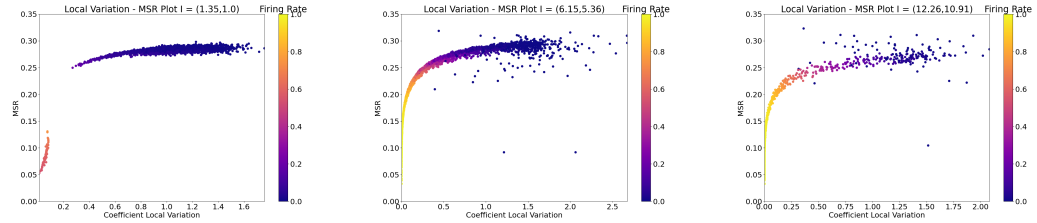


Figure 8.35: Local Variation - MSR correlation for $q = 0.5$

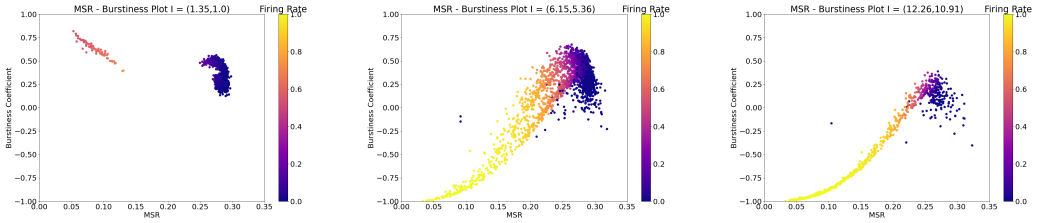


Figure 8.36: MSR - Burstiness for $q = 0.5$

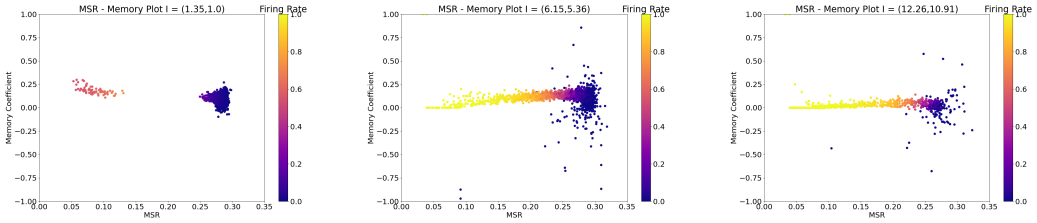


Figure 8.37: MSR - Memory for $q = 0.5$

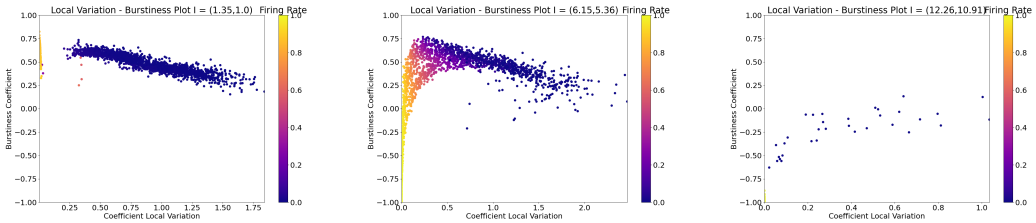


Figure 8.38: Local Variation - Burstiness correlation for $q = 0.8$

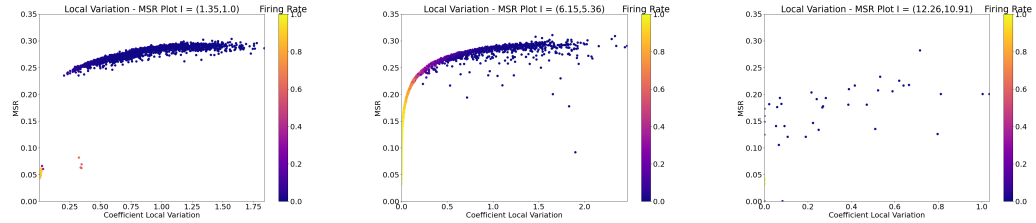


Figure 8.39: Local Variation - MSR correlation for $q = 0.8$

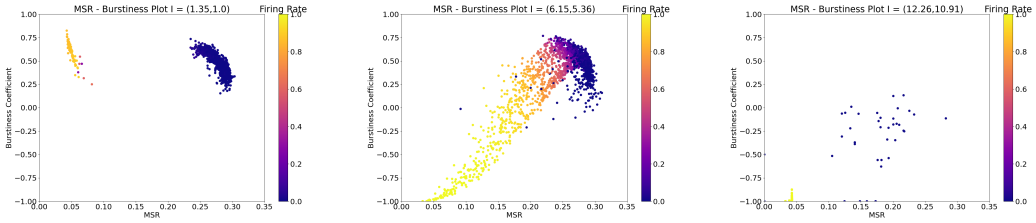


Figure 8.40: MSR - Burstiness for $q = 0.8$

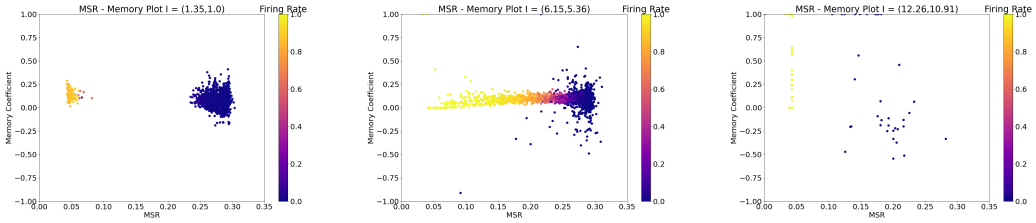


Figure 8.41: MSR - Memory for $q = 0.8$

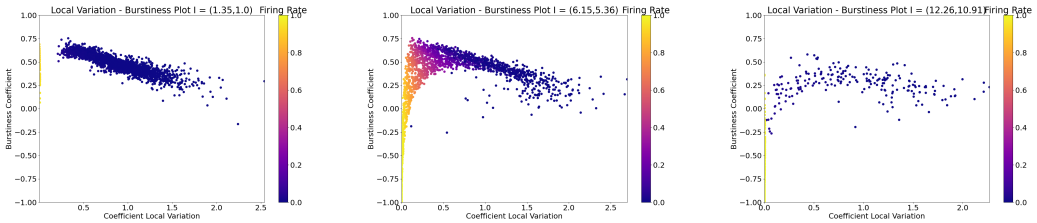


Figure 8.42: Local Variation - Burstiness correlation for $q = 1.0$

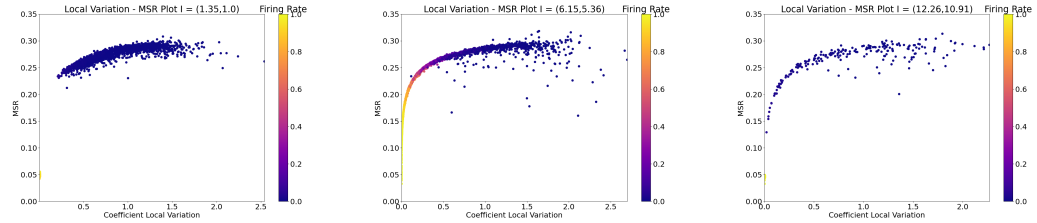


Figure 8.43: Local Variation - MSR correlation for $q = 1.0$

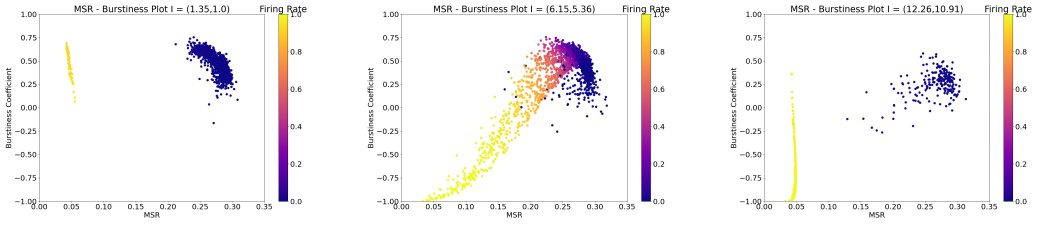


Figure 8.44: MSR - Burstiness for $q = 1.0$

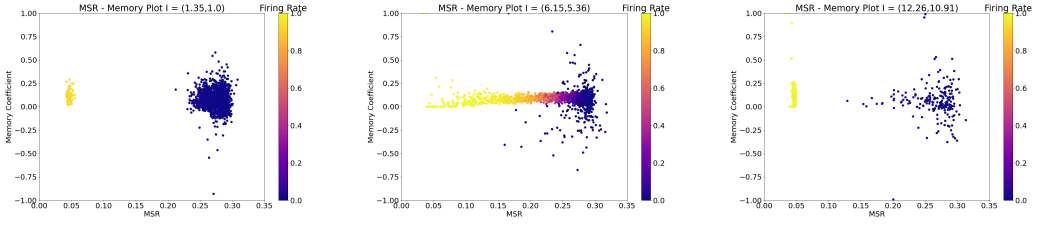


Figure 8.45: MSR - Memory for $q = 1.0$

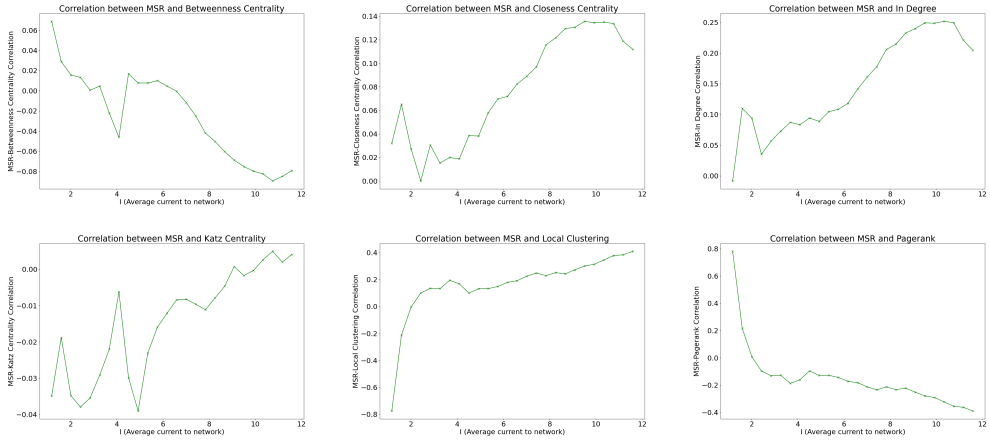


Figure 8.46: MSR-Node Centralities correlation for $q = 0.1$

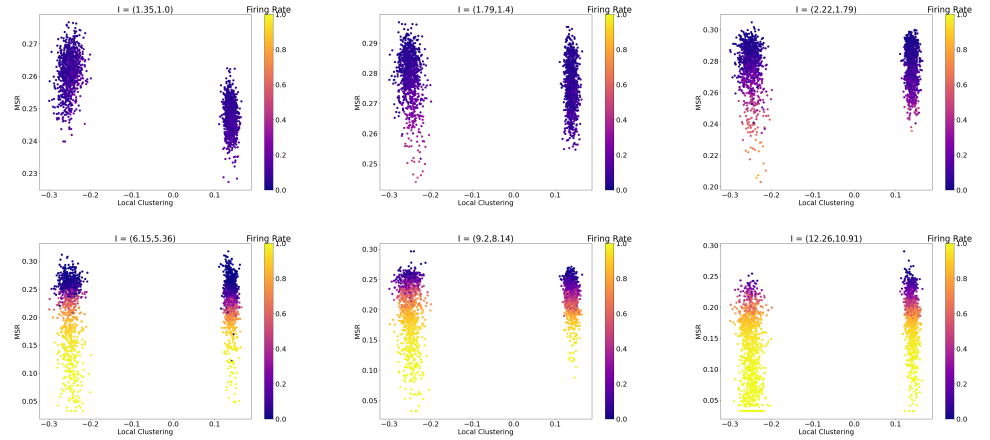


Figure 8.47: MSR-Local Clustering correlation for fixed currents, $q = 0.1$

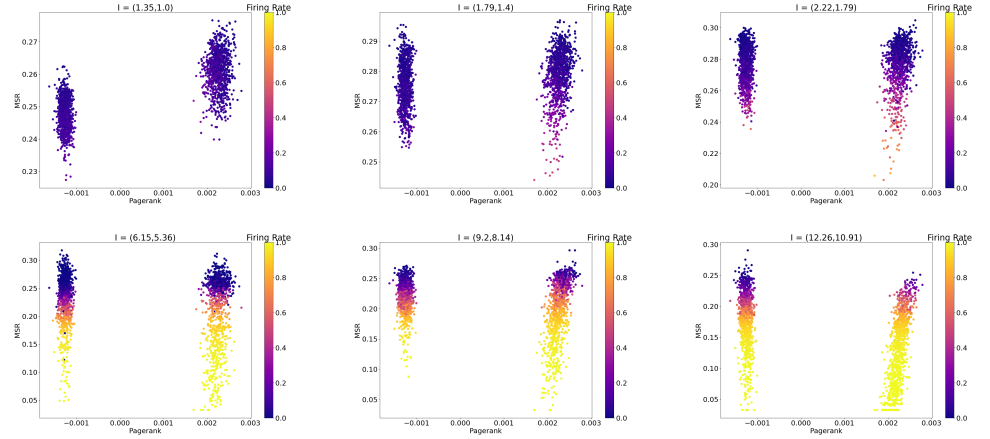


Figure 8.48: MSR-Pagerank correlation for fixed currents, $q = 0.1$

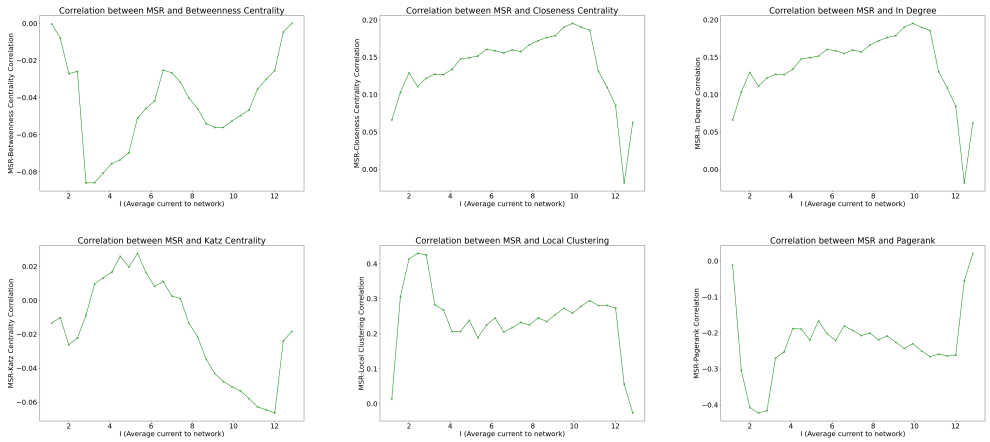


Figure 8.49: MSR-Node Centralities correlation for $q = 0.3$

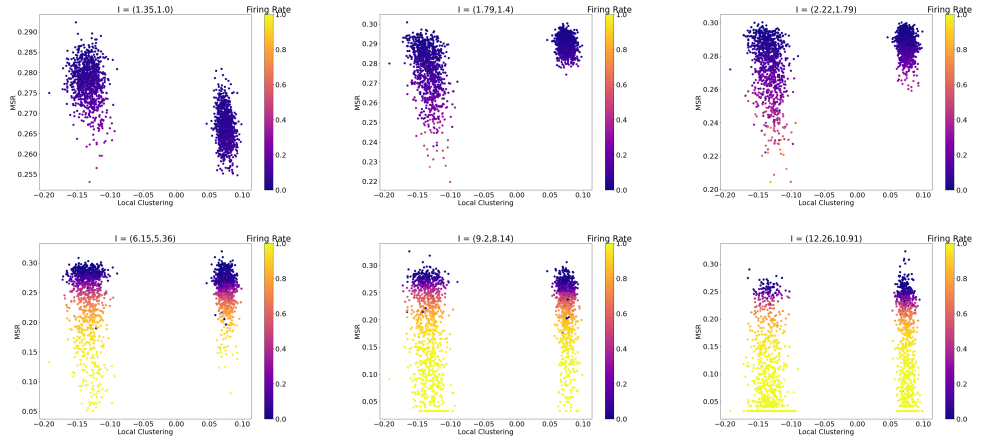


Figure 8.50: MSR-Local Clustering correlation for fixed currents, $q = 0.3$

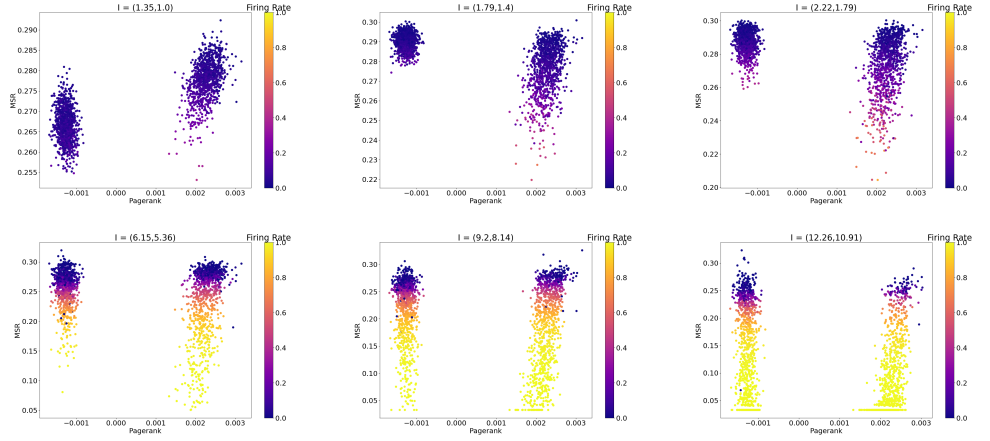


Figure 8.51: MSR-Pagerank correlation for fixed currents, $q = 0.3$

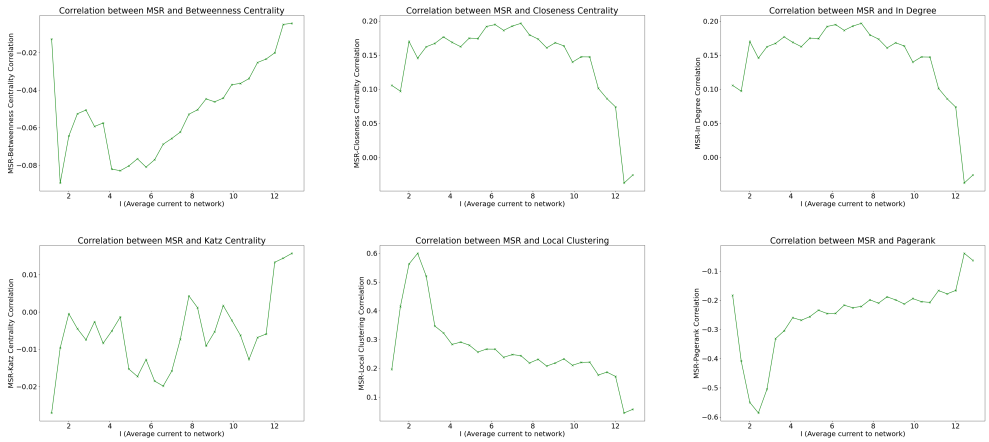


Figure 8.52: MSR-Node Centralities correlation for $q = 0.5$

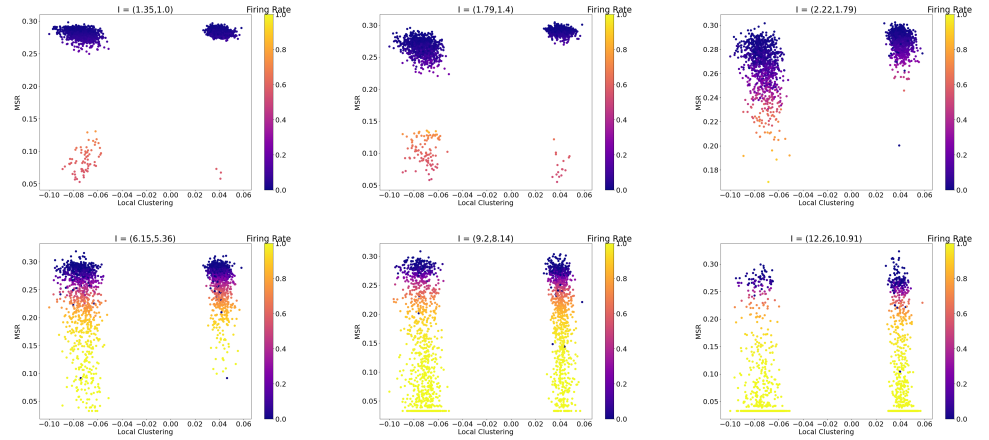


Figure 8.53: MSR-Local Clustering correlation for fixed currents, $q = 0.5$

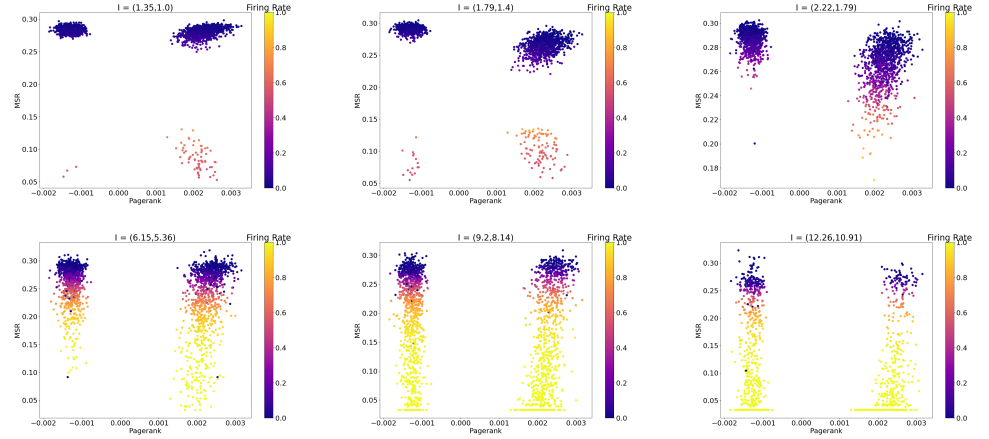


Figure 8.54: MSR-Pagerank correlation for fixed currents, $q = 0.5$

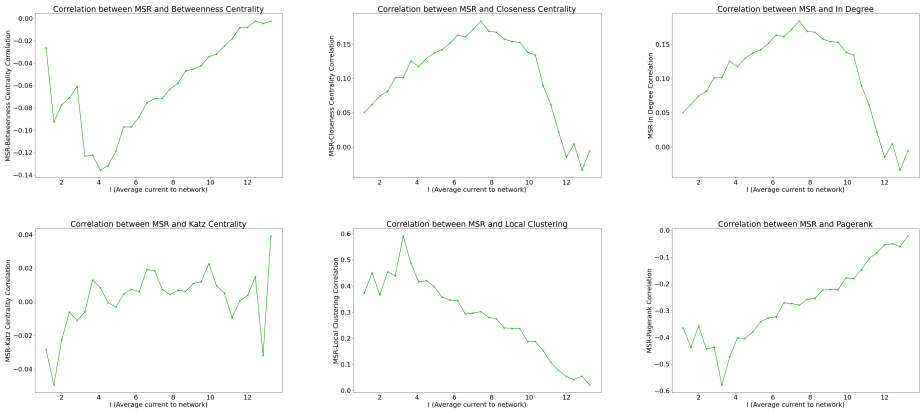


Figure 8.55: MSR-Node Centralities correlation for $q = 0.8$

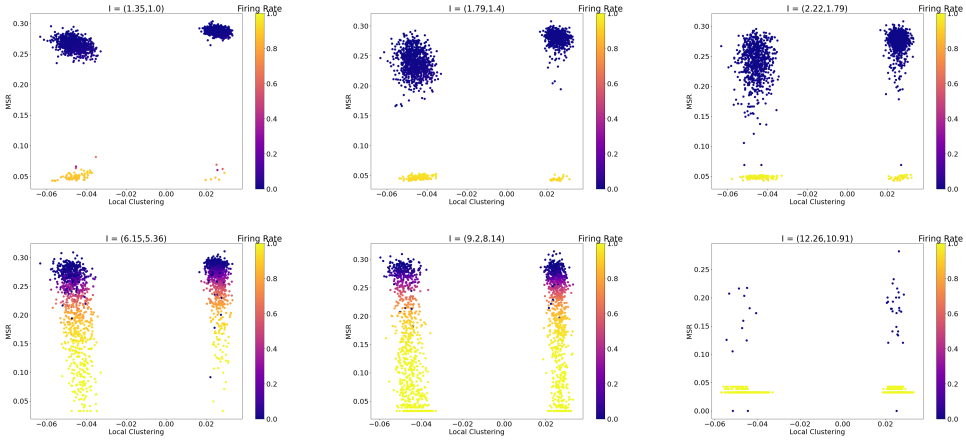


Figure 8.56: MSR-Local Clustering correlation for fixed currents, $q = 0.8$

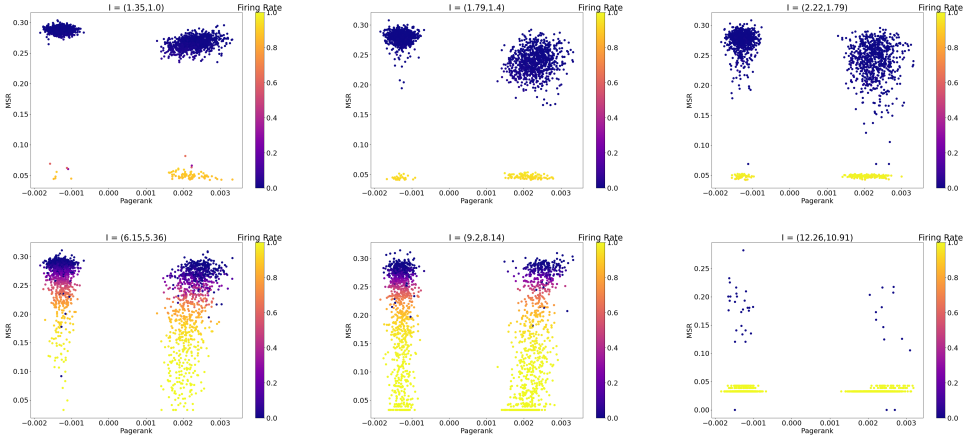


Figure 8.57: MSR-Pagerank correlation for fixed currents, $q = 0.8$

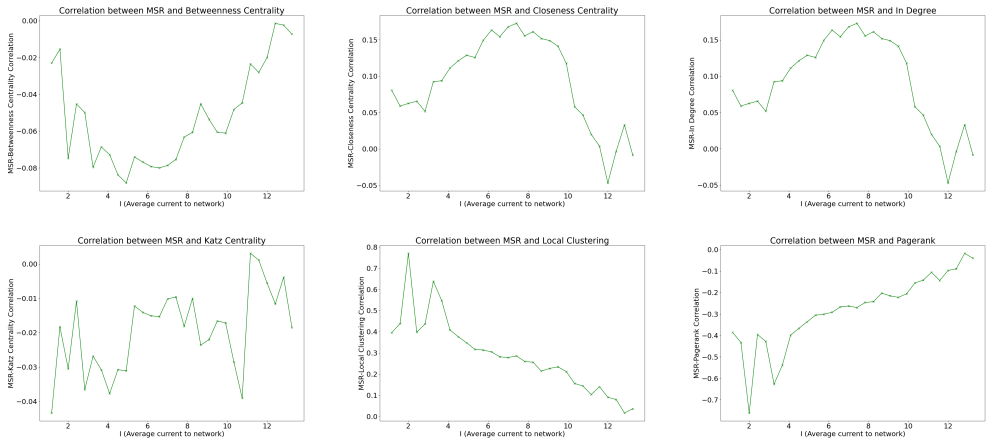


Figure 8.58: MSR-Node Centralities correlation for $q = 1.0$

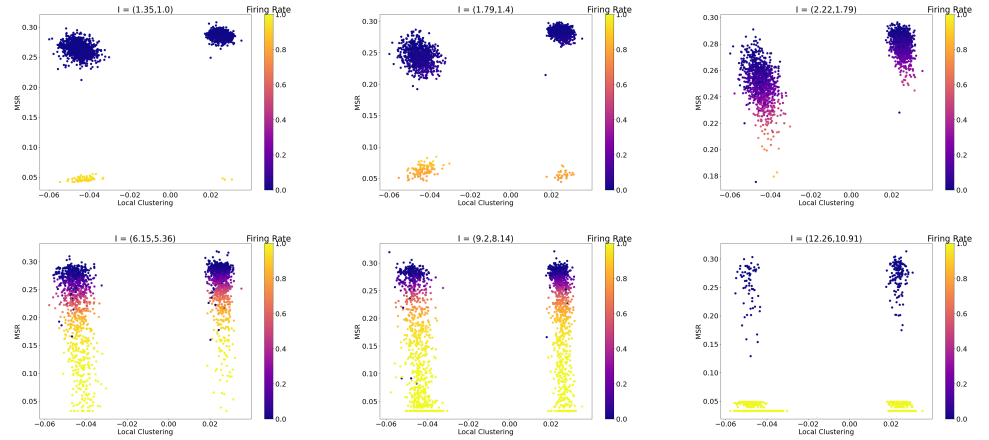


Figure 8.59: MSR-Local Clustering correlation for fixed currents, $q = 1.0$

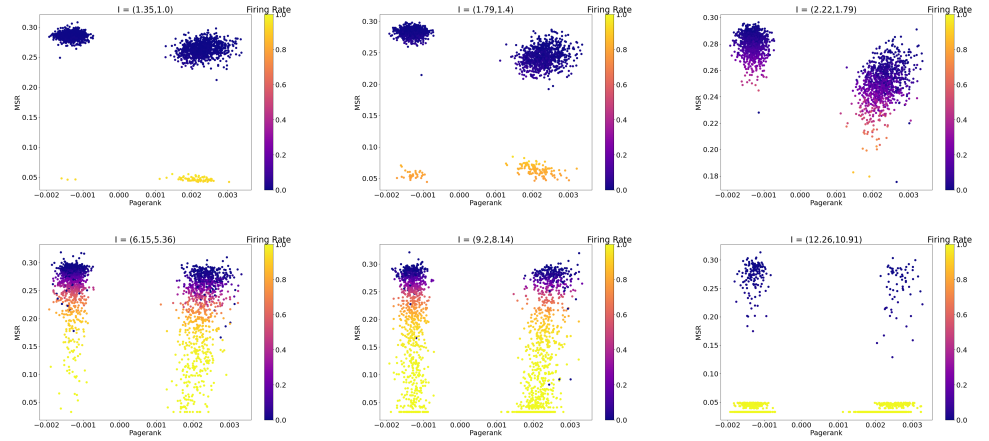


Figure 8.60: MSR-Pagerank correlation for fixed currents, $q = 1.0$

Bibliography

[1] Abeles, M. (1991). *Corticonics: Neural circuits of the cerebral cortex*. Cambridge University Press.

[2] Albert, R. & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1), 47.

[3] Amit, D. J. & Brunel, N. (1997). Dynamics of a recurrent network of spiking neurons before and following learning. *Network: Computation in Neural Systems*, 8(4), 373–404.

[4] Amit, D. J., Gutfreund, H., & Sompolinsky, H. (1985). Spin-glass models of neural networks. *Physical Review A*, 32(2), 1007.

[5] Anishchenko, A. & Treves, A. (2006). Autoassociative memory retrieval and spontaneous activity bumps in small-world networks of integrate-and-fire neurons. *Journal of Physiology-Paris*, 100(4), 225–236.

[6] Ben-Yishai, R., Bar-Or, R. L., & Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. *Proceedings of the National Academy of Sciences*, 92(9), 3844–3848.

[7] Bloch, F., Jackson, M. O., & Tebaldi, P. (2019). Centrality measures in networks. *Available at SSRN 2749124*.

[8] Bruce, A., Gardner, E., & Wallace, D. (1987). Dynamics and statistical mechanics of the hopfield model. *Journal of Physics A: Mathematical and General*, 20(10), 2909.

[9] Burns, B. D. & Webb, A. (1976). The spontaneous activity of neurones in the cat's cerebral cortex. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 194(1115), 211–223.

[10] Cover, T. M. & Thomas, J. A. (1991). Information theory and statistics. *Elements of Information Theory*, 1(1), 279–335.

[11] Cubero, R. J., Jo, J., Marsili, M., Roudi, Y., & Song, J. (2019). Statistical criticality arises in most informative representations. *Journal of Statistical Mechanics: Theory and Experiment*, 2019(6), 063402.

[12] Cubero, R. J., Marsili, M., & Roudi, Y. (2018). Minimum description length codes are critical. *Entropy*, 20(10), 755.

[13] Cubero, R. J., Marsili, M., & Roudi, Y. (2020). Multiscale relevance and informative encoding in neuronal spike trains. *Journal of computational neuroscience*, 48(1), 85–102.

[14] Douglas, R. J., Martin, K., & Whitteridge, D. (1991). An intracellular analysis of the visual responses of neurones in cat visual cortex. *The Journal of physiology*, 440(1), 659–696.

[15] Fletcher, J. M. & Wennekers, T. (2018). From structure to activity: using centrality measures to predict neuronal activity. *International journal of neural systems*, 28(02), 1750013.

[16] Freberg, L. (2009). *Discovering biological psychology*. Cengage Learning.

[17] Gerstner, W., Kistler, W. M., Naud, R., & Paninski, L. (2014). *Neuronal dynamics: From single neurons to networks and models of cognition*. Cambridge University Press.

[18] Ginzburg, I. & Sompolinsky, H. (1994). Theory of correlations in stochastic neural networks. *Physical review E*, 50(4), 3171.

[19] Goh, K.-I. & Barabási, A.-L. (2008). Burstiness and memory in complex systems. *EPL (Europhysics Letters)*, 81(4), 48002.

[20] Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature*, 436(7052), 801–806.

[21] Holt, G. R., Softky, W. R., Koch, C., & Douglas, R. J. (1996). Comparison of discharge variability in vitro and in vivo in cat visual cortex neurons. *Journal of neurophysiology*, 75(5), 1806–1814.

[22] Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, 79(8), 2554–2558.

[23] Huttenlocher, P. R. (2002). Neural plasticity: The effects of environment on the development of the cerebral cortex.

[24] Jaynes, E. T. (2003). *Probability theory: The logic of science*. Cambridge university press.

[25] Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S., Hudspeth, A. J., & Mack, S. (2000). *Principles of neural science*, volume 4. McGraw-hill New York.

[26] Liao, X., Vasilakos, A. V., & He, Y. (2017). Small-world human brain networks: perspectives and challenges. *Neuroscience & Biobehavioral Reviews*, 77, 286–300.

[27] Mainen, Z. F. & Sejnowski, T. J. (1995). Reliability of spike timing in neocortical neurons. *Science*, 268(5216), 1503–1506.

[28] Moini, J. & Piran, P. (2020). *Functional and Clinical Neuroanatomy*. Elsevier.

[29] Moser, E. I., Kropff, E., & Moser, M.-B. (2008). Place cells, grid cells, and the brain's spatial representation system. *Annu. Rev. Neurosci.*, 31, 69–89.

[30] Newman, M. E. (2003). The structure and function of complex networks. *SIAM review*, 45(2), 167–256.

[31] Nishimori, H. (2001). *Statistical physics of spin glasses and information processing: an introduction*. Number 111. Clarendon Press.

[32] Ostroumova Prokhorenkova, L. & Samosvat, E. (2014). Global clustering coefficient in scale-free networks. In *International Workshop on Algorithms and Models for the Web-Graph* (pp. 47–58).: Springer.

[33] Shinomoto, S., Miura, K., & Koyama, S. (2005). A measure of local variation of inter-spike intervals. *Biosystems*, 79(1-3), 67–72.

[34] Shinomoto, S., Shima, K., & Tanji, J. (2003). Differences in spiking patterns among cortical neurons. *Neural computation*, 15(12), 2823–2842.

[35] Skaggs, W., Mcnaughton, B., & Gothard, K. (1992). An information-theoretic approach to deciphering the hippocampal code. *Advances in neural information processing systems*, 5.

[36] Stensola, H., Stensola, T., Solstad, T., Frøland, K., Moser, M.-B., & Moser, E. I. (2012). The entorhinal grid map is discretized. *Nature*, 492(7427), 72–78.

[37] Sydney, A., Scoglio, C., Youssef, M., & Schumm, P. (2010). Characterising the robustness of complex networks. *International Journal of Internet Technology and Secured Transactions*, 2(3-4), 291–320.

[38] Tezuka, T. (2018). Multineuron spike train analysis with r-convolution linear combination kernel. *Neural Networks*, 102, 67–77.

[39] Tucker, A. (1994). *Applied combinatorics*. John Wiley & Sons, Inc.

[40] van Putten, M. J. (2020). *Dynamics of Neural Networks*. Springer.

[41] Van Vreeswijk, C. & Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*, 274(5293), 1724–1726.

[42] van Vreeswijk, C. & Sompolinsky, H. (1998). Chaotic balanced state in a model of cortical circuits. *Neural computation*, 10(6), 1321–1371.

[43] Varshney, L. R., Chen, B. L., Paniagua, E., Hall, D. H., & Chklovskii, D. B. (2011). Structural properties of the caenorhabditis elegans neuronal network. *PLoS computational biology*, 7(2), e1001066.

[44] Watts, D. J. & Strogatz, S. H. (1998). Collective dynamics of 'small-world'networks. *nature*, 393(6684), 440–442.

[45] Yu, S., Huang, D., Singer, W., & Nikolić, D. (2008). A small world of neuronal synchrony. *Cerebral cortex*, 18(12), 2891–2901.

[46] Zhang, P., Wang, T., & Yan, J. (2021). Pagerank centrality and algorithms for weighted, directed networks with applications to world input-output tables. *arXiv preprint arXiv:2104.02764*.

Acknowledgments

Once in the dark, a person must walk, so that along the path a light might sparkle and provide a new perspective on its surroundings. If the person keeps walking, it might find another lonely wanderer, with her/his own light, and together they might brighten a wider section of the landscape. Eventually, if enough people embrace the uncertainty in the interest of discovery, the environment will still remain highly clouted in darkness, but the lightened portions will be cohesive enough to shed an understanding on its nature, and perhaps motivate others to join in the search. I would like, first and foremost, to thank all the wonderful people that have shared part of their research experience with me during my stay in Trondheim, at the Kavli Institute for System Neuroscience, and in particular my co-supervisor Prof. Yasser R., who has persistently fostered my academic curiosity to as wide range of subjects as possible, without loosing track of the actual research goal. Accordingly, I would also like to thank my supervisor here in Padova, Prof. Suweis S.S., and his Ph.D. student Giorgio, who have always shown availability for constructive discussions on how the project was evolving, and have provided valuable insights into how to expand it.

I would like to offer my deepest thanks to my parents, who have provided both emotional and financial support during my stay abroad, and who have visited me, bringing a little piece of Italy in the remote lands of Norway. In the same spirit, I would like to thank my extended family and family friends, who have often displayed interest in my whereabouts and my quests, and have provided emotional support during the journey. In addition, I would

like to express my gratitude to my girlfriend, Elena, who throughout the entire period has been virtually, and later physically, at my side, and with whom I have shared most of my achievements and defeats. Finally, I would like to thank all my friends for the support that they have and continue to show, despite (or maybe because) of my idiosyncrasies.

And now who's ready for the next adventure?