UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

A study of cortical network models with realistic connectivity

Marina Vegué Llorente

A study of cortical network models with realistic connectivity

Tesi Doctoral

Marina Vegué Llorente

Director: Dr. Alexander Roxin (CRM)

Ponent: Dr. Antoni Guillamon Grabolosa (UPC)

Doctorat en Matemàtica Aplicada

Universitat Politècnica de Catalunya

Febrer de 2018

M. Vegué has developed her research at the Centre de Recerca Matemàtica (CRM) as a scientific collaborator.

She has received financial support through the "la Caixa" Fellowship Grant for Doctoral Studies, "la Caixa" Banking Foundation, Barcelona, Spain.

Al Simon, a qui vaig conèixer a l'inici d'aquesta tesi
i amb qui espero compartir molts més anys
de converses sense paraules.

I també als milers d'animals anònims que han donat
la seva vida en benefici de la ciència del cervell.
Perquè el seu sofriment és important
i no podem donar-li l'esquena.

Agraïments

Aquesta tesi és el resultat d'una etapa de formació i de reflexió, durant la qual he après bastant, tant sobre el funcionament del cervell com sobre molts altres aspectes de l'existència. És per això que estic en deute amb algunes persones, no només per la seva vinculació directa amb el contingut de la tesi sinó també per haver estat a prop i haver format part de la meva vida durant aquest temps.

En primer lloc he d'agrair la dedicació del meu director de tesi, l'Alex, el seu rigor científic i la seva visió sempre crítica de la neurociència. També el suport del Toni, que potser és a qui més han anat a parar els meus dubtes sobre el sentit de tot plegat i que en diverses ocasions m'ha animat a continuar quan jo ho veia tot negre. Vull donar les gràcies a Nicolas Brunel i a Srdjan Ostojic, perquè em van acollir en els seus respectius grups de recerca en dues estades a l'estranger, a Chicago i a París. Aquestes experiències han eixamplat el meu món, en aspectes tant científics com personals. I també al Yonatan, l'Ulises, la Ho Ling i el Maurizio, investigadors del grup de neurociència computacional de la Universitat de Chicago, per haver-me fet sentir com a casa. La Fundació "la Caixa" i la Fundació Ferran Sunyer i Balaguer han finançat la recerca a través d'una beca doctoral i una beca de mobilitat, i per tant han sigut peces indispensables. Agraeixo també al CRM haver-me acollit durant aquests anys com a part del seu equip investigador.

Algunes de les persones que han estat a prop meu durant aquests anys han deixat una empremta inesborrable. Mai oblidaré la primera època al CRM, l'acollida de l'Esther i el Roberto, i l'amistat que va sorgir d'allà i que va donar lloc a tants sopars i lectures apassionades sobre els misteris de la biologia. Els companys de grup, Yota, Bernat i Genís, heu fet, amb el vostre humor, molt més divertides les reunions i les sortides neurocientífiques. També ha estat important l'ambient càlid i festiu entre els investigadors i els estudiants de doctorat del CRM, Álvaro, Tomás, Isabel, Juan, Alberto, Gemma, Víctor, Carmelo, Helena i Claudia, que heu acompanyat i heu fet més fàcils els dies monòtons en què no surt res. Gràcies també al Lluís Alsedà per tots els consells sobre la vida i la ciència entre passadissos, i per tenir un dels somriures que més s'encomanen. I, sobretot, sempre agrairé infinitament el suport indispensable de la Núria, companya, amiga, i gran persona, a qui tinc un apreci molt especial.

Han sigut molt importants per mi les reunions periòdiques amb la comunitat de neurociència computacional de Barcelona, a través de *journal clubs*, xerrades i múltiples altres encontres científics. D'aquestes trobades he extret una concepció més àmplia de com és la recerca actualment, i he pogut desenvolupar la meva pròpia visió de la ciència. Són moltes les persones que n'han format part, i no podria anomenar-les totes sense deixar-me algú. Entre elles, vull donar les gràcies especialment al Txema, a l'Alberto i a la Maria, amb els quals he gaudit discutint de tot allò que vingués al cas. També a l'Adrià, interlocutor essencial davant els fenòmens humans inexplicables, i peça clau en la meva lluita pel respecte animal. I, sens dubte, a l'Iñigo, per les seves ganes de canviar el món i per haver compartit des d'un principi la passió per l'enigma de la consciència.

No hauria arribat fins aquí sense l'ajuda de les persones que m'han format durant tots els anys passats. No puc deixar de mencionar el paper que alguns professors d'escola i d'institut van tenir en fomentar el meu interès per la ciència. A ells vull dedicar un petit però sincer homenatge. Tinc un record molt especial del Llorenç Guasch, la Maribel Mestre, la Sagrario Vidales, la Neus Margalef i el Jesús Vega; amb vosaltres vaig descobrir que les matemàtiques són l'exercici intel·lectual més divertit. I guardaré sempre a la memòria les classes de l'Elina de Miguel, el seu entusiasme i la seva singular devoció per la biologia.

Evidentment no seria qui sóc sense els meus pares, Marta i Pere. M'heu ensenyat a viure i heu fet néixer en mi el sentit crític i l'interès i l'admiració pel món. I sempre heu cregut en mi. Dono també les gràcies a la meva àvia Ana, lectora incansable i amb una curiositat que mai ha deixat de sorprendre'm. I, per descomptat, a l'estimadíssima Elisa, germana, confident i amiga.

Finalment, al Carles, simplement per ser-hi, pel camí que hem recorregut junts i per tot el que vindrà.

Resum

L'estructura té un paper fonamental a l'hora de determinar els tipus d'operacions que els circuits neuronals poden dur a terme. Entendre les lleis que defineixen la connectivitat de les xarxes del cervell i les seves implicacions en la dinàmica neuronal és, per tant, un pas important en la comprensió del funcionament d'aquestes xarxes. Els circuits locals del còrtex, que es creu que suporten les computacions essencials i bàsiques de la funció cerebral, estan organitzats de manera altament ordenada i estereotipada, i aquesta arquitectura, en termes molt generals, s'ha conservat al llarg de les diferents espècies, de les diverses àrees cerebrals i dels individus. Una bona manera de representar matemàticament aquesta família de xarxes és mitjançant models definits per una sèrie de lleis de connectivitat que inclouen un cert grau d'aleatorietat. Les lleis reflecteixen el patró estructural comú, mentre que l'aleatorietat ha de ser interpretada com la variabilitat quan es comparen diferents xarxes del conjunt.

Cada vegada hi ha més evidència experimental que els circuits locals del còrtex estan lluny del model aleatori més simple, segons el qual les connexions apareixen de manera independent amb una probabilitat fixada. Aquesta troballa es fonamenta en un conjunt d'observacions a les quals ens referim col·lectivament com la "no aleatorietat" dels circuits corticals. En aquesta tesi hem explorat fins a quin punt diverses arquitectures alternatives (xarxes amb agrupació, xarxes amb connectivitat dependent de la distància i xarxes definides a través d'una certa distribució de graus d'entrada i de sortida) podrien ser compatibles amb les propietats de no aleatorietat. Hem mostrat que tots els models estructurals alternatius que havíem proposat poden explicar les observacions esmentades, per tant aquestes propietats no aporten gaire informació sobre el tipus d'organització subjacent. Això es deu principalment al fet que les dades reals provenen d'anàlisis molt restringides, en les quals l'estructura s'estudia a partir de mostres locals formades per poques neurones. Vam buscar un estadístic local que permetés, malgrat aquestes dificultats, distingir entre les diverses estructures alternatives, i l'hem trobat en el coeficient de correlació entre els graus d'entrada i de sortida en mostres petites, que hem anomenat *sample degree correlation* (SDC) en anglès. L'anàlisi d'aquesta mesura en dades reals suggereix que els microcircuits corticals tenen una configuració heterogènia −en el sentit que semblen diferir dels models simples proposats− i estan modelats possiblement per factors dependents de la distància física entre neurones però també per principis addicionals que actuen de manera no simètrica.

Finalment, hem estudiat algunes de les conseqüències dinàmiques d'imposar una estructura heterogènia en models d'activitat neuronal. Aquesta heterogeneïtat apareix en els nostres models a través de la distribució conjunta de graus d'entrada i de sortida a la xarxa completa. Fent ús d'aproximacions de camp mitjà i de l'anàlisi espectral, hem mostrat que les distribucions de grau amb elevada variància i correlació positiva poden tenir un efecte rellevant en la dinàmica neuronal, fet que suggereix que aquest tipus d'heterogeneïtat estructural podria facilitar uns modes de computació més rics en comparació dels models aleatoris estàndard.

Resumen

La estructura juega un papel fundamental en el tipo de operaciones que los circuitos neuronales pueden llevar a cabo. Entender las leyes que definen la conectividad de las redes del cerebro y sus implicaciones en la dinámica neuronal es, por tanto, un paso importante en la comprensión del funcionamiento de estas redes. Los circuitos locales del córtex, que se cree realizan las computaciones esenciales y básicas de la función cerebral, están organizados de manera altamente ordenada y estereotipada, y esta arquitectura, en términos muy generales, se ha conservado a lo largo de las diferentes especies, de las distintas áreas cerebrales y de los individuos. Una manera de representar matemáticamente esta familia de redes es mediante modelos definidos por una serie de leyes de conectividad que incluyen un cierto grado de aleatoriedad. Las leyes reflejan el patrón estructural común, mientras que la aleatoriedad debe ser interpretada como la variabilidad cuando se comparan distintas redes del conjunto.

Cada vez hay mayor evidencia experimental a favor de que los circuitos locales de la corteza difieren del modelo aleatorio más simple, según el cual las conexiones aparecen de manera independiente con una probabilidad fija. Este hallazgo se basa en un conjunto de observaciones a les cuales nos referimos colectivamente como la "no aleatoriedad" de los circuitos corticales. En esta tesis hemos explorado hasta qué punto distintas arquitecturas alternativas (redes con agrupación, redes con conectividad dependiente de la distancia y redes definidas a través de una cierta distribución de grados de entrada y de salida) podrían ser compatibles con las propiedades de no aleatoriedad. Mostramos que todos los modelos estructurales alternativos que habíamos propuesto pueden explicar las observaciones mencionadas, por tanto estas propiedades no aportan demasiada información sobre el tipo de organización subyacente. Esto se debe principalmente al hecho de que los datos reales provienen de análisis muy restringidos, en los cuales la estructura se estudia a partir de muestras locales formadas por pocas neuronas. Nos propusimos encontrar una medida estadística local que permitiera, a pesar de estas dificultades, distinguir entre las distintas estructuras alternativas, y la hemos encontrado en el coeficiente de correlación entre los grados de entrada y de salida en muestras pequeñas, que hemos llamado *sample degree correlation* (SDC) en inglés. El análisis de esta medida en datos reales sugiere que los microcircuitos corticales tienen una configuración heterogénea −en el sentido de que se alejan de los modelos simples propuestos− y están modelados posiblemente por factores dependientes de la distancia física entre neuronas pero también por principios adicionales que actúan de manera no simétrica.

Finalmente, hemos estudiado algunas de las consecuencias dinámicas de imponer una estructura heterogénea en modelos de actividad neuronal. Esta heterogeneidad aparece en nuestros modelos a través de la distribución conjunta de grados de entrada y de salida en la red completa. Usando aproximaciones de campo medio y análisis espectral, hemos mostrado que las distribuciones de grado con elevada variancia y correlación positiva pueden tener un efecto relevante en la dinámica neuronal, hecho que sugiere que este tipo de heterogeneidad estructural podría facilitar unos modos de computación más ricos en comparación con los modelos aleatorios estándar.

Abstract

Structure is fundamental in shaping the types of computations that neuronal circuits can perform. Explaining the laws that determine the connectivity properties of brain networks and their implications in neuronal dynamics is therefore an important step in the understanding of how brains operate. The local circuits of cortex, which are considered to carry out the basic and essential computations for brain functioning, exhibit a highly stereotyped and organized architecture, which is, in very general terms, conserved across different species, brain areas and individuals. An appropriate way to mathematically represent this family of networks is by means of models defined by a set of connectivity laws that include a certain degree of randomness. These laws reflect the common structural scaffold, whereas the randomness should be interpreted as the variability across the different networks in the ensemble.

There is growing experimental evidence that the local circuits of cerebral cortex are far from the simplest random model, according to which connections appear independently with a fixed probability. This evidence is based on a set of observed features that have been collectively called the "nonrandomness" of the cortical circuitry. In this thesis we have explored to what extent several alternative architectures (clustered networks, networks with distance-dependent connectivity and networks that exhibit a given in/out-degree distribution) could be compatible with the reported nonrandom features. We showed that all these structural models can explain the experimental observations, which implies that these nonrandom properties do not provide much information about the underlying organization. This is mainly due to the fact that real data are collected from sparse neuronal samples due to experimental limitations. We sought a local measure that can nevertheless help to distinguish between different alternatives, and we found it in the *sample degree correlation* (SDC), or the correlation coefficient between in- and out-degrees in small groups of neurons. The analysis of the SDC in real data suggests that cortical microcircuits are heterogeneous in structure and possibly shaped through a mixture of distance-dependent and non-symmetrical organizational principles.

We finally explored some of the dynamical consequences of imposing a heterogeneous structure in models of neuronal activity. This heterogeneity appears through an arbitrary joint in/out-degree distribution in the entire network. By means of both mean-field approximations and spectral analysis, we demonstrate that broad and positively correlated degree distributions can have an important effect on neuronal dynamics, which suggests that this particular type of structural heterogeneity might allow for richer network computations as compared to standard random models.

Keywords: neural networks, cortical connectivity, microcircuits, structural motifs, degree distribution, degree correlation, dynamics, mean-field, stationary state, spectrum

Contents

.

"No es disparatado decir que la misma consciencia que nos permite investigar el funcionamiento de nuestras células puede provenir de las capacidades coordinadas de millones de microorganismos que evolucionaron simbióticamente hasta convertirse en el cerebro humano."

Lynn Margulis y Dorion Sagan, *Microcosmos* [1]

[1] L. Margulis y D. Sagan. *Microcosmos. Cuatro mil millones de años de evolución desde nuestros ancestros microbianos.* Tusquets Editores, Barcelona, 1995.

Introduction

Background

The study of the nervous system underwent a profound transformation after the work of Camillo Golgi and Santiago Ramón y Cajal during the late nineteenth century. Golgi developed a method of silver impregnation which, when applied to brain preparations, resulted in the staining of only a small, random proportion of cells. This technique makes it possible to visualize entire neurons, with what we now recognize as dendrites and axons, in high contrast under the microscope. The Golgi technique was then used by Cajal, whose thorough examinations of nervous tissue preparations drove him to the formulation of the *neuron doctrine*, which states that the nervous systems are composed of a juxtaposition of many individual and independent units: the neurons. This "atomized" view of the nervous structures, which extended the already formulated cell theory, establishes that all the complexity and capabilities of animal brains arise from the combination and interaction of many of these functional and structural "building blocks", which communicate with each other through tiny empty clefts or synapses. The neuron doctrine contrasted with the so-called *reticular theory*, according to which the nervous system is a single continuous net of tissue. Ironically, Golgi was a passionate defender of the reticular theory, and he supported it even during the lecture he pronounced at the Nobel ceremony, after Golgi himself and Cajal were jointly awarded the Nobel Prize in Physiology or Medicine for their discoveries.

The neuronal theory established the foundations of modern neuroscience. In the following decades, many advances led to a deeper understanding of the properties of single neurons, their electrical behavior and the molecular principles behind synaptic transmission. These achievements were paralleled by the progresses of genetics and molecular biology, together with the development of a rich repertoire of new image and molecular techniques. One of the key achievements in the course of this transformation was provided by the work of Eric Kandel and collaborators on

the molecular mechanisms underlying memory storage. In the 1950s, learning was considered by many a higher-order cognitive process that had to be addressed from the perspectives of psychology or psychiatry. Nonetheless, the studies of Kandel and others on the giant sea slug *Aplysia* −which, despite having a relative simple nervous system, exhibits basic forms of learning− proved not only that memory can be successfully studied on these organisms, but also that it has a molecular substrate which eventually shapes the strength of individual synapses. Such discoveries showed that both short- and long-term memory can be related to specific molecular events which link neuronal activity with certain modifications of membrane channels, receptors and even the physical structure of synaptic buttons [Kandel, 2001]. More recent studies have also revealed that memory can involve the formation of new connections or the pruning of existing ones [Zuo et al., 2005; Hofer et al., 2009]. Differences between these two forms of learning were also related to distinct mechanisms: whereas the molecular processes that lead to short-term memory involve the chemical modification of preexisting proteins, long-term learning requires new protein synthesis. Moreover, further research proved that the very same molecular agents identified in *Aplysia* participate in the formation and maintenance of memories in more complex organisms such as rodents [Kandel, 2001].

These findings are of particular significance. On the one hand, they related a *behavioral* property (memory) with a *physiological* process (structural modification of synapses). The idea that learning relies on the plastic modification of synaptic strengths suggests that it is *structure* (and not neurons per se) what defines the type of computations that a brain can perform. On the other hand, the mentioned discoveries indicate that one of the highest order processes performed by large mammalian brains −complex learning− shares its basic mechanisms with one of the simplest organisms endowed with a nervous system. Such a "universality" of memory −which extended the already recognized universality of other biological principles like the genetic code or the physiology of the cell− brought an important issue to the fore: the high capabilities of mammalian and, in particular, human brains are not due to unique mechanisms but are presumably the result of both *numbers* and *structure*. The success lies in having a huge amount of units connected in the appropriate manner.

The importance of structure

These experiments on the physiological basis of memory were key to defining one of the most important paradigms of neural science: that the repertoire of computations that brains can perform

is defined by the precise connectivity patterns among their neurons. Anatomical studies on the organization of neurons in the mammalian brain had provided additional evidence for this. The work of Cajal on the anatomy of nervous tissue had already revealed that the mammalian cortex is a highly ordered structure, with different kinds of neurons located in specific positions along the vertical axis and a clear horizontally stable structural pattern (Fig. 1). Neuroanatomical studies during the early twentieth century provided a detailed description of the so-called *laminar structure of the cortex*: the cortex is organized into six different horizontal layers defined by the presence of distinct types of neurons and fibers. In the sixties, from his experiments on the cat's neocortex, Vernon Mountcastle proposed a principle of *functional organization of the neocortex*, according to which neurons that lie in close proximity across vertical cylinders that span the different cortical layers form functional units. For example, in the sensory cortex, neurons that compose a unit have almost identical responses to the stimulation of peripheral receptors and share receptive fields [Mountcastle, 1957]. He called these units *cortical minicolumns*. Subsequent experiments such as the ones performed by David Hubel and Torsten Wiesel on the visual cortex of the cat further confirmed these ideas [Hubel and Wiesel, 1962]. Such findings suggested that the cortex is composed of an iterative juxtaposition of many stereotypical functional and structural modules [Mountcastle, 1997]. Different modules might differ in their fine structure according to differences in functionality, but the general architectural pattern is highly stable [Harris and Shepherd, 2015]. Stability occurs not only across distinct brain regions but also across different species. The modular organization of the cortex is a widely conserved evolutionary invention, and the differences between the brains of different species rely mainly on the total surface area rather than on its local structure.

The discovery of a conserved architecture for the cerebral cortex has interesting implications in our current understanding of how the mammalian brain works. The fact that areas which encode very different types of information (i.e., visual versus auditory stimuli) exhibit a common structural pattern seems to indicate that the kinds of computations needed to store and process such information are similar and independent of the qualitative nature of its content. The universality of structure might imply a universality of the basic modes of computation.

From that perspective, the functional differences between brain areas might be the result of a distinct connectivity with other cortical or extracortical regions and with peripheral receptors. It is also plausible that these differences are determined by the fine structure within the above-mentioned stereotypical modules. The experiments of Kandel and colleagues on *Aplysia* showed that whereas the general connectome of the sea slug is hard-wired (that is, it is highly conserved

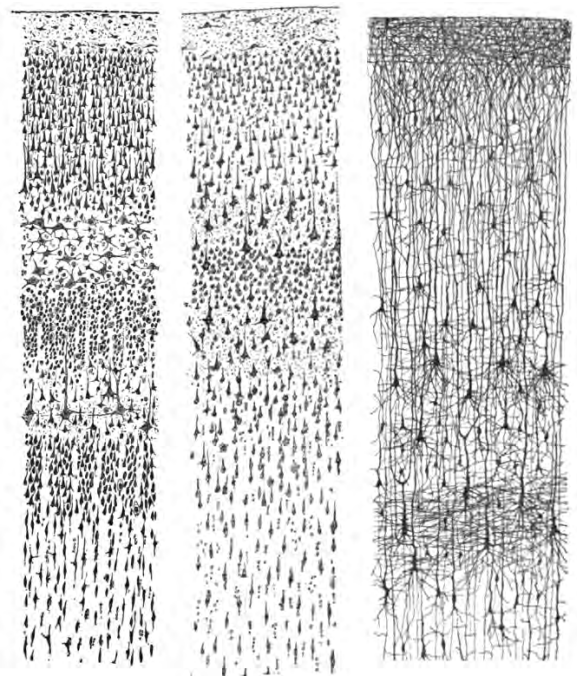

Figure 1: Three drawings of the human cortex by Santiago Ramón y Cajal. All of them correspond to vertical cross-sections. **Left.** Visual cortex of an adult (Nissl stain). **Middle.** Motor cortex of an adult (Nissl stain). **Right.** Cortex of a 1.5 month old child (Golgi stain). Image published in [Cajal, 2013].

among individuals, probably as a result of a precise genetic program), this connectivity profile can be fine-tuned so as to adapt behavior according to the animal's past experiences. The repertoire of possible behaviors of the animal is limited and predefined by its "by default" circuitry, but the system itself can suffer minor modifications that will determine the specific properties and relations of such stereotyped behaviors. This concept goes beyond *Aplysia* and could be applied to the cortex as well: whereas the general connectivity plan defines what kind of information can be stored, processed and transmitted, the fine structure might be responsible for codifying the specific content of such information.

Structure and function

It is therefore widely accepted that the functionality of neural circuits depends on their connectivity. We have also pointed out the hypothesis that cortical networks are organized according to a general −and probably genetically predefined− "scaffold" that is conserved along the lifetime of the individual and is shared by distinct brain areas and even across species. We postulate that such a common structural plan shapes and constrains function in a wide sense whereas the fine structure defines the precise content of the encoded information. Memory can be regarded as the process by which experience imprints its unique trace on top this general, stable organization.

The above observations suggest that a deep understanding of how brains process information will

necessarily require a thorough analysis of how neuronal networks are arranged. Structure and function are just two manifestations of a common issue. One possible way to shed some light on how brain networks work might be to focus on the local networks mentioned before. Understanding the functioning of those microcircuits can be a key step towards the comprehension of how they interact with other circuits to produce complex computations.

Consequently, a study on the structure or topology of the local circuitry is needed. A possible strategy to elucidate the basic operational mechanisms of these networks is therefore to understand what the general rules that define their structure are. The main objective of this thesis is precisely to explore the question of how cortical local networks are built in a broad, statistical sense, and also to explore the relationship between certain statistical features and network dynamics. A statistical approach is important for highlighting those basic features that are independent of the particular organism, brain area or individual studied. The standard way of studying network architecture in a statistical sense is by introducing random connectivity models, that is, models defined by a set of laws that also include a certain amount of randomness. The laws constrain the statistical properties of the set of possible networks that result from the model, whereas the randomness has to be interpreted as the variability across the networks in the ensemble. Whether the true origin of such variability is random or not is a fundamental question that we do not address here.

Insights from experiments and network modeling

Mathematical modeling is a powerful tool to explore and analyze complex systems, and it provides a means for a precise and quantitative study of nature. Ideal models are simple enough to be tractable but complex enough to provide information that could not have obtained without the use of a mathematical language. Abstract models have proven to be successful in many areas of brain science. One paradigmatic example is the well-known Hodgkin-Huxley model, named after Alan Lloyd Hodgkin and Andrew Fielding Huxley, who proposed a set of differential equations to describe the ionic mechanisms underlying the formation of action potentials in nerve cells [Hodgkin and Huxley, 1952]. They were awarded the Nobel Prize in Physiology or Medicine for their work in 1952.

Models of neural networks are used to extract information about the collective behavior of the system (the neuronal ensemble) from a set of rules that define how single neurons behave and how the interactions between them are. Classical models are designed to study synchronization mech-

anisms, to explore the different types of available dynamics, or even to address how functional processes as working memory or decision making could be implemented, to cite just some examples. In the majority of such models, the connectivity structure has traditionally been assumed to be either homogeneous or simply random. By *homogeneous* we mean that all the neurons receive a fixed number of connections from the network, although the precise connectivity configuration might be chosen randomly. On the other hand, *simply random* (or *random*, without additional specifications) tends to be used to denote a scenario in which each connection appears independently with a fixed probability p, which is a parameter of the model. The last is also known as the Erdös-Rényi model, the name given after the mathematicians Paul Erdös and Alfréd Rényi, who described and studied a closely related model [Erdős and Rényi, 1959].

Connection density has been estimated to be small in cortical slices (see, for example, [Holmgren et al., 2003; Song et al., 2005; Le Bé and Markram, 2006]), leading to the idea that brain networks are *sparse* although the total number of synaptic contacts that a given neuron makes tends to be large. Sparseness is represented in simple random models by assuming that p is a small, constant parameter or even that it scales inversely with the number of neurons N: $p \propto 1/N$. In any of these scenarios, the neuron-to-neuron variability in the number of connections received relative to the mean is small in the large N limit, which indicates that when dealing with large networks, random models can be considered to be almost homogeneous. This is the reason why in certain contexts it is common to use the term *homogeneous* for both random and homogeneous topologies. Such terminology is also employed in contraposition to the word *heterogeneous*, which refers to networks with a non-negligible degree of inter-neuronal structural variability.

Network models with homogeneous connectivity structure have been widely used to study neuronal dynamics. They have been successful in reproducing general features of the spontaneous activity of neurons reported by electrophysiological studies. The high irregularity of the spiking process in real neurons, for example, can be captured by randomly connected networks of excitatory and inhibitory neurons where the amount of excitation and inhibition received by single cells cancel in the mean and where fluctuations are responsible for the spiking activity [van Vreeswijk and Sompolinsky, 1996; Renart et al., 2010].

Data on the precise connectivity between individual neurons in cortex, however, have shown that homogeneous models for network structure are overly simplistic. There is consistent evidence that the connectivity patterns between pyramidal neurons in the rodent cortex significantly deviate from those schemes. This evidence comes from the study of statistical connectivity features in

small groups of neurons in brain slices [Song et al., 2005; Perin et al., 2011]. One of such findings is that the amount of reciprocal connections is much larger than what would be expected in simple homogeneous or random models [Song et al., 2005]. Importantly, these "nonrandom" features have been found in different animals and brain areas, so they might be footprints of the above-mentioned "stereotypical connectivity laws" which constrain cortical structure in a wide sense. The implications that more realistic topologies would have in neuronal dynamics are still largely unknown.

In the last years, some studies have addressed the relation between structure and function in neuronal network models. The finding of the above-mentioned nonrandom features −in particular, the over-representation of reciprocally connected neuronal pairs [Song et al., 2005] and the so-called "common neighbor rule", according to which neurons that share more neighbors tend to be connected more frequently [Perin et al., 2011]− has been interpreted as evidence for clustering in those networks. Clustering implies that neurons belong to communities and they are preferentially connected with neurons in the same community. The effect of clustering on neuronal dynamics has been studied theoretically [Deco and Hugues, 2012; Litwin-Kumar and Doiron, 2014; Mazzucato et al., 2015]. In these papers, the authors simulated networks of excitatory and inhibitory neurons in which the excitatory-to-excitatory connectivity was clustered. They found that an asynchronous, balanced state can be also reproduced by this alternative architecture. Moreover, the spontaneous state of the network consists of a random and transient activation of the different clusters. The reason is that the connectivity structure allows the dynamics to have multiple attractors, which are characterized by the selective activation of a subset of clusters, and noise makes the network state stochastically "jump" from one to another. This produces slow temporal fluctuations in the firing rates that are also highly variable from trial to trial, in accordance with real spontaneous dynamics [Churchland et al., 2010]. The introduction of a transient stimulation that targets a fraction of neurons in a given cluster favors the activation of the whole cluster and it is therefore accompanied by a drastic reduction of inter-trial variability, which is again consistent with experiments [Churchland et al., 2010]. Therefore, clustering is a structural feature that can explain the reduction in trial-to-trial variability induced by a stimulus that has been observed experimentally.

The role of structural aggregation has also been analyzed in a recent work that related both clustering and the presence of weight-hubs (i.e., neurons whose incoming connections exhibit large synaptic weights) with two functional properties reported in experiments: the presence of irregular transitions between up- and down-states and the spreading of activity induced by

optogenetic stimulation in certain cortical layers [Setareh et al., 2017].

Firing measurements during spontaneous activity have revealed long-tailed rate distributions [Hromádka et al., 2008; O'Connor et al., 2010]. Such distributions could emerge from single neuron properties or from certain connectivity structures between neurons, or both. If a neuron receives inputs from pre-synaptic neurons whose activity is almost uncorrelated, the total synaptic input received is normally distributed (as a consequence of the Central Limit Theorem). If different neurons have homogeneous connectivity properties, the synaptic inputs are also normally distributed *across neurons*. If, in addition, the relationship between the synaptic input and the output firing rate is linear, then the instantaneous firing rates of different neurons should be normally distributed as well. Koulakov et al. [2009] examined connectivity-based strategies to reproduce heavy-tailed rate distributions in network models where the input-output relationship is linear. They found that such distributions are exhibited by networks in which individual synaptic weights follow a lognormal distribution (consistent with experiments like the ones reported in [Song et al., 2005]) and where the incoming weights to a given neuron are positively correlated (which means that some neurons tend to receive stronger connections than others, so that the homogeneity hypothesis needed in the normality argument does not hold). They concluded that the presence of correlated incoming synaptic weights can reproduce the experimental findings. But heavy-tailed rate distributions can also be explained by a non-linear relationship between input and output firing rates in models with homogeneous connectivity [Roxin et al., 2011].

The nonrandomness found in the cortical circuitry also involves the over or underrepresentation of several three-neuron motifs (or subgraphs) relative to the classical random model [Song et al., 2005]. Other authors have examined how the presence of different two-synapse motifs modulates dynamics in models of coupled oscillators [Zhao et al., 2011]. In this work, the authors constructed networks with a prescribed proportion of bidirectional, convergent, divergent and chain motifs, and then assessed to what extent the ability of the network to synchronize is affected by these proportions. They showed that synchronizability is facilitated by the presence of chain motifs whereas it is suppressed when the network is rich in convergent motifs. A mathematical framework for assessing the effect of motifs on global activity correlations or *coherence* has been described in [Pernice et al., 2011; Hu et al., 2014].

A possible interpretation of the experimental findings regarding the structure of cortical circuits might be that the mentioned nonrandom features arise as a consequence of neurons having in- and out-degrees (i.e., number of incoming and outgoing connections) that are distributed differently

from what would be expected in homogeneous or simply random models. The purely homogeneous network exhibits Delta in-degree and Binomial out-degree distributions, whereas in the Erdös-Rényi model both degrees follow Binomial distributions. In both scenarios, the in-degree and the out-degree of individual neurons are independent variables. Although degree distribution is not a local property and is therefore very difficult to estimate in a real network without the knowledge of the complete structure, many alternative −and more heterogeneous− degree distributions are in principle possible to occur in cortical circuits. Individual pyramidal neurons in cortex exhibit distinct amounts of coupling to the overall activity in the surrounding network [Okun et al., 2015]. This heterogeneity is, moreover, related to their likelihood to receive connections from neighboring pyramidal cells: neurons with high coupling tend to receive more connections from the local network than those with a low level of coupling. Thus, diversity in population coupling might be a functional consequence of structural degree heterogeneity in the circuits of cerebral cortex.

The role that marginal in- and out-degree distributions play in dynamics of integrate-and-fire and rate models has been analyzed by Roxin [2011]. It was shown that the variance of in-degrees has an important impact in the global dynamical state of the network, specifically in the ability of the system to undergo global oscillations, whereas the out-degree variance shapes pairwise correlations in the synaptic currents. Apart from marginal degree distributions, degree correlations might play an important role in neuronal activity. In network theory, the term *assortativity* is used to denote the property by which connected nodes tend to have similar degrees [Newman, 2003]. Schmeltzer et al. [2015] showed that assortative in-degree correlations improve the sensitivity to weak stimuli in model neural networks. Degree correlations between in-degrees and out-degrees of individual neurons have also been shown to substantially affect neuronal dynamics [Vasquez et al., 2013; Nykamp et al., 2017].

The previously mentioned work directly links specific structural features with particular dynamical properties. An interesting and complementary approximation to the problem of structure and function is to seek relations between the configuration of brain networks and their performance in certain tasks. In a recent paper, Brunel [2016] found that model networks that have been selected for being optimal in terms of memory storage exhibit some of the reported nonrandom features, a result that could suggest a role of these features in learning. Another recently published work analyzed the functional consequences of the micro-anatomy of cerebellum [Litwin-Kumar et al., 2017]. Neurons in cerebellum possess unique anatomical and organizational properties: mossy fibers originated mainly in brainstem nuclei provide inputs to granule cells (the most abundant

excitatory cells in the brain), which in turn project to Purkinje cells. The granule cell layer thus seems to act as a "relay station" for the information that travels from mossy fibers to Purkinje cells. The cytoarchitecture is such that granule cells outnumber mossy fibers by a factor of about 200 and each granule cell receives input from 4-5 mossy fibers, whereas the number of granule cells that provide input to a single Purkinje cell is large [Eccles et al., 1967]. Litwin-Kumar et al. [2017] used analytical arguments to show that the intermediate granule cell layer could amplify the signals conveyed by mossy fibers to favor high-dimensional representations even at the sparse connectivity density between the fibers and granule cells reported in experiments. Such an enhanced dimensionality might facilitate posterior computations by Purkinje cells, as has been shown to occur in other brain regions [Rigotti et al., 2013].

The previous studies exemplify a reductionist approach, based on the study of very well defined neuronal systems where specific structural principles are controlled and some of the resulting dynamical consequences can be precisely quantified. One of the problems of assessing the effect of particular connectivity features on neuronal activity is that the considered models usually represent a very small fraction of all the possible topologies. It is therefore difficult to clearly define a role of specific structural features on dynamics in general. As a way to overcome this obstacle, Pernice et al. [2013] randomly constructed a vast collection of different networks, not subject to any *a priori* architectural prototype, and performed a statistical analysis on the entire network ensemble to link particular structural properties with specific dynamical features. They found, for example, that in- and out-degrees of neurons have a clear impact on dynamics, in agreement with [Roxin, 2011]. Isolating the contribution of other individual features like connectivity motifs, however, is not completely possible because their numbers are not independent quantities. Using a similar strategy, Tomm et al. [2014] generated thousands of network structures and asked which ones are compatible with data about both structure and neuronal dynamics (from layer 2/3 of barrel and visual rodent cortices). They showed that the connectivity constraints require degree distributions to be heterogeneous within the excitatory subnetwork and that activity constrains impose positive correlations between the excitatory input weights onto single neurons and between the output synaptic weights from single excitatory neurons to inhibitory fast spiking neurons. These approaches are certainly interesting but have the drawback that an analytical treatment of the models and their functional impact tends to be problematic, and this severely limits the comprehension of such structural-functional relations.

A very different approximation to the problem of how structure and function interact is provided by the Blue Brain Project. This project seeks to reproduce, with the highest possible level of

realism, the functioning of an entire cortical microcircuit. To do so, a vast collection of anatomical and electrophysiological data has been used to develop a very detailed reconstruction of a certain volume of the rat somatosensory cortex [Markram et al., 2015]. The dynamical simulation of this ambitious *in silico* experiment shows activity patterns that resemble those found both *in vivo* and *in vitro*, including transitions from asynchronous to synchronous states.

Although this approach has many potentialities, such as creating a realistic *in silico* environment where different experiments can be performed, thus possibly reducing the need of animal experimentation, its elevated complexity makes it difficult to clearly link specific structural properties with function. We believe that a systematic examination of simplified and well controlled network models deserves further attention and can provide complementary insights into the extent to which neuronal architecture shapes brain functioning.

Objectives

In light of the previous observations, several questions remain unresolved. First, what are the nonrandom features of connectivity telling us about the structure of the cortical microcircuitry? Can we define alternative models that explain such properties? To what extent these findings constrain the possible underlying architectures? We address these issues in Chapter 1, where we depart from the assumption that cortical microcircuitry arises from a predefined set of connectivity rules that nevertheless admit a certain degree of stochasticity. This "noise" represents the variability that we would find when comparing circuits from different animals and brain areas or even as the natural perturbations that experience imprints in connectivity. Our goal is to be able to extract information about such structural principles from the partial and incomplete information provided by the available experimental findings. We do so by means of a statistical and probabilistic approach to the study of random graph models.

A second topic to be addressed concerns functionality. Why should cortical circuits exhibit such structural features? What are the consequences in terms of dynamics of assuming connectivity profiles which differ from the classical homogeneous and simply random scenarios? Do alternative topologies have computational advantage of any kind with respect to classical structures? Although these questions go far beyond the scope of this thesis, we have tried to tackle some less ambitious but related problems.

In Chapter 2 we study networks of spiking neurons where connectivity is highly heterogeneous,

that is, where there is a large variability in terms of the number of connections received and originated from different neurons. We also analyze topologies where the number of incoming and outgoing connections of individual cells are correlated. This type of connectivity goes beyond homogeneous and random models because in those networks the variances in the number of incoming and outgoing connections are small compared to their means. We extend the well-known mean-field formulation [Brunel, 2000; Roxin et al., 2011] to such more general scenarios and we study how alternative topologies can modify the repertoire of firing rates in the stationary state.

One of the approaches for predicting the qualitative behavior of a system of ordinary differential equations (ODE) is to linearize the system around a fixed point and study the eigenvalues of the resulting operator. Firing rate models are defined through systems of ODEs which specify the time-evolution of the firing rates of the neurons in a network. These models are perhaps less realistic than spiking models but can possibly capture general properties of neuronal networks that are independent of the precise details of the system. Therefore, computing the spectral properties of the linearized versions of such models can help to study properties of their dynamical states and how these states depend on the connectivity imposed. In Chapter 3 we consider the problem of predicting the set of eigenvalues of heterogeneous random connectivity matrices. Such matrices correspond to the linearized operators found in certain types of firing rate models, so the possibility to predict their spectra opens new doors to the study of neuronal networks with alternative structure.

Chapter 1

On the structure of cortical microcircuits inferred from small sample sizes

The work in this chapter has been published [Vegué et al., 2017].

1.1 Introduction

Network architecture shapes the way in which information is transmitted and stored in neuronal circuits. In the mammalian cortex, complex functions such as sensory processing, decision making, memory storage and even abstract reasoning are thought to be the result of a highly structured network topology. Therefore, understanding the structure of cortical microcircuits may be a key step towards a deep understanding of how the brain performs such tasks.

The organization of cortical microcircuits varies across brain areas and species, and undergoes continual plastic modifications during the lifetime of a given individual as a result of experience [Trachtenberg et al., 2002; Zuo et al., 2005; Le Bé and Markram, 2006; Hofer et al., 2009]. It is accepted, however, that these circuits also exhibit certain regularities, the canonical example of which is a well defined vertical organization into layers. The existence of conserved connectivity principles suggests the notion of a neocortex composed of a juxtaposition of similarly structured building blocks [Szentágothai, 1978; Mountcastle, 1997; Silberberg et al., 2002], which are then

dynamically adjusted to respond to the precise demands of every subsystem, in a continuously changing environment.

In the last decades, much effort has been devoted to elucidating the structure of cortical microcircuits. Intracellular recording techniques have made it possible to assess the presence of monosynaptic connections between pairs of neurons in cortical slices directly [Mason et al., 1991; Markram et al., 1997; Holmgren et al., 2003; Song et al., 2005; Perin et al., 2011]. The morphological examination of synaptic contacts via electron microscopy can also in principle provide ground-truth connectivity [Denk and Horstmann, 2004; Bock et al., 2011; Kleinfeld et al., 2011; Kasthuri et al., 2015], although current throughput is too small to allow for the reconstruction of microcircuits. Finally, some studies have sought to infer network connectivity from observations of the neuronal dynamics [Nykamp, 2007; Pajevic and Plenz, 2009; Stetter et al., 2012; Sadovsky and MacLean, 2013; Tomm et al., 2014]; the accuracy of such methods generally depends on the how closely the real data might conform to specific model assumptions. Nonetheless, the data acquired through slice electrophysiology still currently represent the most accurate picture of cortical microcircuitry available.

One important limitation of cell recording techniques, however, is that they currently allow for the study of only small groups of neurons simultaneously. Therefore, microcircuitry reconstructions necessarily require an inference process from partial data. Despite these limitations, experimental studies have brought to light some fundamental common principles, such as that the connections tend to be sparse, with connection rates between pyramidal neurons in the range 5-15% [Mason et al., 1991; Markram et al., 1997; Holmgren et al., 2003; Le Bé and Markram, 2006; Wang et al., 2006; Ko et al., 2011]. Recent work has also determined specific connection rates depending on the pre- and post-synaptic cell types [Hill et al., 2012; Jiang et al., 2015]. Interestingly, there is increasing evidence that the connectivity between pyramidal neurons in different areas and layers is far from the Erdös-Rényi (ER) random network model, where connections appear independently with a fixed probability p. These so-called "nonrandom" features include an excess of reciprocal connections, which can be quantified by the ratio between the number of bidirectional connections and the expected number of such connections in ER networks with equivalent connection rates (R). R has been reported to be around 2-4 in visual cortex [Mason et al., 1991; Song et al., 2005], 3-4 in somatosensory cortex [Markram et al., 1997; Le Bé and Markram, 2006] and 4 in medial prefrontal cortex [Wang et al., 2006]. Additional evidence for this nonrandomness is the overrepresentation of highly connected motifs [Song et al., 2005; Perin et al., 2011] and the finding that the connection probability between neuron pairs increases with the number of neighbors they

share [Perin et al., 2011]. Some initiatives are seeking to leverage these data in order to construct realistic microcircuit models for numerical simulation [Hill et al., 2012; Markram et al., 2015; Reimann et al., 2015; Ramaswamy et al., 2015]. On the other hand, a recent theoretical study has shown that some of these features arise naturally in network models that maximize the number of stored memories [Brunel, 2016].

In this chapter we study several broad classes of network structure that could potentially explain the observed nonrandomness. These include clustered networks [Litwin-Kumar and Doiron, 2014], spatially structured networks [Holmgren et al., 2003; Perin et al., 2011; Jiang et al., 2015] and networks with strong heterogeneity in the number of incoming and outgoing connections of neurons [Roxin, 2011; Timme et al., 2016].

Surprisingly, all of these network classes are compatible with the reported nonrandomness. In fact, we have found that networks with qualitatively distinct *global* structure could be nearly indistinguishable when all the available information comes from the study of small groups of neurons, as in experiment. In particular, current measures of cortical circuit structure from slice experiments, such as motifs, common neighbors, or connection density in small groups, cannot be directly used to distinguish between these network classes. We therefore propose here a new measure, the sample degree correlation (SDC), which provides a unique fingerprint for each network class, based only on the analysis of small samples of neurons. Using the SDC we show that microcircuit data from rat somatosensory cortex [Perin et al., 2011] are incompatible with any of these network classes. Rather, the data have lead us to develop a more general network class which reduces to the previous models under certain constraints. Our results suggest that the nonrandom features of cortical microcircuits reflect a combination of spatially-decaying connectivity and additional non-spatial structure which, however, is not simple clustering.

1.2 Canonical network models for cortical circuits

We first asked ourselves to what extent simple, canonical models of network topology could reproduce the salient statistics from actual cortical circuits in slice experiments. The simplest possible sparsely connected network model is the so-called Erdös-Rényi (ER) network, for which connections between neurons are made independently with a fixed probability p. However, data show that cortical circuits are not well described by the ER model, and in particular, the occurrence of certain cortical motifs is above what would be expected from ER. Therefore, we consider other

candidate network models which go beyond ER, as detailed below (see Fig. 1.1).

All the networks are treated as directed graphs with N neurons. We assume that the network's size N is large and that the network is *sparse*, meaning that its connection density p is "small". We use the following notations: $i \to j$: a connection exists from neuron i to neuron j; $i \mapsto j$: a connection exists from i to j but not from j to i; $i \leftrightarrow j$: there is a bidirectional connection between i and j. The network models we consider are as follows:

(i) **ER network with additional bidirectional connections (ER-Bi).** This model has just two parameters: the probability of a unidirectional connection p_{uni} and that of a bidirectional connection p_{bid}. For each neuronal pair (i, j) the connectivity between i and j is created independently of the other pairs and according to these probabilities. In the case of unidirectional connections, the two possible patterns $i \mapsto j$ and $j \mapsto i$ are chosen at random with equal probability.

(ii) **Network with clusters and homogeneous membership (Cl)**. Each neuron belongs to one or more clusters and cluster membership is homogeneous across the network. This means that, for any neuron i, the number of other neurons that share a cluster with i is almost constant. More precisely, if n_i denotes the number of neurons that are at least in one of the clusters of i,

$$\frac{\sqrt{\text{Var}(n_i)}}{\mathbb{E}[n_i]} \to 0 \tag{1.2.1}$$

as $N \to \infty$ (these moments are defined over the network realization but also *across* the network, because we consider models where the connectivity rules are the same for all the neurons). The typical example is a network with a fixed number of clusters $C \ll N$ where each neuron belongs to one cluster that is chosen uniformly at random. In this case, $n_i \sim$ Binomial$(N-1, 1/C)$, so

$$\frac{\sqrt{\text{Var}(n_i)}}{\mathbb{E}[n_i]} \;=\; \sqrt{\frac{C-1}{N-1}} \;\to\; 0. \tag{1.2.2}$$

Connections are generated independently with probability p_+ when neurons are in the same cluster and p_- otherwise, $p_- < p_+$.

(iii) **Network with clusters and heterogeneous membership (Cl-Het).** Neurons belong to a (possibly) variable number of clusters and cluster membership is heterogeneous across neurons, which means that Eq. (1.2.1) does not necessarily hold. The probability of connection within and between clusters is as for the Cl model. In our simulations we have

considered networks with $C \ll N$ clusters where each neuron has a probability $p_c = 1/C$ of belonging to any given cluster. Therefore, neurons can be simultaneously in different clusters and clusters may have non empty overlap. In this case Eq. (1.2.1) does not hold as we show in Materials and Methods at the end of this chapter.

(iv) **Network with a distance-dependent connectivity (Dis).** There is a notion of distance in the network and connections are made independently with a probability that decays with the distance between neurons:

$$P(i \to j \,|\, r_{ij} = r) \quad = \quad p(r), \tag{1.2.3}$$

where r_{ij} is the distance between neurons i and j and $p(r)$ is a decreasing function of r. We assume that distances are homogeneously distributed in the network, i.e., that the proportion of neurons that are a given distance away from a neuron i does not vary substantially with i. This condition is analogous to requirement (1.2.1) for clustered networks. When it does not hold, the model belongs to the Cl-Het class in terms of the properties studied here.

(v) **Network defined by a joint distribution of in- and out-degrees (Deg).** The in-degree and the out-degree of a neuron are the total number of incoming and outgoing connections it has. This model includes networks whose in- and out-degrees follow a prescribed joint distribution, which could be correlated. Important parameters of this model are the mean degree $\langle K \rangle$ (which has to be the same for both degrees because the sum of in-degrees and the sum of out-degrees are equal in any directed network), the degree variances σ^2_{in}, σ^2_{out} and degree correlation coefficient ρ. The connection probability once the network degrees are known can be approximated by

$$P(i \to j \,|\, K_j^{\text{in}} = k, K_i^{\text{out}} = k') \quad \simeq \quad \frac{kk'}{N\langle K \rangle} \tag{1.2.4}$$

(see Materials and Methods for details).

Figure 1.1 shows a schematic representation of example networks from each of these models.

1.2.1 Representation of 2- and 3-neuron motifs relative to random

We asked whether the deviation in the number of two-neuron motifs relative to random that has been reported previously (e.g. [Song et al., 2005]) could be explained by any of the models

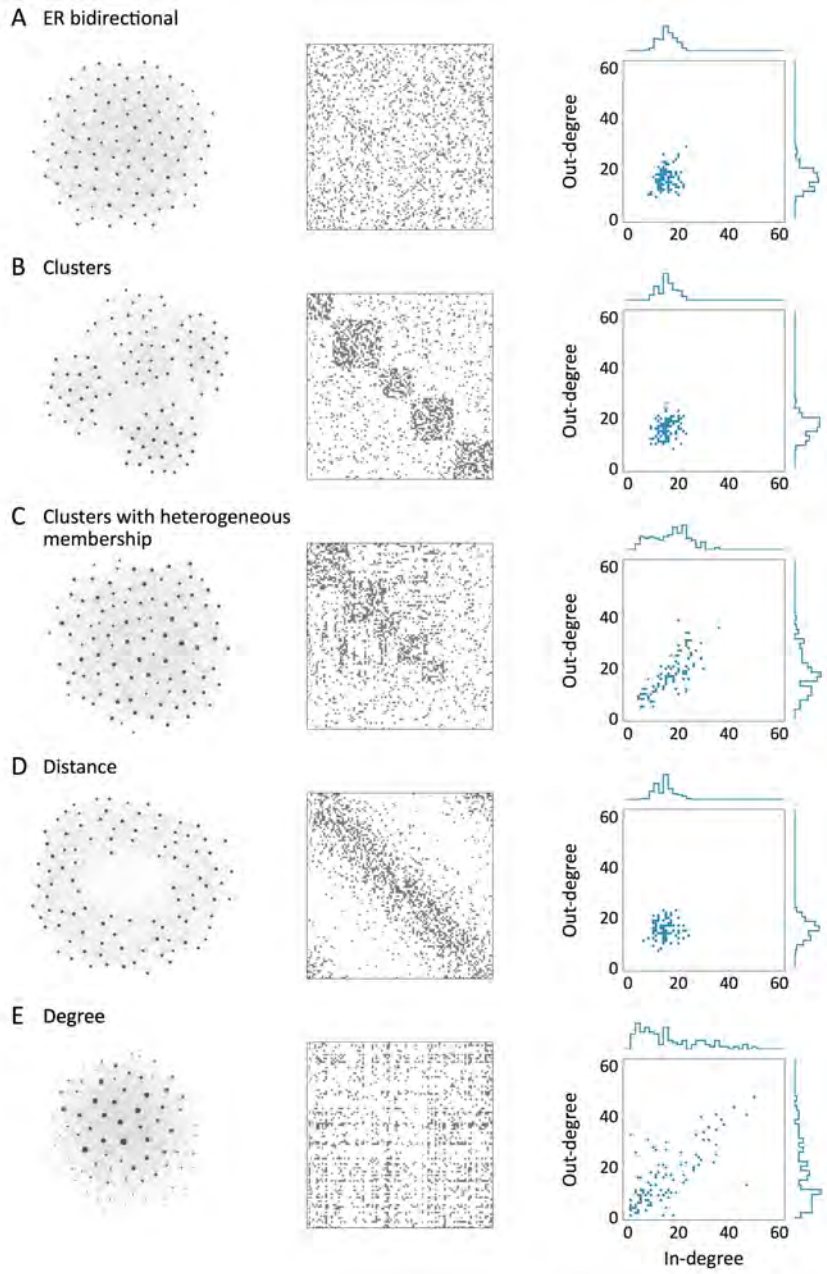

Figure 1.1: Schematic representation of the models: connectivity (left), adjacency matrix (middle) and in/out-degree distribution (right). The nodes in the left column are arranged according to the ForceAtlas algorithm using Gephi software [Bastian et al., 2009]. The size of each node is proportional to the sum of its degrees and the direction of the connections has been omitted for simplicity. In all the networks, $N = 100$, $p = 0.15$, $R = 2$.

Model	p	R
ER bidirectional	$p_{\text{bid}} + \frac{p_{\text{uni}}}{2}$	$\frac{1}{p^2}\ p_{\text{bid}}$
Clusters	$f_+ p_+ + f_- p_-$	$\frac{1}{p^2}(f_+ p_+^2 + f_- p_-^2)$
Distance	$\langle p(r) \rangle$	$\frac{1}{p^2}\ \langle p^2(r) \rangle$
Degree	$\frac{\langle K \rangle}{N}$	$\left(1 + \rho \frac{\sqrt{\sigma_{\text{in}}^2 \sigma_{\text{out}}^2}}{\langle K \rangle^2}\right)^2$

Table 1.1: Sparseness (p) and fraction of bidirectional connections relative to random (R) in the different models. In the models with clusters, $f_- = 1 - f_+$ and f_+ is the fraction of neuronal pairs that are in the same cluster (see Materials and Methods for details). The brackets $\langle\rangle$ in the Dis model represent averages over the distribution of distances in the network. See the main text for a description of the other parameters.

presented here. Given the sparseness p of a network model (that is, the expected number of connections divided by the total number of possible connections), we denote by R the expected number of reciprocal connections relative to that in ER(p), which can be calculated for each model as shown in Table 1.1 (see Materials and Methods for details). The expected number of uni-directionally connected and unconnected pairs is then uniquely determined once p and R are known.

Once p has been fixed, all models can account for a wide range of values in R, including the specific values reported in [Song et al., 2005; Wang et al., 2006; Mason et al., 1991; Markram et al., 1997; Le Bé and Markram, 2006], see Fig. 1.2 B and C (in Fig. 1.2 C we have taken the values of p and R reported in [Song et al., 2005]). The numbers of three-neuron motifs relative to ER-Bi are also qualitatively similar across models, and consistent with experiment, with the exception of ER-Bi which has no additional structure beyond two-neuron motifs (Fig. 1.2 C, bottom).

An important question to be addressed here is to what extent the experimental results are sensitive to the sampling procedure. Data are collected through simultaneous patch-clamp recordings and hence can only record from a small number of cells at a time. The motif counts are local properties whose *averages* do not depend on the sample size, but the results can be highly variable if the number of samples studied is not large enough. In order to mimic the experiment by Song et al. [2005], we computed p and R not only from the study of the whole network but also through 163 samples of 4 neurons per network over 5 networks. As shown in Figs. 1.2 B and C (grey bars), the estimates of the 2-neuron motif counts are quite close to the real counts in networks of $N = 2000$ neurons, which suggests that the magnitudes p and R are well approximated even when only a small fraction of the total network is known. Although the results of 3-neurons motifs were roughly consistent between the full analysis and that from small sample sizes, they were much more variable than the 2-neuron motifs.

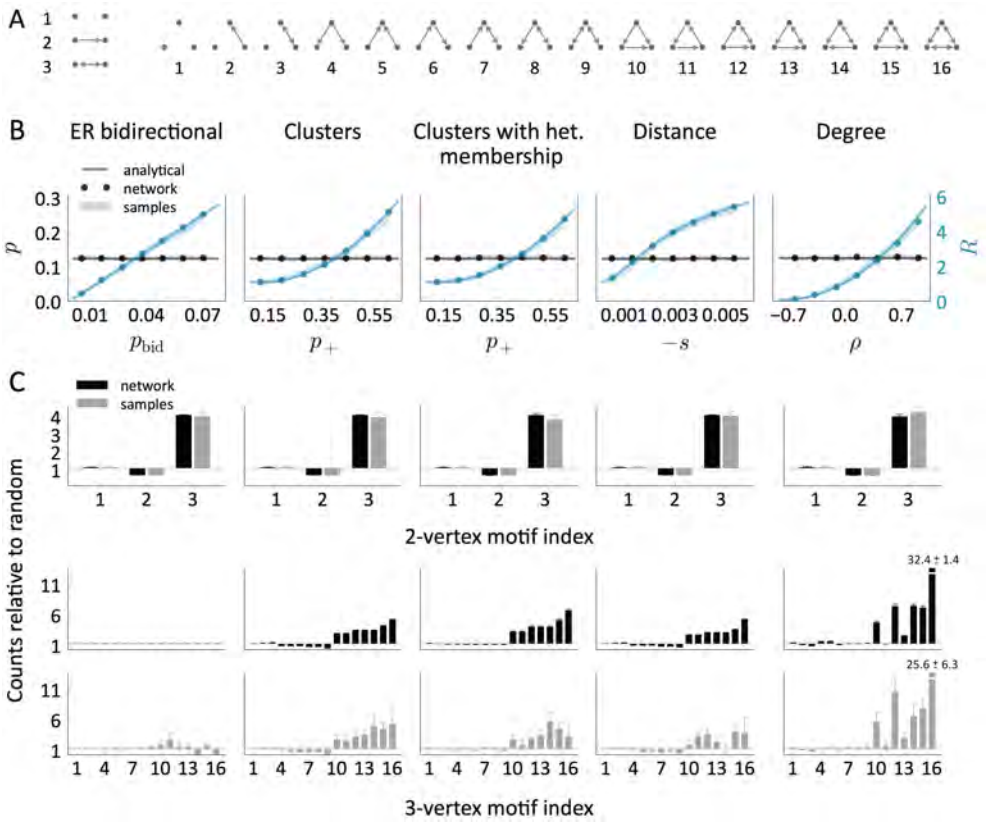

Figure 1.2: Counts of 2- and 3-neuron motifs relative to random models. **A** Representation of all the possible 2- and 3-neuron motifs. **B** Sparseness (p) and expected number of reciprocal connections relative to random (R) as a function of a model parameter. In all the models except the Deg, an additional parameter was varied (p_{uni}, p_-, p_-, t, respectively) to keep p constant. In the Dis model, neurons are arranged in a ring and the connection probability as a function of distance r is defined by the sigmoid function $p(r) = 1 - \frac{1}{1+e^{2s(r-t)}}$, so t is the point where the absolute slope is maximal and $-s$ is this absolute slope. **C** Counts of all the 2- and 3-neuron motifs relative to random models (ER and ER-Bi, respectively) in networks with $p = 0.12$, $R = 4$. We used 5 different networks of size $N = 2000$ per condition. The computations were performed both on the whole network and on 163 samples of size 4 per network. Shaded regions and error bars indicate mean ± SEM.

Nonetheless, at least in the example networks shown in Fig. 1.2 C, it seems that the particular distribution of triplet motifs might provide a means of classifying the different models. In subsequent sections we will show that there is a particular combination of dual and triplet motifs from which we can extract information about the network class, independently of the choice of other parameters.

1.2.2 Connection probability as a function of the number of common neighbors

A common neighbor to neurons i and j is a third neuron which is connected to both i and j. Perin et al. [2011] have shown that the probability of connection between pairs of cortical neurons increases with the number of common neighbors they have (the so-called "neighbor rule"). Fig.

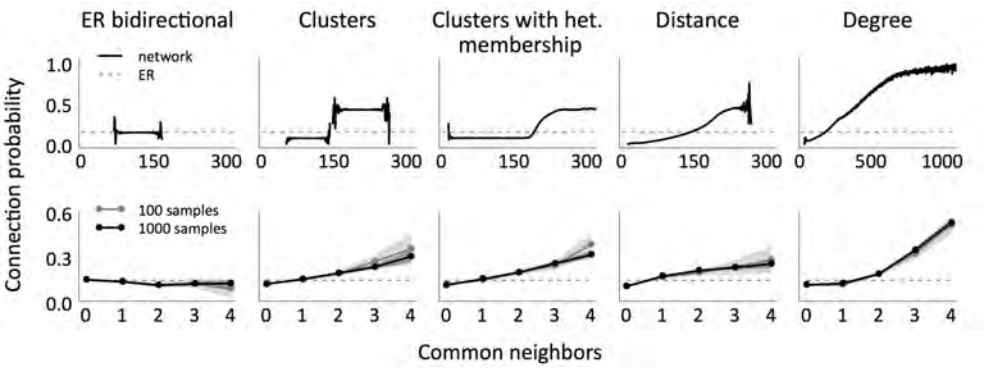

Figure 1.3: Connection probability as a function of the number of common neighbors for the different models, in the whole network (top) and in samples of size 12 (bottom). In all the cases, $N = 2000$, $p = 0.14$, $R = 2$. The analyses were performed on 5 networks and the shaded regions indicate the resulting mean $\pm$ SEM. In the sample analyses we took 20 samples per network (100 in total, grey) and 200 samples per network (1000 in total, black). The dotted lines show the expected probability if it were independent of the number of common neighbors, as in the ER and ER-Bi models.

1.3 (top) shows the connection probability as a function of common neighbors for examples from each model class from the analysis of a network of $N = 2000$ neurons where p and R are close to the values reported in [Perin et al., 2011]. In the ER-Bi model, as in the classical ER model, all the pairs are connected independently and according to the same rule, so the number of common neighbors does not provide any information about the "laws" controlling a given connection. All the other models, however, exhibit the common neighbor rule for a general choice of the network parameters. Interestingly, the precise shape of this dependence is quite distinct for different models, indicating it might provide a signature for inferring the full network structure from this one measure. However, these qualitative differences between models largely vanish when realistic sample sizes are analyzed (Fig. 1.3, bottom). It is important to keep in mind that the curves shown in Fig. 1.3 are for a particular choice of network from each model class. The exact shape of the curves will depend on that choice. In general, we can say that given small sample sizes one will observe a monotonically increasing dependence of the connection probability on the number of common neighbors for all models but ER-Bi. Specifically, for clustered (distance dependent) models, neuron pairs with more common neighbors are more likely to belong to the same cluster (be closer together), which increases the probability of connection. In the Degree model neuron pairs with more common neighbors are more likely to have large degrees, which again increases the probability of connection.

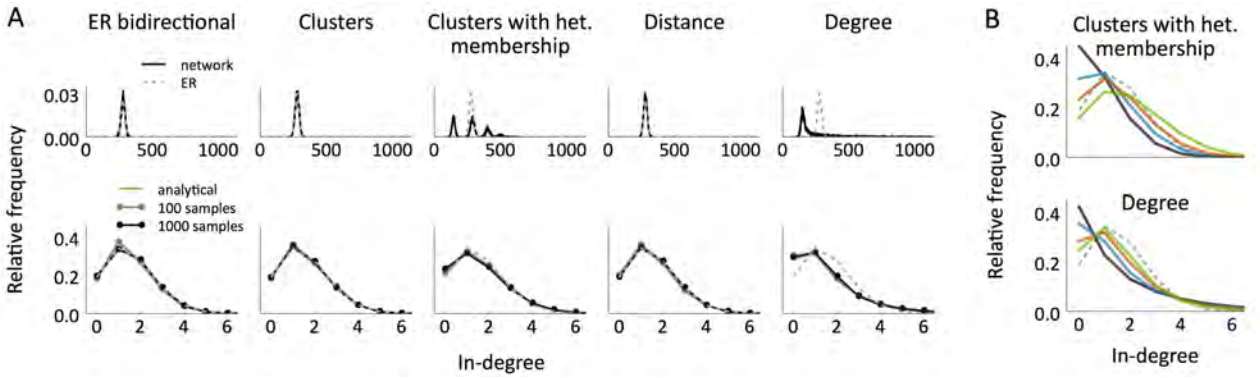

Figure 1.4: In-degree distribution of the different network models. **A** In-degree distribution in the whole network (top) versus in-degree distribution in samples of size 12 (bottom) and comparison with the distributions exhibited by the ER model (dotted lines). The networks and samples used are the same as in Fig. 1.3. The shaded regions indicate mean ± SEM. **B** In-degree distributions in samples of size 12 for different networks generated according to the Cl-Het (top) and the Deg (bottom) models, all of them with $N = 2000$, $p = 0.14$, $R = 2$.

1.2.3 Degree distributions and higher-order connectivity

Figure 1.4 A (top) shows the in-degree distributions exhibited by example networks from the different models for physiological values of p and R. For both the Cl-Het and Deg models the distribution differs dramatically from that of the equivalent ER network. Nonetheless, and as was the case with the common-neighbor rule, when the distributions are constructed from realistic sample sizes (here 12), all models are qualitatively similar, see Fig. 1.4 A (bottom). In fact, due to additional degrees of freedom that both the Cl-Het and the Deg models have, it is possible to define networks with a fixed p and R whose distributions are nevertheless very different (Fig. 1.4 B). In some situations, the distribution is quite close to ER/ER-Bi cases.

Finally, real data also exhibit a significant over-representation of densely connected groups [Perin et al., 2011]. We therefore also studied the distribution of the number of connections in small groups of neurons and found that all models, with the exception of ER-Bi, could account for these findings, see Fig. 1.5.

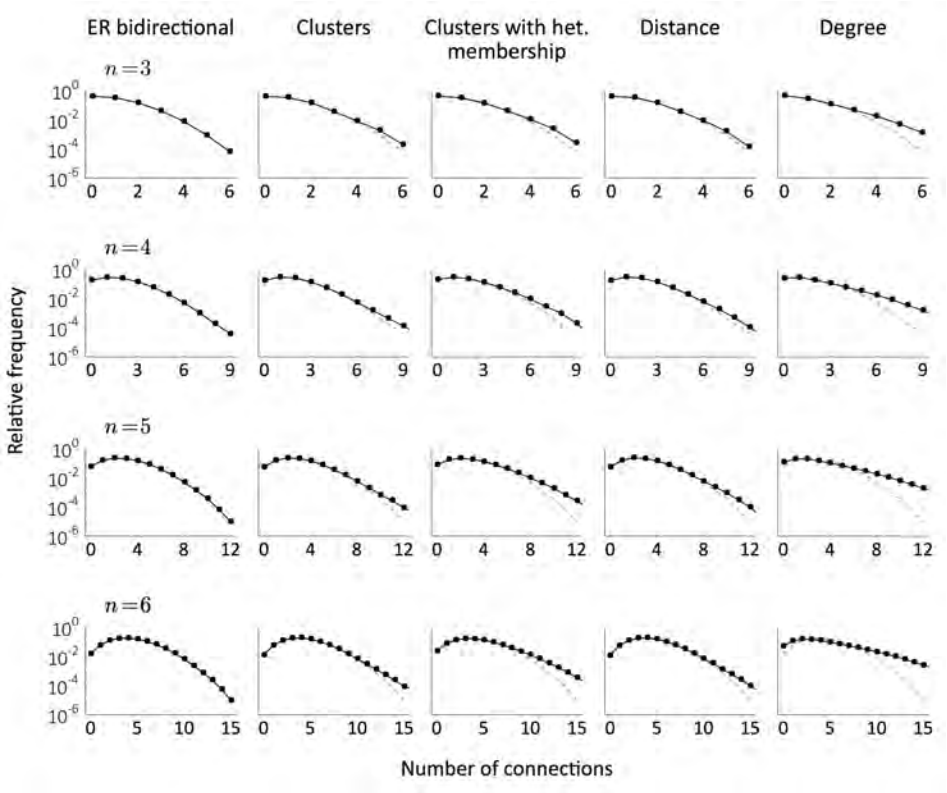

Figure 1.5: Distribution of the total number of connections in samples of sizes $n \in \{3, ..., 6\}$ for the different models (black) compared to the distribution obtained in ER bidirectional networks with the same p and R (dashed grey). The parameters are the same as in Figs. 1.3 and 1.4. The analyses were performed on 5 networks per condition and the computations come from 10^5 random samples for each network.

1.3 A method for distinguishing between network models using measures from small sample sizes

1.3.1 The sample degree correlation (SDC)

We sought a measure, based on small sample sizes, which would allow us to distinguish between the classes of topological models defined here. In other words, we looked for a way to infer general topological properties of the network when only local information is available. We found such a measure in the *sample in/out-degree correlation*

$$\mathrm{SDC} = \frac{\mathrm{Cov}(k_i^{\mathrm{in}}, k_i^{\mathrm{out}})}{\sigma^2}, \tag{1.3.1}$$

where i represents a random neuron from the sample, k_i^{in} and k_i^{out} are its in- and out-degrees in the sample and $\sigma^2 = \sqrt{\mathrm{Var}(k_i^{\mathrm{in}})\mathrm{Var}(k_i^{\mathrm{out}})}$. The SDC therefore depends on the variances and covariances of the sample degrees. The in-(out-)variance in turn depends on the occurrence of convergent (divergent) motifs, while the covariance depends on the occurrence of chain and reciprocal motifs. All of these quantities can be calculated analytically for the network classes we have considered here, and the SDC is finally expressed as a function of p, R, σ^2 and the sample size n, as shown in Materials and Methods.

In particular, we can group the five network types into three classes based on the functional form of the SDC: (1) The ER-Bi, Cl, and Dis models, (2) The Cl-Het model and (3) the Deg model, see Materials and Methods for details. We can additionally use the common-neighbor rule to distinguish between the ER-Bi (which shows no dependence) and the Cl and Dis models (which do). It can be shown that all these classes of networks have SDC $\equiv 0$ whenever $R = 1$, which means that networks that do not show an over-representation of bidirectional connections cannot be distinguished in terms of the SDC. Therefore, as long as $R > 1$, in principle we can distinguish between all models, except for the Cl and Dis models. This is not surprising given that the Cl is nothing but a particular case of the Dis where the distance is binary.

We applied this "SDC criterion" to networks of size $N = 2000$ generated randomly according to the four classes of models presented here (grouping Cl and Dis), with p and R chosen uniformly in the ranges $[0.05, 0.23]$ and $[1.5, 4.1]$, respectively. We used the SDC to distinguish between the different model classes by taking samples of size $n' = 12$. To do so, we first estimate p and R from the sample set and then the SDC and σ^2 for each $n \leq n'$. We finally compare the computed

SDC at each n with the predictions for the three different models, see Fig. 1.6 B and Materials and Methods for details.

We computed σ^2 and the SDC over this range of sizes by using the estimated values of p, R and the occurrence of convergent, divergent and chain motifs (through the quantities $Conv$, Div and $Chain$ defined in Materials and Methods, Eq. (1.6.16)). An alternative approach is to generate n-neuron subsamples from the original samples of size n' and directly compute σ^2 and the SDC for each $n \leq n'$. We also checked that the performance is almost the same for the second method when $n' = 12$ (data not shown). The advantage of estimating σ^2 and the SDC instead of calculating them directly is that it allows one to implement the criterion even when the original samples are small (e.g. $n' = 3, 4$). To further distinguish between the ER-Bi and Cl/Dis classes we studied if the connection probability increases with the number of common neighbors in the n' samples.

The efficacy of this classification criterion increases with the number of samples considered, m. Fig. 1.6 C, D shows the performance as a function of m. The rate of success is above the chance level (chance here is 25%) for all models already for $m = 2$ samples and reaches 94% for $m = 300$. As long as the network size is large compared to the sample size, the classification accuracy is independent of system size (Fig. 1.7). This simply means that it can be applied to real data without the need to worry about the true size of functional cortical circuits.

1.3.2 Analysis of the SDC in data from rat somatosensory cortex

We implemented our SDC criterion in the data obtained by Perin et al. [2011] from pyramidal neurons of the rat somatosensory cortex. The data come from 6, 9, 5, 10 and 10 groups of 8, 9, 10, 11 and 12 neurons, respectively. As previously reported in [Perin et al., 2011], these data show a clear dependency of connection probability on intersomatic distance. The estimated connection density and number of reciprocal connections relative to random were $p = 0.144$, $R = 2.575$. The analysis of the SDC revealed a relationship which deviates from any of the previously defined models, as shown in Fig. 1.8 A. Although the form of the SDC appears close to that of the Dis model (Fig. 1.8 A left), the degree variance from the data σ^2, which should be that of a Binomial distribution, differs strongly from the theoretical value (Fig. 1.8 A right). Note that the degree variance for the other two classes of network is a free parameter and hence here is estimated directly from the data.

This finding could be interpreted to mean that the SDC is not a sufficient criterion to discard the

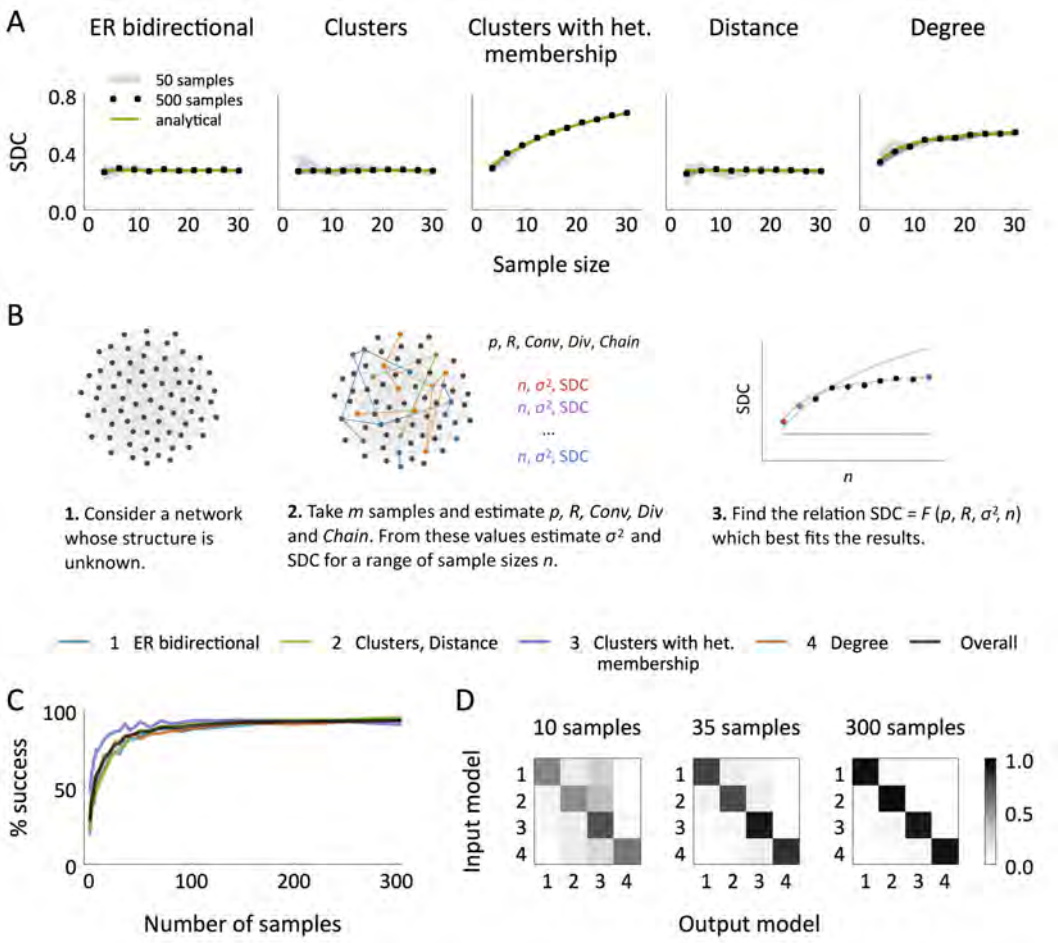

Figure 1.6: Sample in/out-degree correlation (SDC) as a measure to distinguish between classes of networks. **A** SDC in samples of 3 to 30 neurons in the different models. In all the networks, $N = 2000$, $p = 0.12$, $R = 3$. We computed the empirical correlations using 50 and 500 samples per network for each sample size. Every analysis was performed independently in 5 different networks and the shaded region indicates the resulting mean ± SEM. **B** Schematic representation of the algorithm proposed to distinguish between the model classes: (1) ER-Bi, (2) Cl/Dis, (3) Cl-Het and (4) Deg. **C** Success rate of the algorithm performed on randomly generated networks with $N = 2000$, $p \in [0.05, 0.23]$, $R \in [1.5, 4.1]$ as a function of the number of samples considered m. All the samples had size $n' = 12$. Each success rate was computed over 2000 experiments. **D** Frequencies of all the possible input-output combinations in the experiments shown in C, for three choices of the number of analyzed samples. Each frequency is normalized by the frequency of the input model so that the sum of every row is 1.

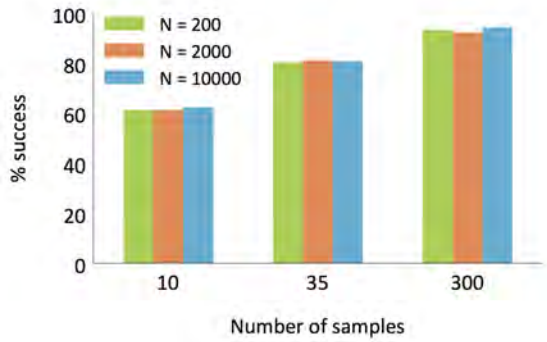

Figure 1.7: Overall success rate of the SDC criterion as a function of the size of the network N and the total number of 12-neuron samples considered.

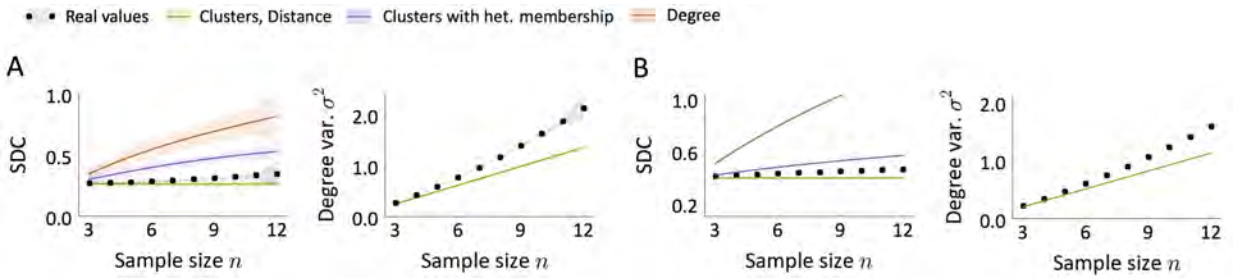

Figure 1.8: Sample in/out-degree correlation SDC and geometric mean of the sample degree variances σ^2 as a function of the sample size n. **A** Values calculated directly from the data of [Perin et al., 2011]. **B** Inferred values from the motif counts presented in [Song et al., 2005]. The black curves correspond to the observed SDC and σ^2, whereas colors show the expected SDC (σ^2) in networks generated according to the studied models with the same p, R, σ^2 (p, R) as in data. Shaded regions in **A** indicate mean ± SEM computed with the Bootstrap method.

Cl/Dis family (it seems that we need to know σ^2 as well). But this is not the case: if the data came from a Cl/Dis network, the computed SDC would be fitted equally well by the Cl/Dis and Cl-Het formulas (just because Cl/Dis models are a subclass of the Cl-Het model). Therefore, the green and purple lines of Fig. 1.8 A should coincide. The fact that they differ from one another indicates that σ^2 deviates from what we would expect for Cl/Dis networks.

Since the SDC can be extrapolated when the counts of two- and three-neuron motifs are known, we calculated the expected SDC in putative samples of 3 to 12 neurons from the motif distribution described in [Song et al., 2005] (Fig. 1.8 B), which corresponds to layer 5 pyramidal neurons in rat visual cortex. The connection density and the number of reciprocal connections relative to random in this case are $p = 0.116$ and $R = 4$. The results are qualitatively similar to the ones computed directly from the data of [Perin et al., 2011]. This suggests the underlying network structure itself may be similar.

1.4 A general class of network model

We discovered that all of the models, with the exception of the ER-Bi model, which could be rejected already by its failure to capture triplet motifs and the neighbor rule, belong to a more general class of model. Specifically, in what we dub *Modulator* networks, the probability of a connection from neuron i to neuron j is $P(i \to j \,|\, x_i = x, x_j = y) = g(x, y)$, where x_i and x_j, the *modulators*, are properties associated with neurons i and j and g is a *modulatory function*. We assume that the modulators of individual neurons are independent random variables that come from a common distribution. These properties (or sets of properties) might represent spatial position, axonal/dendritic length, neuronal type defined by the expression of some proteins, presence

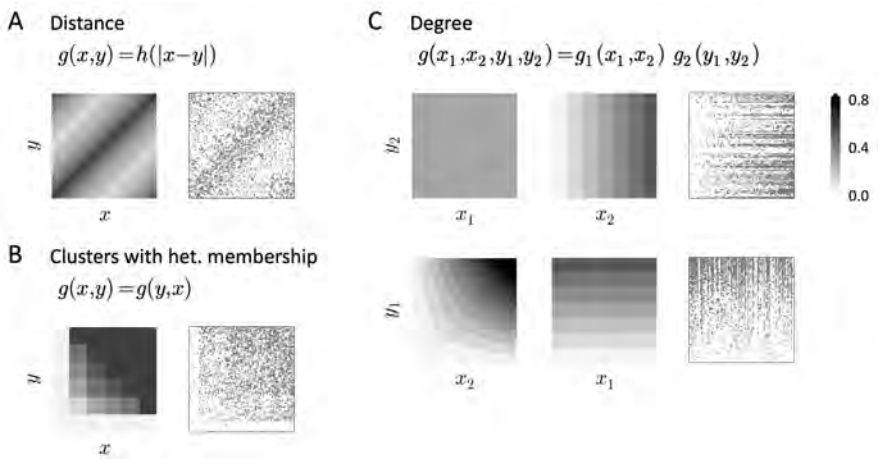

Figure 1.9: Examples of modulatory functions g and adjacency matrices of the three main classes of networks described in this study. Notice that the row ordering in the adjacency matrices has been inverted to be coherent with the g plot. **A** In the Dis model, the modulators x and y represent the spatial position of pre- and post-synaptic neurons, respectively. **B** In the Cl-Het model, x and y represent the number of clusters to which pre- and post-synaptic neurons belong. **C** In the Deg model, (x_1, x_2) represents the pair of in- and out-degrees of the pre-synaptic neuron and (y_1, y_2) are the degrees of the post-synaptic neuron. The adjacency matrices result from ordering neurons according to their out-degree (top) and their in-degree (bottom). See Materials and Methods for details and parameter values used.

of neuromodulators in the medium, amount of input received from other brain areas, stimulus selectivity or even information related to the past history of neurons, to cite just some possibilities. The Modulator model, therefore, represents any general scenario in which connections appear with higher or lower probability depending on features of the two neurons involved. The models we have considered so far are special cases of this more general modulator framework. This is illustrated in Fig. 1.9, which shows three sample networks from the Dis, Cl-Het and Deg classes.

In the clustered and distance-dependent models that we have considered, $g(x, y) = g(y, x)$ is reflection symmetric. In this case the modulators are the position or membership in a cluster (or group of clusters), e.g. Fig. 1.9 B. It can be shown that any Modulator network with a symmetric g exhibits the same SDC as the Cl-Het model. If, additionally, $g(x, y)$ can be assumed to be independent from one neuronal pair to another (as in our Cl and Dis models when a small sample is considered, where $g(x, y)$ only depends on the distance $|x - y|$, see Fig. 1.9 A), the formula reduces to the Cl/Dis case. In the Deg model g is separable, i.e, $g(x, y) = g_1(x)g_2(y)$, and the modulator itself is the pair of in- and out-degrees. The g function is just the product of the pre-synaptic out-degree and post-synaptic in-degree, normalized by the appropriate factor (Fig. 1.9 C). In Materials and Methods we show that the SDC of any Modulator network with separable g has the form of the SDC of the Deg model.

Therefore, the SDC criterion not only makes it possible to distinguish between the families Cl/Dis,

Cl-Het and Deg, but allows for a classification into three major types of Modulator networks, defined by different properties and symmetries. The fact that the data are not fit by any of the models indicates that real cortical circuits have features which violate the reflection symmetry and separability of the function g.

Since the estimated SDC lies in between the predicted SDC for the Dis/Cl and Cl-Het models (Fig. 1.8 A), one would be tempted to think that a hybrid network from these two classes would be compatible with data. Such a model, however, would still belong to the class of Modulator networks with symmetric g and would therefore exhibit the same SDC as the Cl-Het class (purple line in Fig. 1.8 A). This suggests that not only is there additional structure in the data beyond the distance dependence of connection probabilities, but that this structure is not simple clustering.

1.4.1 Data are consistent with a network with spatial dependence and hierarchical clustering

We were able to obtain an excellent fit to all relevant topological statistics in the data with a Modulator network. Specifically, we considered a network in which the probability of connection between pairs was

$$P(i \to j \,|\, x_i = x, x_j = y, r_{ij} = r) = p(r)\, g(x, y), \tag{1.4.1}$$

where $p(r)$ depends on the physical distance r between pairs, and the modulator component $g(x, y)$ is not reflection symmetric. This model is itself a two-dimensional Modulator network in which one dimension is physical space, and the other represents a property of the neurons not captured by their spatial location, see Fig. 1.10 A. We assumed that the distribution of distances in samples obtained from the model is close to the sampled distribution in the data (Fig. 1.10 B, left) and that the $\{x_i\}_i$ modulators are independent of distances. We assume a Gaussian distribution of the modulator and take $g(x, y)$ to be the weighted sum of the p.d.f. of two bivariate Gaussians, one of which breaks the reflection symmetry, see Fig. 1.10 A and Materials and Methods for details. This choice is equivalent to other possible distributions of the modulator as long as g is also properly transformed. The model successfully captures the observed distance-dependency of the connection probabilities (Fig. 1.10 B right). Note, in particular, that it reproduces the over-representation of reciprocal connections as a function of distance (Fig. 1.10 B right inset). A pure Dis model cannot explain this finding; although the value of R evaluated globally would be greater than 1, for any given distance it would be identically 1. Therefore, the increased R as a function of distance is a clear signature of additional structure, captured here by our modulator

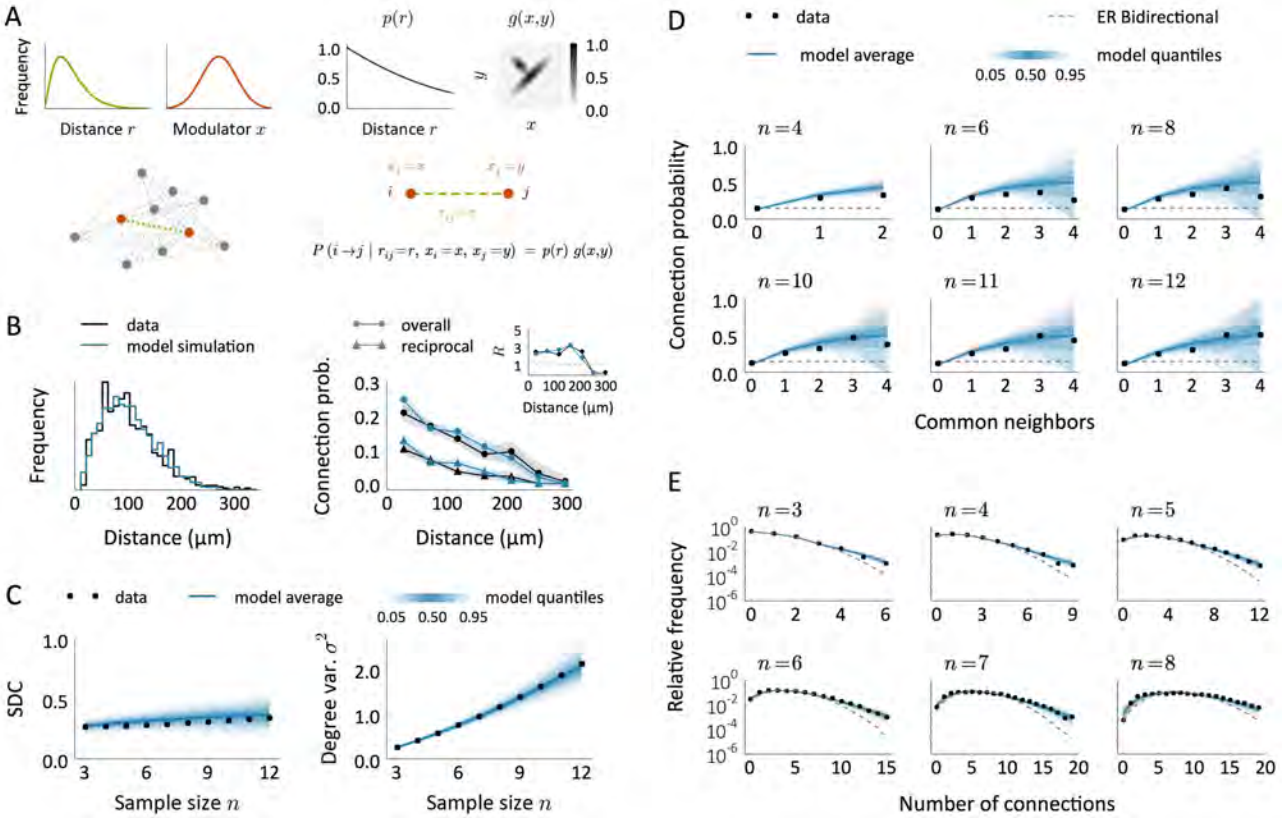

Figure 1.10: **A** Schematic of a model to explain the observed data. First, neurons are arranged in space so that distances between neuronal pairs follow a given distribution (green). Each neuron has also an associated modulator whose distribution is shown in red. Given a distance-decaying probability $p(r)$ and a function $g = g(x, y)$, connections are created independently with probability $P(i \to j \,|\, r_{ij} = r,\, x_i = x,\, x_j = y) = p(r)\, g(x, y)$. **B** Intersomatic distance distribution and connection probabilities as a function of distance in the data (black) and in the model (blue). Inset: number of reciprocal connections relative to random R as a function of distance. The model results come from a single replica of the real experiment and shaded regions indicate mean ± SEM. **C** Sample degree correlation SDC and geometric mean of the sample degree variances σ^2 as a function of sample size n in the data (black) and in the model (blue). The blue shaded regions indicate quantiles computed from a set of 200 replicas of the real experiment, each performed on an independent network. **D, E** Comparison between model and data in terms of the common neighbor rule (D) and the distribution of the total number of connections (E) in samples of size n. Dashed lines show the prediction for ER-Bi networks.

function. The Modulator model also reproduces the sample degree correlation and variance (Fig. 1.10 C), as well as the common neighbor rule and the density of connections in groups of few neurons (Fig. 1.10 D and E).

What is the interpretation of the modulator in this network? The modulator acts as an identifier for each neuron, and neurons with similar modulators will connect in similar ways. Indeed, if the modulator is symmetric we recover a continuous version of a clustered network with heterogeneous membership (Cl-Het). Therefore, the symmetric part of $g(x, y)$ (see plot in Fig. 1.10 A) can be interpreted as clustering: neurons with similar values of x are more likely to connect to one-another than to neurons with different values (although this preference decreases for extreme values of x). However, the presence of asymmetry in g indicates that connections between clusters are actually hierarchical. Specifically, in our example, neurons with low x are likely to connect to similar neurons, and also to neurons with large x. On the other hand, neurons with large x are likely to connect with similar neurons, but not to neurons with low x. We further checked that this is actually captured by a model where the distance-independent modulatory component is based on discrete hierarchical clustering. It was sufficient to consider a homogeneous distribution into three clusters where connection probability within cluster 2 is higher than within clusters 1 and 3 and where connection probabilities between different clusters are low except for cluster 1, which has a strong preference to project to cluster 3, as in our continuous model (data not shown). Although these two model versions are essentially the same, the continuous one incorporates a higher variability in the modulatory variables which could resemble real modulatory mechanisms that operate through continuous variables such as concentration of molecules or the amount of input received from other brain areas. In conclusion, the data are consistent with a network in which neurons are connected according to the physical distance between them and their membership in a clustered structure, independent of distance, which itself exhibits hierarchical features.

1.5 Conclusion

We have presented three major classes of network models that are compatible with the "nonrandomness" reported so far in cortical microcircuits [Song et al., 2005; Perin et al., 2011]. The first is based on a similarity principle: pairs of neurons have associated a notion of distance which modulates the likelihood of the connections between them, in the sense that similar neurons tend to be connected more frequently than different ones. The connections appear independently once the distances between neuronal pairs are known. Distance in this context can represent not only

a spatial proximity but any other measure of similarity, for example based on input received from other areas or stimulus selectivity. This family also includes networks where neurons are classified homogeneously into clusters so that connections form preferentially between cells that are in the same cluster. In the second model, neurons are assigned to clusters but there is heterogeneity both in the cluster size and in the number of clusters to which different neurons belong. Connections form with higher likelihood between neurons that coincide in any of the clusters. The third family corresponds to networks where in- and out-degrees of single neurons follow a prescribed joint probability distribution.

Our results show that the three classes of networks can exhibit both an excess of reciprocal connections relative to random and the so-called common neighbor rule for a wide range of parameters. In the case of networks with a specified degree distribution, in- and out-degrees must be positively correlated for the bidirectional connections to be over-represented, meaning that neurons that receive more synapses from the network tend to be the ones that have more outgoing connections, i.e. they are hubs. All of the models can also be similar in terms of the marginal degree distribution in small samples and are in qualitative agreement with previously reported results concerning the number of connections in groups of few neurons. The first important conclusion of our study is therefore that these nonrandom features, rather than being a footprint of a specific topology, seem to arise naturally from several qualitatively distinct types of models.

One of the major difficulties of inferring structural principles from real data is that functional neuronal networks likely encompass thousands of neurons, whereas simultaneous patch-clamp experiments, which provide ground truth for synaptic connectivity, provide samples of only a few neurons at a time. Although the models presented here are based on very different principles, they can be almost indistinguishable from one another given only small sample sizes. Thus, even structures that are distinct globally can exhibit similar properties locally.

A natural question is whether it is possible to define a local measure −i.e., a measure that can be estimated from the study of small samples− that could be used to distinguish between models. We have found such a measure in the *sample degree correlation* (SDC), the correlation coefficient between sample in- and out-degrees. The SDC is, in fact, a particular nonlinear combination of triplet motifs which allows us to correctly classify network models without recourse to training classifiers numerically. Interestingly, the SDC depends on precisely those second-order network statistics which have been recently used to develop dynamical mean-field models for neuronal networks with structure beyond the ER network [Zhao et al., 2011; Nykamp et al., 2017].

Note that a machine learning approach to this problem would require training a classifier on particular instantiations of networks from a given network class; each class encompasses a vast range of possible networks. Therefore, training sets would not likely be representative of the class as a whole. A major advantage of our approach, in contrast, is that it allows us to classify networks regardless of the details of every model candidate, which can be difficult to estimate in real situations. For example, in the Distance model the exact shape of the function $p(r)$ is irrelevant for estimating the SDC, which only depends on the overall connection probability and the over-representation of reciprocal connections. We have also shown that these three model classes are particular cases of a very general model according to which single neurons have an associated property that modulates the connection probability. We call such a property a *modulator*.

We estimated the SDC for distinct data sets from both rat somatosensory cortex and rat visual cortex and found that the structure in those cortical circuits fell outside all three classes of model network in a qualitatively similar way, see Fig. 1.8. These observations therefore suggest that if the underlying network topology can be interpreted in the Modulator framework, then the modulatory function $g(x, y)$, which defines the probability of finding a connection from a neuron with modulator x to a neuron with modulator y, can be neither symmetric nor separable.

Finally, we obtained an excellent fit to the first data set by considering a more general Modulator network in which the probability of connection between neurons depended both on the physical distance between them, as well as on an additional modulator unrelated to distance. In the second data set there is no evidence of distance dependency of connectivity [Song et al., 2005] but the qualitative similarity between data sets in terms of the SDC suggests that a similar non-spatial modulator mechanism might be common to both of them. The structure of this non-spatial modulator could be interpreted as hierarchical clustering, in which connectivity between clusters is asymmetric. However, we cannot rule out that other choices of modulators, which would lead to other interpretations, might provide equally good fits to the data.

The classes of networks that we have explored here are simple enough to be treated analytically. Nature is certainly more complex, and clearly cortical microcircuits are shaped by other principles, including ongoing synaptic plasticity. We have not considered these mechanisms here. Nevertheless, independent of the mechanisms which shape cortical microcircuitry, if the topology of the resultant network can be reduced to a modulatory mechanism, then our results show that this modulation involves both a distance dependence and an additional non-spatial component which is asymmetric.

1.6 Materials and Methods

1.6.1 Heterogeneity in Cl and Cl-Het models

The clustered networks that we have implemented have $C \ll N$ clusters. In the Cl model, each neuron belongs to one single cluster, which is chosen uniformly at random. Therefore, the expected fraction of pairs in the same cluster is $f_+ = \frac{1}{C}$.

In the Cl-Het model, each neuron has a probability $p_c = 1/C$ of belonging to any given cluster and the expected fraction of pairs in the same cluster is

$$f_+ \quad = \quad 1 - (1 - p_c^2)^C. \tag{1.6.1}$$

If n_i is the number of neurons that are at least in one of the clusters of i,

$$\begin{array}{lcl} \mathbb{E}[n_i] & = & (N-1)f_+, \\ \mathrm{Var}(n_i) & \approx & (N-1)[(N-2)(2f_+ - 1 + (p_c^3 - 2p_c^2 + 1)^C) - (N-1)f_+^2 + f_+], \end{array} \tag{1.6.2}$$

so, if C is fixed and N is large,

$$\frac{\sqrt{\mathrm{Var}(n_i)}}{\mathbb{E}[n_i]} \quad \approx \quad \sqrt{\frac{(N-2)(2f_+-1+(p_c^3-2p_c^2+1)^C)-(N-1)f_+^2+f_+}{(N-1)f_+^2}} \quad \to \quad \sqrt{\frac{(p_c^3-2p_c^2+1)^C+2f_+-1}{f_+^2} - 1} \quad > \quad 0.6 \tag{1.6.3}$$

for $C \geq 2$. This means that there is a non negligible variability across neurons in terms of cluster membership, contrary to what happens to the Cl model, which fulfills Eq. (1.2.1).

1.6.2 Properties of the Deg model

Connection probability once the degrees are known

The degree sequence $S = \{(K_i^{\mathrm{in}}, K_i^{\mathrm{out}})\}_{i=1}^N$ of any directed network has the property that the sum of all the in-degrees equals the sum of all the out-degrees. We denote this sum by M, which gives the total number of connections (edges) in the network:

$$M \quad = \quad \sum_{i=1}^{N} K_i^{\mathrm{in}} \quad = \quad \sum_{i=1}^{N} K_i^{\mathrm{out}}. \tag{1.6.4}$$

Here we consider sparse networks, and in particular we assume that the statistics of the degree sequence are fixed as N increases. Given a degree sequence S, a realization of the Deg model

is obtained by uniformly choosing an assignment of incoming edges to outgoing edges. This is equivalent to choosing a random permutation of M elements. Such permutation defines how we have to pair the incoming edges with the outgoing edges to construct the network (notice that the family of networks generated in this way contains graphs with self-loops and multiple edges; these networks form, however, a negligible fraction in the large N limit). Then, the probability that vertex (or neuron) i connects to j equals the probability that a randomly chosen permutation maps one of the first K_i^{out} elements to one of the first K_j^{in} elements. This probability is

$$\begin{aligned} f(K_j^{\text{in}}, K_i^{\text{out}}, M) &= 1 - \tfrac{M-K_j^{\text{in}}}{M} \cdot \tfrac{M-K_j^{\text{in}}-1}{M-1} \cdot \ldots \cdot \tfrac{M-K_j^{\text{in}}-K_i^{\text{out}}+1}{M-K_i^{\text{out}}+1} \\ &= 1 - \tfrac{(M-K_j^{\text{in}})!\,(M-K_i^{\text{out}})!}{M!\,(M-K_j^{\text{in}}-K_j^{\text{out}})!}. \end{aligned} \tag{1.6.5}$$

Expanding $f(x, y, M)$ in powers of $1/M$ we have

$$\begin{aligned} f(x,y,M) &= \tfrac{c_1(x,y)}{M} + \tfrac{c_2(x,y)}{M^2} + \tfrac{c_3(x,y)}{M^3} + \mathcal{O}(1/M^4), \\ c_1(x,y) &= xy, \\ c_2(x,y) &= -\tfrac{1}{2}xy(x-1)(y-1), \\ c_3(x,y) &= \tfrac{1}{6}xy(x-1)(y-1)(1-2x-2y+xy). \end{aligned} \tag{1.6.6}$$

In the sparse limit considered here, the magnitude of these coefficients does not depend on N, whereas M scales linearly with N. Thus, we can consider the first order approximation

$$P(i \to j \,|\, K_j^{\text{in}} = k, K_i^{\text{out}} = k', M = m) \simeq \frac{kk'}{m}. \tag{1.6.7}$$

This expression can be used only *once* we know which is the total number of edges in the network, M. In practice, we deal with the family of networks whose degree sequence follows a given distribution, so M is a random variable and we have to replace $1/m$ in (1.6.7) by $\mathbb{E}[1/M]$. In general it is not true that the expectation of the reciprocal of a random variable equals the reciprocal of the variable's expectation, but taking into account that the standard deviation of M relative to its mean tends to zero as $N \to \infty$, we can approximate $\mathbb{E}\left[\frac{1}{M}\right] \simeq \frac{1}{\mathbb{E}[M]}$, which leads to the following approximation for the connection probability:

$$P(i \to j \,|\, K_j^{\text{in}} = k, K_i^{\text{out}} = k') \simeq \frac{kk'}{\mathbb{E}[M]} = \frac{kk'}{N\langle K\rangle}. \tag{1.6.8}$$

The connection density in this model is clearly

$$p \;=\; \tfrac{\langle K\rangle}{N}. \tag{1.6.9}$$

Frequency of bidirectional connections relative to random

In a network with $M = m$ edges, the probability that a bidirectional connection exists between two vertices i and j once their degrees are known is

$$P(i \leftrightarrow j \,|\, K_i^{\text{in}} = k, K_j^{\text{in}} = l, K_i^{\text{out}} = k', K_j^{\text{out}} = l', M = m) \;=\; f(k, l', m) f(l, k', m-1), \tag{1.6.10}$$

which leads to

$$P(i \leftrightarrow j \,|\, K_i^{\text{in}} = k, K_j^{\text{in}} = l, K_i^{\text{out}} = k', K_j^{\text{out}} = l') \;\simeq\; \frac{kk'll'}{\mathbb{E}[M]^2} \tag{1.6.11}$$

and

$$P(i \leftrightarrow j) \;\simeq\; \frac{\mathbb{E}[K^{\text{in}} K^{\text{out}}]^2}{\mathbb{E}[M]^2}. \tag{1.6.12}$$

Using Eq. (1.6.9) for the sparseness in this model, the expected fraction of bidirectional connections relative to random is

$$R \;=\; \frac{P(i \leftrightarrow j)}{p^2} \;\simeq\; \frac{\mathbb{E}[K^{\text{in}} K^{\text{out}}]^2}{\langle K\rangle^4} \;=\; \left(1 + \rho \frac{\sqrt{\sigma_{\text{in}}^2 \sigma_{\text{out}}^2}}{\langle K\rangle^2}\right)^2, \tag{1.6.13}$$

where σ_{in}^2, σ_{out}^2 and ρ stand for the in/out-degree variances and the Pearson correlation coefficient of individual in/out-degrees, respectively.

1.6.3 Properties of the Modulator model

In the Modulator model each neuron i has an associated parameter x_i and the connections are made independently with probability

$$P(i \to j \,|\, x_i = x, x_j = y) = g(x, y), \tag{1.6.14}$$

where $\{x_i\}_{i=1}^N$ are independent and identically distributed random variables. All the previous models except the ER-Bi can be interpreted, at least locally, as particular cases of this model.

In clustered networks (Cl and Cl-Het), x_i denotes the cluster membership of neuron i, whereas in the Dis model, x_i represents the "position" of neuron i. In both of these cases the connection probability depends on a notion of distance between pairs, so the function g is symmetric: $g(x, y) = g(y, x)$. Moreover, in a random sample of the Cl and Dis models, coexistence in a cluster or distance can be assumed to be independent from pair to pair, as long as the sample size is small compared to the network size. In the Cl-Het model this is not the case by virtue of the neuron-to-neuron heterogeneity in cluster membership: the likelihood of a connection from a neuron i is highly dependent on the number of other neurons in the network that share a cluster with i (the quantity n_i defined before). Since this quantity varies significantly from neuron to neuron, connections from neuron i cannot be assumed to appear independently. In the particular case in which the clusters of neuron i are chosen independently with a fixed probability, this heterogeneity is captured by the number of clusters to which each neuron belongs, which can be considered the effective modulatory variable.

In the Deg model, the connection probability from neuron i to neuron j once the degrees are known can be approximated by Eq. (1.6.8). Additional connections from neuron i can be assumed to be made independently as long as $k \gg 1$. This independence assumption can be extended up to a group of n neurons as long as the degrees are large compared to n and $n \ll N$. Then, the Deg model becomes a special case of the Modulator model in which $x_i = (x_i^{\text{in}}, x_i^{\text{out}})$ is the 2-dimensional vector of the degrees of i and $g(x, y) = g_1(x)g_2(y)$, where $g_1(a, b) = \frac{b}{\sqrt{N\langle K\rangle}}$, $g_2(a, b) = \frac{a}{\sqrt{N\langle K\rangle}}$.

In/out-degree correlation in small samples

Given a random sample of a network, we define the sample degree correlation (SDC) as the Pearson correlation coefficient between in- and out-degrees of individual neurons in the sample:

$$\text{SDC} = \frac{\text{Cov}(k_i^{\text{in}}, k_i^{\text{out}})}{\sqrt{\text{Var}(k_i^{\text{in}})\text{Var}(k_i^{\text{out}})}}, \tag{1.6.15}$$

where i represents a random neuron and k_i^{in}, k_i^{out} are the in- and out-degrees of i in the sample (clearly they depend on the sample size; we use lower case letters to distinguish them from the network degrees).

In order to compute the SDC in our models we first need to introduce the following statistics.

Given any network and random nodes i, j, k, we define

$$\begin{aligned} p &:= P(i \to j) \\ R &:= P(i \leftrightarrow j)/p^2 \\ Conv &:= P(j \to i, k \to i)/p^2 \\ Div &:= P(i \to j, i \to k)/p^2 \\ Chain &:= P(j \to i, i \to k)/p^2. \end{aligned} \tag{1.6.16}$$

Note that these quantities do not trivially coincide with the motifs first defined in [Song et al., 2005] and reproduced here in Fig. 1.2 A. For example, the occurrence of the convergent motif number 5 above chance in Fig. 1.2 A can be written as

$$3\,\frac{P(j \to i,\ k \to i,\ \text{no other connections})}{\left(\frac{p_{\text{uni}}}{2}\right)^2 (1 - p_{\text{uni}} - p_{\text{bid}})}, \tag{1.6.17}$$

where $p_{\text{uni}} = 2p(1 - pR)$, $p_{\text{bid}} = p^2R$ and the factor 3 accounts for the different permutations of i, j and k which produce the same topological configuration. The motifs needed to compute the SDC are not conditioned on the presence or absence of any additional structure in the neuron triplet, merely the existence of, for example, a convergent motif. Therefore, $Conv$ is actually a weighted sum of the counts of all motifs in Fig. 1.2 A containing at least one convergent node, i.e. 5, 7, 9-10, 12-16.

The in- and out-degrees of a node i in a sample of size n can be expressed as

$$k_i^{\text{in}} = \sum_{j \neq i} X_{ij}, \qquad k_i^{\text{out}} = \sum_{j \neq i} X_{ji}, \tag{1.6.18}$$

where $X_{ij} = 1$ whenever $j \to i$ and $X_{ij} = 0$ otherwise (the sums in (1.6.18) are over the n indices of the neurons in the sample). Explicitly computing the sample degree variances and the covariance between in- and out-degrees of neuron i from expression (1.6.18) we find

$$\begin{aligned} \text{Var}(k_i^{\text{in}}) &= (n-1)p[(n-2)p \cdot Conv + 1 - (n-1)p] \\ \text{Var}(k_i^{\text{out}}) &= (n-1)p[(n-2)p \cdot Div + 1 - (n-1)p] \\ \text{Cov}(k_i^{\text{in}}, k_i^{\text{out}}) &= (n-1)p[(n-2)p \cdot Chain + pR - (n-1)p]. \end{aligned} \tag{1.6.19}$$

In the ER-Bi model, the pair to pair independence implies that $Conv = Div = Chain = 1$ and

$$\text{SDC} = \tfrac{p}{1-p}(R-1). \tag{1.6.20}$$

In the Modulator model, the quantities p, R, $Conv$, Div, $Chain$ can be rewritten in terms of moments of g:

$$\begin{aligned} p &= \langle g(x,y)\rangle \\ R &= \langle g(x,y)g(y,x)\rangle/p^2 \\ Conv &= \langle g(x,y)g(z,y)\rangle/p^2 \\ Div &= \langle g(x,y)g(x,z)\rangle/p^2 \\ Chain &= \langle g(x,y)g(y,z)\rangle/p^2, \end{aligned} \tag{1.6.21}$$

where $\langle\rangle$ indicates an average over the distribution of x, y, z, which are independent and identically distributed random variables. We have the following particular cases:

(i) If $g(x,y)$ is independent of $g(x,z)$, $g(z,x)$, $g(z,y)$ and $g(y,z)$, then $Conv = Div = Chain = 1$ and

$$\text{SDC} = \tfrac{p}{1-p}(R-1). \tag{1.6.22}$$

In the Cl and Dis models, the property of being in the same cluster (Cl) and the distance between a pair (Dis) can be assumed to be independent from one pair to another when N is large, so (1.6.22) is a good approximation of the sample degree correlation as long as $n \ll N$.

(ii) If g is symmetric, that is, $g(x,y) = g(y,x)$, then $Conv = Div = Chain$ and

$$\text{SDC} = \tfrac{p}{1-p}(R-1) + \tfrac{1-pR}{1-p}\left(1 - \tfrac{(n-1)p(1-p)}{\sqrt{\text{Var}(k_i^{\text{in}})\text{Var}(k_i^{\text{out}})}}\right). \tag{1.6.23}$$

This is the case of the Cl-Het model. Note that in the Cl/Dis models g is also symmetric, so this expression for SDC is a generalization of (1.6.22), which is recovered whenever $\sqrt{\text{Var}(k_i^{\text{in}})\text{Var}(k_i^{\text{out}})} = (n-1)p(1-p)$.

(iii) If g is multiplicative, that is, $g(x,y) = g_1(x)g_2(y)$, then $Chain^2 = R$ and

$$\text{SDC} = (n-1)\tfrac{p^2(n+\sqrt{R}-1)\left(\sqrt{R}-1\right)}{\sqrt{\text{Var}(k_i^{\text{in}})\text{Var}(k_i^{\text{out}})}}. \tag{1.6.24}$$

The Degree model fits within this case.

Notice that since the SDC can be explicitly calculated from p, R, $Conv$, Div and $Chain$, network models that have the same p, R, $Conv$, Div and $Chain$ but differ in higher-order statistics cannot be distinguished by means of the SDC.

From (1.6.19) it is clear that SDC = 0 for all n if, and only if, $R = Chain = 1$. In fact, $R = 1$ implies $Chain = 1$ in all the model categories studied in this thesis. Cases ER-Bi, (i) and (iii) are obvious. In the second case, that is, when g is symmetric, $R = 1$ implies

$$\langle g(x,y)^2 \rangle \;=\; \langle g(x,y)g(y,x) \rangle \;=\; \langle g(x,y) \rangle^2, \tag{1.6.25}$$

so

$$\langle (g(x,y) - \langle g(x,y) \rangle)^2 \rangle \;=\; 0, \tag{1.6.26}$$

which indicates that $g(x,y)$ is constant (except, maybe, in a zero measure set). This means that the model reduces to ER and, in particular, $Chain = 1$. Therefore, in the model categories defined here, we find that the conditions SDC = 0 for all n and $R = 1$ are equivalent.

1.6.4 Implementation of the networks used in simulations

Generation of distance-dependent networks

In the simulations of Figs. 1.1 to 1.6 A we considered neurons arranged in periodic rings where $r \in \{0, 1, \cdots, [N/2]\}$ and

$$p(r) \;=\; 1 - \tfrac{1}{1+e^{2s(r-t)}}, \tag{1.6.27}$$

which defines a decreasing sigmoid function whose absolute slope is maximal at $r = t$ and its value is $-s$. In the simulations of Fig. 1.6 C, D we also included two-dimensional periodic lattices where $r \in \{0, 1, \cdots, [\sqrt{N}/2]\}$ and $p(r)$ was given by Eq. (1.6.27).

Generation of networks from a prescribed in/out-degree distribution

To generate networks according to the Deg model we have used the following method: given a joint distribution defined by $\tilde{f}_{(\text{in, out})}$, we independently assign to each node i a pair $(\tilde{K}_i^{\text{in}}, \tilde{K}_i^{\text{out}})$. Then we create each connection $i \to j$ independently with probability $\frac{\tilde{K}_j^{\text{in}} \tilde{K}_i^{\text{out}}}{N\langle \tilde{K} \rangle}$. The final degrees in the network satisfy $\langle K_i^{\text{in}} | \tilde{K}_i^{\text{in}} \rangle = \tilde{K}_i^{\text{in}}$ and $\langle K_i^{\text{out}} | \tilde{K}_i^{\text{out}} \rangle = \tilde{K}_i^{\text{out}}$. Despite the resulting degree distribution in the network is no longer given by $\tilde{f}_{(\text{in, out})}$, the statistics $\langle K \rangle$ and $\text{Cov}(K^{\text{in}}, K^{\text{out}})$ are preserved (assuming that N is large and $\tilde{K}^{\text{in/out}} \ll N$). The degree variances become larger, in particular $\sigma^2_{\text{in/out}} = \tilde{\sigma}^2_{\text{in/out}} + \langle \tilde{K} \rangle$, and this results in the correlation coefficient being smaller, $\rho < \tilde{\rho}$.

In all our simulations, the variables $\tilde{K}^{\text{in}}$, $\tilde{K}^{\text{out}}$ followed Gamma distributions with a shift of magnitude $D > 0$. In almost all our simulations they had to be positively correlated and we defined them in the following way: if $X \sim \text{Gamma}(\kappa_1, \theta)$ and $Y, Z \sim \text{Gamma}(\kappa_2, \theta)$ $(\kappa, \theta > 0)$ are independent random variables, we set

$$\begin{aligned} \tilde{K}^{\text{in}} &= D + X + Y \\ \tilde{K}^{\text{out}} &= D + X + Z. \end{aligned} \tag{1.6.28}$$

$\tilde{K}^{\text{in}}$ and $\tilde{K}^{\text{out}}$ follow D-shifted Gamma$(\kappa = \kappa_1 + \kappa_2, \theta)$ distributions and their correlation coefficient is $\tilde{\rho} = \kappa_1/\kappa$. In Fig. 1.2 B we also constructed networks with negative degree correlation. In this case we first generated $\tilde{K}^{\text{in}}$ and $\tilde{K}^{\text{out}}$ independently and then we inversely ordered the two sequences $\{\tilde{K}_i^{\text{in}}\}_{i=1}^N$ and $\{\tilde{K}_i^{\text{out}}\}_{i=1}^N$. By reordering a fraction of values in one of the two sequences we could adjust the correlation coefficient.

Parameter values for Modulator networks shown in Figure 1.9

For all three networks $N = 100$, $p = 0.3$. In the Dis model the modulatory variable represents spatial position and g is a function of the distance $|x - y|$. The ordering of neurons in the adjacency matrix corresponds to their position in a ring. In the Cl-Het model, x and y represent the number of clusters to which pre- and post-synaptic neurons belong and g is a symmetric function. In particular, $g(x, y) = p_+(1 - f_-(x, y)) + p_- f_-(x, y)$, where $f_-(x, y)$ is the probability that the two neurons do not coincide in a cluster given x, y. Explicitly, $f_-(x, y) = \frac{(C-x)!(C-y)!}{C!(C-x-y)!}$ if $x + y \leq C$ and 0 otherwise (C is the total number of clusters in the network). Neurons in the adjacency matrix have also been ordered according to the number of clusters to which they belong. In this example, $C = 5$ and each neuron was assigned to each cluster with a fixed probability, so the fraction of neurons that belong to $k \in \{0, 1, \cdots, C\}$ clusters is not uniform. This is why the width of the different domains of the adjacency matrix and the g plot do not coincide. In the Deg model, (x_1, x_2) represents the pair of in- and out-degrees of the pre-synaptic neuron and (y_1, y_2) are the degrees of the post-synaptic neuron. $g(x_1, x_2, y_1, y_2) = cx_2y_1$, so g is separable with respect to the pre- and post-synaptic variables. We show different projections of g: $g(x_1, -, -, y_2)$ (top left), $g(-, x_2, -, y_2)$ (top right), $g(-, x_2, y_1, -)$ (bottom left) and $g(x_1, -, y_1, -)$ (bottom right). The adjacency matrices result from ordering neurons according to their out-degree (top) and their in-degree (bottom).

	μ_1	μ_2	σ_1^2	σ_2^2	ρ
f_1	0.0	0.0	0.3	0.3	0.92
f_2	-0.5	0.5	0.07	0.07	-0.62

Table 1.2: Parameters of the modulatory function

Definition of the model that fits the data

In the proposed model to fit the data of [Perin et al., 2011], connections are created independently with probability $P\left(i \to j \,|\, r_{ij} = r,\, x_i = x,\, x_j = y\right) = p(r)\, g(x, y)$. The distance dependency has the form

$$p(r) = a + br + cr^2, \tag{1.6.29}$$

where r is the normalized distance $r = \frac{d - d_{\min}}{d_{\max} - d_{\min}} \in [0, 1]$ that is computed from the real distance d in μm and minimal and maximal distances derived from the data, $d_{\min} = 10\mu$m, $d_{\max} = 350\mu$m. We took $a = 1$, $b = -1.04$, $c = 0.21$. The modulatory part is

$$g(x, y) = f_1(x, y) + f_2(x, y), \tag{1.6.30}$$

where f_1 and f_2 have the form

$$f(x, y) \quad = \quad \exp\left(-\frac{\sigma_2^2(x-\mu_1)^2 + \sigma_1^2(y-\mu_2)^2 - 2\rho\sqrt{\sigma_1^2\sigma_2^2}(x-\mu_1)(y-\mu_2)}{\sigma_1^2\sigma_2^2(1-\rho^2)}\right) \tag{1.6.31}$$

and their parameters are shown in Table 1.2. The modulators $\{x_i\}_i$ are independent from neuron to neuron and are drawn from a Gaussian distribution with mean 0 and standard deviation 0.5.

To obtain a distribution of distances in the simulated data close to the sampled distances in the experiment, we directly generated samples as in the real experiment. In each sample, the first neuron was located in the origin of coordinates and the others were sequentially located on the same plane at a position obtained by drawing a random angle $\alpha \in [0, 2\pi)$ and a radius r from a Gamma(κ, θ) distribution, $\kappa = 3.26$, $\theta = 0.08$. The radius was then rescaled as $d = d_0 + (d_1 - d_0) * r$, $d_0 = 16\mu$m, $d_1 = 250\mu$m. We avoided having neurons too close in space by checking, at every step, if the last neuron was closer than a limit distance $d_{\lim} = 14\mu$m to the already created neurons in the sample. In this case we chose a new position.

1.6.5 Statistical analysis

Implementation of the SDC criterion on a random network generator

In Fig. 1.6 C, D we applied the SDC criterion on networks generated randomly according to the models ER-Bi, Cl/Dis, Cl-Het and Deg. We chose a network class and values for $p \in [0.05, 0.23]$ and $R \in [1.5, 4.1]$ uniformly at random. In the ER-Bi model these parameters determine p_{uni} and p_{bid}. If the chosen class was Cl/Dis, we chose one of these two models with equal probability. In the Cl case, we selected the number of clusters randomly and then computed p_+ and p_- to get the desired p and R. In the Dis case, we chose a dimension (1 or 2) randomly and then placed neurons in periodic lattices of the given dimension. Then we determined the parameters s and t of Eq. (1.6.27) to fit p and R. If the selected model was Cl-Het we did exactly the same as in the Cl case. Finally, in the Deg model we chose D and $\rho > 0$ randomly and then found θ, κ_1 and κ_2 to fit p and R.

To classify a network according to the SDC, we took m random samples of size $n' = 12$ each. From them we estimated p, R, $Conv$, Div and $Chain$ (Eq. (1.6.16)) and computed the connection probability as a function of the number of common neighbors. From p, R, $Conv$, Div and $Chain$ we predicted $\sigma^2 = \sqrt{\text{Var}(k_i^{\text{in}})\text{Var}(k_i^{\text{out}})}$ and $\text{Cov}(k_i^{\text{in}}, k_i^{\text{out}})$ for any sample size n through Eq. (1.6.19). We compared the resulting SDC (seen as a function of n) with the SDC that would result in each of the model classes given the observed p, R and σ^2 (Eqs. (1.6.22), (1.6.23) and (1.6.24)). We determined which of these relationships between SDC and n better described the results by computing the sum of the squared distances between the actual SDC and the model predictions while varying n. The range of n values used to make this comparison is arbitrary. We chose $n \in \{3, \cdots, 12\}$ but the results are essentially the same for other choices. Since the formula for the Cl-Het model generalizes the formula for ER-Bi/Cl/Dis, the SDC of a network of the class ER-Bi/Cl/Dis will be fitted equally well by these two formulas. Thus, whenever the best fit corresponded to the Cl-Het class, we further studied if the SDC increased significantly with n by computing the slope of its linear regression and deciding if it was larger than a critical value s^*, which had been previously determined by means of simulations. If the slope was smaller than s^*, the network was reclassified as ER-Bi/Cl/Dis. Finally, to distinguish between ER-Bi and Cl/Dis networks, we determined if the connection probability in the n' samples increased significantly with the number of common neighbors. Again, this was done by computing a linear regression and comparing the slope with a previously defined threshold.

We further checked that the same algorithm works if σ^2 and $\mathrm{Cov}(k_i^{\mathrm{in}}, k_i^{\mathrm{out}})$ are calculated directly for each n on n-neuron samples instead of being estimated from p, R, *Conv*, *Div* and *Chain*. The n-neuron samples in this case are subsamples of the original samples of size n'. The only limitation of this procedure is that the original sample size n' has to be large enough to make it possible to compute σ^2 for n in the desired range, whereas the estimation of p, R, *Conv*, *Div* and *Chain* only requires 3-neuron samples. A study based on sampling from triplets or quadruplets, however, would not allow us to distinguish between the ER-Bi and Cl/Dis classes using the common neighbor rule.

Implementation of the SDC criterion on data

To apply the SDC criterion on the experimental data from [Perin et al., 2011] we considered all the possible subsamples of the original samples. For each subsample size, we used in- and out-degrees of all the neurons to compute $\sigma^2 = \sqrt{\mathrm{Var}(k_i^{\mathrm{in}})\mathrm{Var}(k_i^{\mathrm{out}})}$ and $\mathrm{Cov}(k_i^{\mathrm{in}}, k_i^{\mathrm{out}})$. Since the expected SDC's for each model class are functions of p, R, σ^2, which in a real situation are estimated quantities, they are prone to estimation errors, as well as the real SDC. We estimated the data SDC, the predicted SDC for the model classes and their standard errors by means of the Bootstrap method with 1000 re-samplings, as detailed below. On the one hand we created 1000 artificial samples with replacement from the set of in/out-degrees for each sample size. From each of these samples we computed σ^2. The mean of this collection of values gives the estimated σ^2, and the standard deviation, a measure of the standard error (SE). The same is done to estimate the real SDC and its SE. On the other hand, we estimated the mean and the SE for p, R and the different functions of p, R that participate in Eqs. (1.6.22), (1.6.23) and (1.6.24) in a similar way (in this case, by re-sampling over the different neuronal pairs in the network). For formulas that involve both σ^2 and p, R, i.e., Eqs. (1.6.23), (1.6.24), we computed upper bounds of the resulting errors from the previous partial errors.

We repeated the same procedure considering the predicted σ^2 and $\mathrm{Cov}(k_i^{\mathrm{in}}, k_i^{\mathrm{out}})$ from p, R, *Conv*, *Div* and *Chain*, where these statistics were computed using all the information in the original samples (i.e., using all the pairs and triplets). The results are almost identical.

It is important to notice that this exhaustive data analysis might introduce biases in the estimation of *Conv*, *Div*, *Chain*, σ^2 and the SDC because the triplets and the nodes involved in computing in/out-degrees partially overlap. To cope with this, we used exactly the same procedure in all the analyses of Figs. 1.6, 1.8, 1.10. The fact that the classification algorithm is pretty accurate

even when the number of studied samples is small (Fig. 1.6 C) indicates that such correlations do not play a very important role. In spite of this, we asked ourselves if the deviation in σ^2 from the Cl/Dis model seen in the data could be due to these effects and not to the fact that the real underlying structure deviates from this simple model. To investigate this issue we simulated networks with the same distance-dependent component exhibited by the data with an additional modulatory component based on clustering. The repetition of many replicas of the real experiment on this model indicated that the observed deviation of σ^2 is statistically significant (p-value < 0.05, Fig. 1.11). This suggests that the discrepancy from a symmetric modulatory model is not due to sparse sampling or correlations derived from data overlaps.

The analysis of the data from [Song et al., 2005] was done by directly computing *Conv*, *Div* and *Chain* from the statistics of 3-neuron motifs shown in the paper.

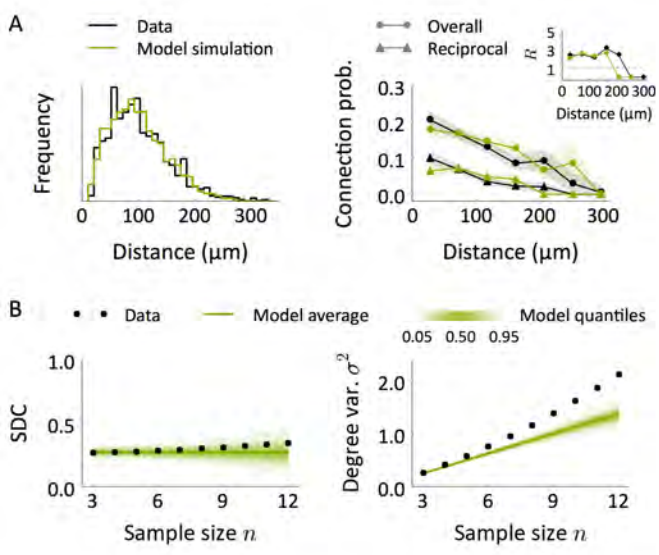

Figure 1.11: Comparison between the data (black) and a null model which incorporates the observed distance-dependency of the data and an additional clustered structure (green). All the plots are equivalent to those shown in Fig. 1.10 B, C except for the change in the model (see the caption of Fig. 1.10 for details).

Chapter 2

A mean-field description of stationary firing rates in networks of spiking neurons with arbitrary degree distributions

2.1 Introduction

One of the aims of neural science is to understand how the observed patterns of neuronal activity originate from the properties of single neurons and the interactions between them. Many experimental studies have revealed that cortical neurons tend to spike at low rates and in a highly irregular manner. The irregularity of the spiking process of a single neuron can be evaluated through the so-called *coefficient of variation* (CV), which is the ratio between the (temporal) standard deviation of the inter-spike intervals (ISIs) and their mean. Electrophysiological data have shown that the CVs of cortical neurons tend to be in the range 0.5-1 [Softky and Koch, 1993], which is a signature of a highly irregular spiking (note that the paradigm of an irregular process, the Poisson process, has CV= 1). One of the questions that arise from these observations is: what are the mechanisms responsible for such irregularity in real networks?

The current hypothesis is that, in physiological conditions, neurons receive a huge number of both excitatory and inhibitory inputs which cancel in the mean, in such a way that the membrane voltage resembles a random walk between the resting and the threshold potentials. Although the

number of total inputs received can be large, the spiking process, that results from the voltage crossing the threshold potential, becomes highly irregular [Shadlen and Newsome, 1994]. Such a *balanced state* has been successfully reproduced in models of neuronal networks without the need for fine-tuning of parameters [van Vreeswijk and Sompolinsky, 1996; Renart et al., 2010]. A precise, macroscopic description of networks of leaky integrate-and-fire (LIF) neurons in the balanced state is provided by mean-field theory, which allows one to predict the firing rates and CVs in the stationary state (that is, the state in which macroscopic quantities do not vary over time) [Brunel, 2000]. This theoretical framework can explain other observed features of neuronal dynamics, as the skewed rate distributions found *in vivo* [Amit and Brunel, 1997a; Roxin et al., 2011].

In the majority of such studies, however, network structure is either homogeneous (that is, every neuron receives the same amount of connections from the network) or random (i.e., connections among neurons are created independently with a fixed probability). In Chapter 1 we have seen that cortical microcircuitry deviates from these overly simple scenarios. It remains an open question to what extent the dynamics exhibited by the cited models is affected by more realistic topologies.

One of the characteristic properties of plain random networks is that they exhibit little structural heterogeneity: the distribution of in- and out-degrees in the network is tightly peaked around the mean value. In Chapter 1 we showed that some of the nonrandom features of real cortical circuits are compatible with networks that are defined through broad in/out-degree distributions which are also positively correlated, although such configurations are unlikely under the light of other local measures. In any case, the analysis of the structure in small groups of neurons revealed that in- and out-degrees are positively correlated, a feature that simple random models cannot reproduce [Vegué et al., 2017]. Therefore, it is of particular interest to study the role that broad degree distributions and degree correlations might play in neuronal dynamics. The effect of broadening in- and out-degree distributions in networks of spiking neurons has been studied by Roxin [2011], who showed that the variance of in-degrees has an important impact in the ability of the network to exhibit global oscillations, whereas the out-degree variance shapes pairwise correlations in the synaptic currents. Broad excitatory distributions of in-degrees can also break down the balanced assumption unless proper compensatory mechanisms are introduced (such as correlations between the number of excitatory and inhibitory connections onto individual neurons or tuning of the inhibitory weights) [Landau et al., 2016]. Nykamp et al. [2017] have analyzed how in/out-degree correlations affect dynamics in models based on firing rates and have shown that a

positive correlation between excitatory in- and out-degrees has a similar effect to increasing the excitatory coupling in the network.

In this chapter we address the problem of extending the mean-field techniques introduced by Amit and Brunel [Amit and Brunel, 1997b; Brunel and Hakim, 1999; Brunel, 2000] to networks of excitatory and inhibitory LIF neurons with a highly heterogeneous structure. In particular, we consider the case in which in- and out-degrees follow a prescribed joint distribution. The extended system of self-consistent equations provides a means to compute the distribution of firing rates and CVs in the stationary state. Our results show good agreement between theory and simulations. We use the derived equations to demonstrate that a positive correlation between in- and out-degrees can have important consequences on dynamics, mainly because it biases the firing rate distribution in the set of available pre-synaptic neurons. This effect has been already pointed out by Nykamp et al. [2017] in the case of firing rate models.

The presence of broad degree distributions can destabilize dynamics and make the stationary state unstable, as long as inhibitory connections are created with a fixed probability. This is due to the fact that in such networks there are neurons which receive a large amount of excitation that is not balanced with inhibition. Heterogeneity in the total amount of excitation received has been observed in the rodent cortex, where compensatory mechanisms at the level of inhibitory synapses are able to maintain a proper balance [Xue et al., 2014]. We have mimicked such a compensation by allowing inhibitory connections to appear with a probability that is modulated by the total excitatory in-degree. Under such circumstances, the network can return to an asynchronous stationary state, very similar to those exhibited by purely random networks. Interestingly, in networks whose degrees are, in addition, positively correlated, transient external inputs can destabilize the stationary state for a period larger than the duration of the stimulus. This finding suggests a possible role of the degree correlation in enhancing the ability of neuronal networks to respond to transient stimulation.

2.2 Background

2.2.1 The model

Let us consider a network of N spiking neurons. An arbitrary neuron i in the network has a membrane voltage V_i which evolves according to

$$\tau \frac{\mathrm{d}V_i(t)}{\mathrm{d}t} = -V_i(t) + \tau I_i(t), \tag{2.2.1}$$

where τ is a time constant and $I_i(t)$ is the synaptic input. Every time V_i reaches a threshold V_θ, the neuron generates an action potential and the voltage is immediately reset to V_r, where it remains for a resting period τ_r.

The input is generated from the spikes of the pre-synaptic neurons to our neuron. We impose that every action potential emitted by the j-th pre-synaptic neuron induces an instantaneous jump in the voltage V_i, of magnitude J_{ij}. This is equivalent to saying that the instantaneous variation of V_i at time t induced by a pre-synaptic spike emitted by neuron j at time t' is $J_{ij}\,\delta(t-t')$, where δ denotes the Dirac delta function. Thus, the external input can be expressed as

$$I_i(t) = \sum_{j=1}^{K_i} J_{ij} \sum_k \delta(t - t_j^k), \tag{2.2.2}$$

where K_i is the number of incoming connections or in-degree of the neuron under study and $\{t_j^1, t_j^2, \ldots\}$ are the spike times of the pre-synaptic neuron j.

We are interested in studying macroscopic properties of this system when we impose a certain topology in the network. The first assumption to make is that individual neurons fire as Poisson processes, so that Eq. (2.2.2) is non-deterministic and Eq. (2.2.1) becomes a *stochastic* differential equation. The original source of this stochasticity will be a set of external neurons which we impose to fire at random times, according to Poisson processes. But even in the absence of such stochastic inputs, if the resulting spiking times have temporal statistics similar to that of Poisson processes, the system can be treated in the same way.

2.2.2 Mean-field equations

As detailed in Appendix B, mean-field analysis provides tools for predicting some statistical properties of the stationary state for networks of this type when the underlying structure is

homogeneous [Brunel, 2000] (a network state is said to be *stationary* if its macroscopic dynamical properties are constant in time). Some of such macroscopic, or statistical, variables are the firing rates of the neurons in the network and the CVs of their ISIs. The firing rate in this context represents the average number of spikes emitted per neuron and per unit time within the set of neurons of the same characteristics (we assume that N is large, so the network has many equivalent neurons). It can also be seen as the average number of spikes per neuron and unit time across many different realizations of the stochastic dynamics. In the stationary state, the firing rate can be evaluated as a temporal average as well.

The results are different depending on the topology of the network: if all the neurons are equivalent and have the same in-degree, the stationary state is characterized by a single stationary firing rate, whereas a variability in terms of in-degrees translates into a variability of stationary firing rates. In the latter case the stationary state is described by a *distribution* of firing rates.

The necessary conditions for the analysis to be correct are:

(i) neurons fire as independent Poisson processes;

(ii) the sizes of the voltage jumps $\{J_{ij}\}_{i,j}$ are small compared with the threshold V_θ so that the voltage can be approximated by a continuous variable.

Condition (i) is approximately fulfilled when the synaptic input is sub-threshold (which induces irregular spiking) and there is a small overlap in the total input received by any pair of neurons (which ensures independence between inputs to different neurons). A small overlap occurs when the connectivity is random and sparse, that is, when the in-degrees are small compared with the system's size N. Condition (ii) depends on the parameter's choice and can therefore be assumed in general. From now on, we will suppose that conditions (i) and (ii) are fulfilled.

Under these assumptions, the stochastic evolution of the membrane voltage V of a single neuron can be described by means of a Fokker-Planck equation, which defines a partial differential equation for the probability density of the voltage as a function of time, $\rho(V,t)$ (see Eqs. (B.1.16), (B.1.15) in Appendix B). The search of a solution which is constant in time (imposing appropriate boundary conditions related to the thresholding mechanism) results in the following expression

for the stationary firing rate ν:

$$\begin{aligned} \nu &= \phi(\mu,\sigma) \\ \phi(\mu,\sigma) &:= \left(\tau_r + \tau\sqrt{\pi}\int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} \exp(u^2)\,\mathrm{erfc}(-u)\,\mathrm{d}u\right)^{-1}, \end{aligned} \tag{2.2.3}$$

where μ and σ^2 are the mean and the variance of the total input received during a time window of length τ (see Appendix B for details).

Totally homogeneous connectivity

When the network is composed of excitatory (E) and inhibitory (I) neurons with the same connectivity and dynamical properties and the E (I) weights are J_E ($-J_I$), all the neurons are statistically equivalent and their firing rate in the stationary state ν is the same. The expression *same connectivity properties* means that all the neurons receive the same number K_E of E connections and the same number K_I of I connections but the precise realization of this connectivity is totally random and therefore uncorrelated from neuron to neuron. We can also assume that each neuron receives external inputs from an independent set of K_{ext} neurons, through synaptic weights J_{ext}, which fire at a constant rate ν_{ext}. As shown in Appendix B, in this scenario μ and σ^2 depend on ν through

$$\begin{aligned} \mu &= \left[(K_E J_E - K_I J_I)\,\nu + K_{\mathrm{ext}} J_{\mathrm{ext}} \nu_{\mathrm{ext}}\right]\tau \\ \sigma^2 &= \left[(K_E J_E^2 + K_I J_I^2)\,\nu + K_{\mathrm{ext}} J_{\mathrm{ext}}^2 \nu_{\mathrm{ext}}\right]\tau, \end{aligned} \tag{2.2.4}$$

so (2.2.3) and (2.2.4) define a self-consistent equation for the stationary firing rate ν.

Erdös-Rényi (ER) connectivity

When connections are generated independently with a fixed probability, neurons are heterogeneous in terms of their in-degrees, and this induces a heterogeneity in the stationary firing rates. The mean input received within a time window of length τ can then be parametrized by a varible $W \sim N(0,1)$: $\mu = \mu(W)$. W captures both the variability coming from the differences in in-degrees and the variability due to the fact that pre-synaptic neurons fire at different rates. In the limit of a large network, both levels of variability can be put together under this common continuous random variable W [Amit and Brunel, 1997a; Roxin et al., 2011].

The stationary firing rates are therefore also parametrized by W through $\nu(W) = \phi(\mu(W),\sigma)$, so

the stationary state is characterized by a *distribution* of firing rates. $\mu(W)$ and σ in turn depend on the mean $\bar{\nu}$ and the variance s^2 of this distribution, so the firing rate is also a function of these parameters: $\nu = \nu(W, \bar{\nu}, s^2) = \phi\left(\mu(W, \bar{\nu}, s^2), \sigma(\bar{\nu}, s^2)\right)$, where

$$\begin{aligned} \bar{\nu} &= \int_{-\infty}^{\infty} \nu(w, \bar{\nu}, s^2) f(w)\, \mathrm{d}w \\ s^2 &= \int_{-\infty}^{\infty} \nu(w, \bar{\nu}, s^2)^2 f(w)\, \mathrm{d}w - \bar{\nu}^2 \end{aligned} \tag{2.2.5}$$

and f is the probability density function of a standard Gaussian random variable. Eq. (2.2.5) constitutes a system of self-consistent equations for $\bar{\nu}$ and s^2, from which the entire firing rate distribution can be reconstructed. If, in addition, E and I neurons are different in terms of other properties, the equations are analogous but depend on the mean and variances of E and I firing rate distributions. In this case the system to be solved has four unknowns and four equations.

We do not provide here the details of how this parametrization can be defined because in the next section we will address this problem in a more general scenario, that reduces to the simply random (ER) topology when in- and out-degrees follow independent Binomial distributions.

2.3 Distribution of firing rates in networks with arbitrary degrees

The aim now is to extend the formalism sketched in the previous section to more general classes of networks. The analytical description of the stationary state provided by the mean-field theory has been so far applied either to networks with homogeneous in-degrees or to networks in which the in-degree heterogeneity is the one provided by ER models. Recall that in an ER(p) network, connections are drawn independently with probability p, so the mean degree is $\langle K \rangle = (N-1)p$ and its variance, $\Delta K^2 = (N-1)p(1-p) = (1-p)\langle K \rangle$. Therefore, the ratio between the standard deviation and the mean of the degree distribution is

$$\frac{\Delta K}{\langle K \rangle} = \frac{\sqrt{(1-p)\langle K \rangle}}{\langle K \rangle} \leq \frac{1}{\sqrt{\langle K \rangle}} \approx 0 \tag{2.3.1}$$

for large values of $\langle K \rangle$ (the exact limit of this ratio depends on how the parameters scale with N: if p is fixed, this clearly goes to zero as $N \to \infty$, whereas if p scales inversely with N, $\langle K \rangle$ is constant and the ratio tends to a constant which is small but not identically zero). We will focus now on networks whose connectivity (at least within the EE subnetwork) is generated to preserve a given in/out-degree distribution, as in the Degree model described in Chapter 1. The

distributions used will be broader than those provided by ER networks, that is, we will consider distributions for which $\langle K \rangle$ is large when N is large but whose ratio $\frac{\Delta K}{\langle K \rangle}$ stays constant and it is not necessarily small. We will also introduce correlations between individual in- and out-degrees and see how these correlations shape the distribution of firing rates in the stationary state.

2.3.1 Networks with arbitrary in-degree distribution

We start by considering the case of networks with arbitrary degree distribution, assuming that the out-degrees are independent of in-degrees. To make the presentation clearer, we assume first that the network is composed of neurons of a single type with homogeneous synaptic weights J.

Let us consider a single neuron. Its firing rate in the stationary state is $\nu = \phi(\mu, \sigma)$, where ϕ is defined by Eq. (2.2.3) and μ, σ are the mean and the standard deviation of the total input received by the neuron within a time window of length τ. We start by computing μ and σ^2.

Let us suppose that our neuron receives inputs from K pre-synaptic neurons. Since the pre-synaptic neurons fire as Poisson processes, if we knew their rates, $\nu_1, \cdots, \nu_K$, these quantities would be

$$\begin{aligned} \mu &= \tau J \sum_{i=1}^{K} \nu_i \\ \sigma^2 &= \tau J^2 \sum_{i=1}^{K} \nu_i. \end{aligned} \tag{2.3.2}$$

As different neurons have different degrees, we expect to have a *distribution* of firing rates in the stationary state in the entire network. Therefore, the term $S := \sum_{i=1}^{K} \nu_i$ varies across neurons due to differences in their connectivity. If the neuron under study is randomly chosen, S will be a random variable that is defined as a sum of independent variables that come from a common firing rate distribution. If K is large enough, the Central Limit Theorem ensures that S will approximately follow a Gaussian distribution with mean $K\bar{\nu}$ and variance Ks^2, where $\bar{\nu}$ and s^2 are the mean and variance of the stationary firing rate distribution:

$$\begin{aligned} \mu &= \tau J S \\ \sigma^2 &= \tau J^2 S \\ S &= K\bar{\nu} + \sqrt{Ks^2}\, W, \qquad W \sim N(0,1). \end{aligned} \tag{2.3.3}$$

Recall that one of our assumptions is that the network is sparse, meaning that $K \ll N$. But we also suppose that the degrees are large quantities when N is large. In this case, we can approximate σ by its leading term $\sqrt{\tau J^2 K \bar{\nu}}$.

Regarding now K as a variable which varies from neuron to neuron, we can express μ and σ as a function of the *quenched* randomness given by the pair (K, W) as

$$\begin{aligned} \mu(K,W) &= \tau J S(K,W) \\ \sigma^2(K) &= \tau J^2 K\bar{\nu} \\ S(K,W) &= K\bar{\nu} + \sqrt{Ks^2}\, W, \end{aligned} \tag{2.3.4}$$

and the stationary firing rate of a neuron with $K = k, W = w$ given the statistics $\bar{\nu}$ and s^2 of the firing rate distribution is

$$\nu(k, w, \bar{\nu}, s^2) = \phi(\mu(k, w), \sigma(k)). \tag{2.3.5}$$

Again, in the large N limit we can treat the degrees as continuous random variables. In this case, as in (2.2.5), the system of equations can be closed through the definition of $\bar{\nu}$ and s^2:

$$\begin{aligned} \bar{\nu} &= \int_0^\infty \int_{-\infty}^\infty \nu(k,w,\bar{\nu},s^2) f_K(k) f_W(w)\, \mathrm{d}w\, \mathrm{d}k \\ s^2 &= \int_0^\infty \int_{-\infty}^\infty \nu(k,w,\bar{\nu},s^2)^2 f_K(k) f_W(w)\, \mathrm{d}w\, \mathrm{d}k - \bar{\nu}^2, \end{aligned} \tag{2.3.6}$$

where f_K and f_W are the probability density functions of the K and W variables. Since W is a standard Gaussian, $f_W(w) = \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2}}$. Notice that the integration limit of the variable K is set to $(0, \infty)$ even though the degree is always bounded by $(N-1)$; we assume that the support of f_K is contained in $[0, N-1]$. To find the distribution of firing rates in this system we have to (numerically) solve Eq. (2.3.6) for the unknowns $(\bar{\nu}, s^2)$. The stationary firing rate of a neuron with $K = k$ and $W = w$ is then just obtained by evaluating Eqs. (2.3.4), (2.3.5).

This formulation can be easily extended to the case of a network composed of E and I neurons where the connectivity rules are different depending on the neuronal type and there is an external source of inputs. Imagine that the external inputs come from a population of neurons which fire as Poisson processes at constant rate ν_{ext}, in such a way that each neuron receives information from a fixed number K_{ext} of these external sources, which are totally uncorrelated. We assume that all the E (I) weights are the same and they take the value J_E ($-J_I$). If $K_{\alpha\beta}$ is the random variable which gives the number of inputs from population β to a neuron within population α,

expression (2.3.4) reads

$$\begin{aligned}
\mu_\alpha(K_{\alpha E}, K_{\alpha I}, W_{\alpha E}, W_{\alpha I}) &= \tau\left[J_E\, S_E(K_{\alpha E}, W_{\alpha E}) - J_I\, S_I(K_{\alpha I}, W_{\alpha I}) + J_{\text{ext}}\, K_{\text{ext}}\, \nu_{\text{ext}}\right] \\
\sigma^2_\alpha(K_{\alpha E}, K_{\alpha I}) &= \tau\left(J_E^2\, K_{\alpha E}\, \bar{\nu}_E + J_I^2\, K_{\alpha I}\, \bar{\nu}_I + J_{\text{ext}}^2\, K_{\text{ext}}\, \nu_{\text{ext}}\right) \\
S_\beta(K, W) &= K\bar{\nu}_\beta + \sqrt{K s^2_\beta}\, W,
\end{aligned} \tag{2.3.7}$$

where μ_α and σ^2_α denote the μ and σ^2 variables associated to population $\alpha \in \{E, I\}$, and $\bar{\nu}_\beta$ and s^2_β refer to the mean and variance of the distribution of stationary firing rates within population β.

Expressions (2.3.7) specify the magnitude of μ_α and σ^2_α as a function of the random variables $K_{\alpha E}$, $K_{\alpha I}$, $W_{\alpha E}$, $W_{\alpha I}$. $K_{\alpha E}$ and $K_{\alpha I}$ represent the heterogeneity in terms of in-degrees, whereas $W_{\alpha E}$ and $W_{\alpha I}$ are normally distributed and reflect the variability in the firing rates of the input neurons. All this variability is *quenched* in the sense that it represents the neuron-to-neuron heterogeneity but it is stable in time. In general situations $K_{\alpha E}$, $K_{\alpha I}$, $W_{\alpha E}$, $W_{\alpha I}$ will be pairwise independent, except, maybe, for the two "structural" variables $K_{\alpha E}$, $K_{\alpha I}$, which could be correlated depending on the connectivity imposed (consider, for example, a network where E and I degrees compensate each other to achieve balance). The expressions for μ_α and σ^2_α also depend on the mean and variance of the distributions of stationary firing rates in the E and I populations. Therefore, the firing rate of a neuron is a function of $K_{\alpha E}$, $K_{\alpha I}$, $W_{\alpha E}$, $W_{\alpha I}$, $\bar{\nu}_E$, s^2_E, $\bar{\nu}_I$, s^2_I and can be recovered when the mean and standard deviation of the firing rate distributions are known. These quantities can be computed by solving the corresponding extended version of Eq. (2.3.6).

2.3.2 Networks whose in- and out-degrees are correlated

In the previous section we have pointed out the dependency of the firing rate of a given neuron on its "characteristic" variables $K_{\alpha E}$, $K_{\alpha I}$, $W_{\alpha E}$, $W_{\alpha I}$. The result is general for any connectivity structure in which typical degrees are large but small compared with the system's size N. Now we want to introduce the effect of having a non-zero correlation between individual in- and out-degrees.

To simplify the arguments, let us consider again the case of a single neuronal population. The variable K is the in-degree of the neuron under study, whereas W represents the variability in the sum of the firing rates of its pre-synaptic neurons. The presented theory supposes that this sum originates from taking many independent realizations of the distribution of firing rates in

the network. Since the number of elements that contribute to the sum is large, it is sufficient to know what the mean and variance of this distribution are. There does not seem to be a role for the out-degree in this formulation.

However, the distribution of firing rates within the set of pre-synaptic neurons to a given neuron can deviate from the distribution of firing rates in the network if in- and out-degrees are correlated. We provide first an intuitive explanation of this fact and we will show it analytically afterwards.

Let us consider the process of taking one neuron at random and then picking one of its in-neighbors at random. Now we focus on the out-degree of the last neuron. This out-degree, as a random variable, does not follow the same distribution as the out-degrees in the network. For example, the network could potentially have many neurons with zero out-degree, but pre-synaptic neurons have out-degree of at least 1. Therefore, the distribution of out-degrees of the in-neighbors of a given neuron is biased. Now we consider the same random process but we look at in-degrees. The result is clear: if in/out-degrees are independent, the distribution will match the in-degree distribution in the network. On the contrary, if there is a correlation between degrees, the previous bias will be inherited by the in-degree distribution. Hence, the pre-synaptic neurons to a given neuron have in-degrees which follow a distribution that does not necessarily match the network distribution. Since the firing rate of a neuron is a function of its in-degree, this translates into a bias in terms of the *firing rates* of the pre-synaptic neighbors.

Let us show this rigorously and compute the bias. For clarity, we will treat the degrees as discrete variables. First, notice that if i and j are two random neurons and K_i^{in}, K_i^{out} K_j^{in}, K_j^{out} are their in- and out-degrees,

$$P\left(j \to i \,|\, K_i^{\text{in}} = x, K_i^{\text{out}} = y, K_j^{\text{in}} = x', K_j^{\text{out}} = y'\right) \approx \frac{xy'}{N\langle K\rangle} \tag{2.3.8}$$

(this relation was derived in Chapter 1, see p. 34; the only difference is that now we consider the probability conditioned on the two pairs of in/out-degrees, but the formula is the same because the in-degree of the pre-synaptic neuron and the out-degree of the post-synaptic neuron do not add any extra information). Now we compute the probability that the firing rate of a pre-synaptic

neuron j to neuron i lies within the range $(\nu, \nu+\delta)$ once we know the in-degree of i:

$$\begin{aligned} & P\left(\nu_j \in (\nu,\nu+\delta) \,|\, K_i^{\text{in}} = x,\, j \to i\right) \\ & \quad = \sum_{x'} P\left(\nu_j \in (\nu,\nu+\delta) \,|\, K_i^{\text{in}} = x, j \to i, K_j^{\text{in}} = x'\right)\, P\left(K_j^{\text{in}} = x' \,|\, K_i^{\text{in}} = x, j \to i\right) \\ & \quad = \sum_{x'} P\left(\nu_j \in (\nu,\nu+\delta) \,|\, K_j^{\text{in}} = x'\right)\, P\left(K_j^{\text{in}} = x' \,|\, K_i^{\text{in}} = x, j \to i\right). \end{aligned} \tag{2.3.9}$$

On the other hand,

$$\begin{aligned} & P\left(K_j^{\text{in}} = x' \,|\, K_i^{\text{in}} = x, j \to i\right) \\ & \quad = \sum_{y'} P\left(K_j^{\text{in}} = x', K_j^{\text{out}} = y' \,|\, K_i^{\text{in}} = x, j \to i\right) \\ & \quad = \sum_{y'} \frac{P\left(j \to i \,|\, K_j^{\text{in}} = x', K_j^{\text{out}} = y', K_i^{\text{in}} = x\right)\, P\left(K_j^{\text{in}} = x', K_j^{\text{out}} = y' \,|\, K_i^{\text{in}} = x\right)}{P\left(j \to i \,|\, K_i^{\text{in}} = x\right)} \\ & \quad = \sum_{y'} \frac{x y'\, P\left(K_j^{\text{out}} = y' \,|\, K_j^{\text{in}} = x'\right) P\left(K_j^{\text{in}} = x'\right)}{x\langle K\rangle} \\ & \quad = \frac{\langle K_j^{\text{out}} \,|\, K_j^{\text{in}} = x'\rangle}{\langle K\rangle}\, P\left(K_j^{\text{in}} = x'\right), \end{aligned} \tag{2.3.10}$$

where we have used the fact that, in the considered networks, the degrees of different neurons are independent variables. Inserting (2.3.10) into (2.3.9) gives

$$\begin{aligned} & P\left(\nu_j \in (\nu,\nu+\delta) \,|\, K_i^{\text{in}} = x, j \to i\right) \\ & \quad = \sum_{x'} P\left(\nu_j \in (\nu,\nu+\delta) \,|\, K_j^{\text{in}} = x'\right)\, \frac{\langle K_j^{\text{out}} \,|\, K_j^{\text{in}} = x'\rangle}{\langle K\rangle}\, P\left(K_j^{\text{in}} = x'\right). \end{aligned} \tag{2.3.11}$$

This is the fundamental result. It indicates that the distribution of firing rates among the pre-synaptic neurons of a given neuron does not necessarily follow the distribution of firing rates in the network. The bias is due to the fact that the expectation of the out-degree of a neuron *conditioned* on its in-degree is not necessarily equal to the expected out-degree. The ratio between this conditional expectation and the expectation itself is what alters the distribution of firing rates. Of course, when the degrees are independent this ratio is 1 and we recover the distribution of firing rates in the network. Notice, also, that the biased distribution is independent of the in-degree of the post-synaptic neuron, K_i^{in}, which means that the distribution of firing rates within the pre-synaptic neighbors of a given neuron is a *network* property (rather than a property associated to each post-synaptic neuron).

As a consequence, the sum of the rates of the pre-synaptic neurons to a given neuron (the variable

S defined in (2.3.3)) is, in fact, a sum over the *biased* distribution of firing rates. We will denote this new distribution with a star (*), so that $\bar{\nu}^*$ and s^* will refer to the mean and the standard deviation of the biased version of the firing rate distribution.

The mean-field equations are the same as before with the difference that we have to replace $\bar{\nu}$ and s^2 by $\bar{\nu}^*$ and $(s^*)^2$ in the definition of μ and σ^2 (Eq. (2.3.4)). Analogously, in Eq. (2.3.7) $\bar{\nu}_\alpha$ and s_α^2 are replaced by $\bar{\nu}_\alpha^*$ and $(s_\alpha^*)^2$, respectively. To self-consistently close the equations, we take into account the definition of the biased rate moments:

$$\begin{aligned} \bar{\nu}^* &= \int_0^\infty \int_{-\infty}^\infty \nu(k, w, \bar{\nu}^*, (s^*)^2) f_K^*(k) f_W(w)\, \mathrm{d}w\, \mathrm{d}k \\ (s^*)^2 &= \int_0^\infty \int_{-\infty}^\infty \nu(k, w, \bar{\nu}^*, (s^*)^2)^2 f_K^*(k) f_W(w)\, \mathrm{d}w\ \mathrm{d}k - (\bar{\nu}^*)^2, \end{aligned} \tag{2.3.12}$$

where f_K^* is the probability density function of the *biased* in-degree:

$$f_K^*(k) = \frac{\langle K^{\mathrm{out}} \,|\, K^{\mathrm{in}} = k \rangle}{\langle K \rangle} f_K(k). \tag{2.3.13}$$

Interpreting the degrees as continuous variables, the conditional expectation is computed as

$$\langle K^{\mathrm{out}} \,|\, K^{\mathrm{in}} = x \rangle = \int_0^\infty y\, f_{\mathrm{out|in}}(y \,|\, x)\, \mathrm{d}y = \frac{1}{f_K(x)} \int_0^\infty y\, f_{\mathrm{in,out}}(x, y)\, \mathrm{d}y, \tag{2.3.14}$$

where $f_{\mathrm{out|in}}$ is the probability density of the out-degree conditioned to the in-degree and $f_{\mathrm{in,out}}$ is the probability density of the joint degree distribution.

Analogous equations for the moments of the rate distributions can be obtained in the two-population scenario described by (2.3.7). We provide some examples in the next section.

2.3.3 Examples

E/I network with ER connectivity for E→I, I→E and I→I connections and arbitrary degree distribution within the EE subnetwork

We first analyze the case of a network composed of excitatory (E) and inhibitory (I) neurons where connections from/to I neurons are generated independently with probability p (i.e., they have ER-like structure). The connectivity within the EE subnetwork, on the contrary, is created

according to a joint in/out-degree distribution with correlation ρ. We assume that the number of external inputs that each neuron receives, K_{ext}, is constant across the entire population.

We denote by $\langle K\rangle_{\alpha\beta}$ and $\Delta K_{\alpha\beta}$ the mean and the standard deviation of the in-degree of a neuron in population α from population β. Since the ER in-degrees have a mean much larger than its standard deviation, the contributions of $\Delta K_{\alpha\beta}$ in expressions of the type $\sqrt{K_{\alpha\beta}}$ in Eq. (2.3.7) can be neglected for $(\alpha,\beta) \in \{(E,I),(I,E),(I,I)\}$ (and thus $\sqrt{K_{\alpha\beta}} \approx \sqrt{\langle K\rangle_{\alpha\beta}}$ in these cases). On the other hand, and representing again the degrees by continuous variables, the remaining random variables $K_{\alpha I}$, $W_{\alpha E}$ and $W_{\alpha I}$ of Eq. (2.3.7) are independent and follow approximately Normal distributions, so they can be grouped together under a single Gaussian random variable W_α (this holds from the fact that the ER degrees follow Binomial distributions, which can be approximated by Gaussians in the large N limit). In the I population, the K_{IE} variable is also independent of the others and Gaussian, so it can be grouped with the other three variables through a new Gaussian variable W_I.

This finally implies that the firing rates of neurons in the E population are parametrized by two quenched random variables (the excitatory in-degree K_{EE} and $W_E \sim N(0,1)$) and the firing rates in the I population, by a single variable ($W_I \sim N(0,1)$) which includes all the contributions coming from the degree and the rate heterogeneity.

We thus obtain

$$
\begin{aligned}
\nu_E(K_{EE}, W_E, \bar{\nu}_E^*, (s_E^*)^2, \bar{\nu}_I, s_I^2) &= \phi\left(\mu_E(K_{EE}, W_E), \sigma_E(K_{EE})\right) \\
\nu_I(W_I, \bar{\nu}_E, s_E^2, \bar{\nu}_I, s_I^2) &= \phi\left(\mu_I(W_I), \sigma_I\right),
\end{aligned}
$$

$$
\begin{aligned}
\mu_E(K_{EE}, W_E) &= \tau\left(J_E K_{EE}\,\bar{\nu}_E^* - J_I\langle K\rangle_{EI}\,\bar{\nu}_I + J_{\text{ext}} K_{\text{ext}}\,\nu_{\text{ext}}\right) + \Delta_E(K_{EE})\,W_E \\
\sigma_E^2(K_{EE}) &= \tau\left(J_E^2 K_{EE}\,\bar{\nu}_E^* + J_I^2\langle K\rangle_{EI}\,\bar{\nu}_I + J_{\text{ext}}^2 K_{\text{ext}}\,\nu_{\text{ext}}\right) \\
\Delta_E^2(K_{EE}) &= \tau^2\left(J_E^2 K_{EE}(s_E^*)^2 + J_I^2(\Delta K_{EI})^2\,\bar{\nu}_I^2 + J_I^2\langle K\rangle_{EI}\,s_I^2\right),
\end{aligned}
$$

$$
\begin{aligned}
\mu_I(W_I) &= \tau(J_E\langle K\rangle_{IE}\,\bar{\nu}_E - J_I\langle K\rangle_{II}\,\bar{\nu}_I + J_{\text{ext}} K_{\text{ext}}\,\nu_{\text{ext}}) + \Delta_I\,W_I \\
\sigma_I^2 &= \tau\left(J_E^2\langle K\rangle_{IE}\,\bar{\nu}_E + J_I^2\langle K\rangle_{II}\,\bar{\nu}_I + J_{\text{ext}}^2 K_{\text{ext}}\,\nu_{\text{ext}}\right) \\
\Delta_I^2 &= \tau^2\left(J_E^2(\Delta K_{IE})^2\,\bar{\nu}_E^2 + J_E^2\langle K\rangle_{IE}\,s_E^2 + J_I^2(\Delta K_{II})^2\,\bar{\nu}_I^2 + J_I^2\langle K\rangle_{II}\,s_I^2\right),
\end{aligned}
\tag{2.3.15}
$$

where $W_E, W_I \sim N(0,1)$.

Now the unknowns are $\bar{\nu}_E, s_E^2, \bar{\nu}_E^*, (s_E^*)^2, \bar{\nu}_I, s_I^2$. The equations that close the system are their

definitions:

$$
\begin{aligned}
\bar{\nu}_E &= \int_0^\infty \int_{-\infty}^\infty \nu_E\left(k, w, \bar{\nu}_E^*, (s_E^*)^2, \bar{\nu}_I, s_I^2\right) f_K(k) f_W(w)\,\mathrm{d}w\,\mathrm{d}k \\
s_E^2 &= \int_0^\infty \int_{-\infty}^\infty \nu_E\left(k, w, \bar{\nu}_E^*, (s_E^*)^2, \bar{\nu}_I, s_I^2\right)^2 f_K(k) f_W(w)\,\mathrm{d}w\;\mathrm{d}k - \bar{\nu}_E^2 \\
\bar{\nu}_E^* &= \int_0^\infty \int_{-\infty}^\infty \nu_E\left(k, w, \bar{\nu}_E^*, (s_E^*)^2, \bar{\nu}_I, s_I^2\right) f_K^*(k) f_W(w)\,\mathrm{d}w\,\mathrm{d}k \\
(s_E^*)^2 &= \int_0^\infty \int_{-\infty}^\infty \nu_E\left(z, w, \bar{\nu}_E^*, (s_E^*)^2, \bar{\nu}_I, s_I^2\right)^2 f_K^*(k) f_W(w)\,\mathrm{d}w\;\mathrm{d}k - (\bar{\nu}_E^*)^2 \\
\bar{\nu}_I &= \int_{-\infty}^\infty \nu_I\left(w, \bar{\nu}_E, s_E^2, \bar{\nu}_I, s_I^2\right) f_W(w)\,\mathrm{d}w \\
s_I^2 &= \int_{-\infty}^\infty \nu_I\left(w, \bar{\nu}_E, s_E^2, \bar{\nu}_I, s_I^2\right)^2 f_W(w)\,\mathrm{d}w - \bar{\nu}_I^2,
\end{aligned}
\tag{2.3.16}
$$

where f_W, f_K and f_K^* are the probability densities of a standard Gaussian random variable (f_W), of the excitatory in-degree of E neurons (f_K) and of the excitatory in-degree of E neurons which are pre-synaptic to a given E neuron (f_K^*). The last density depends on the joint in/out-degree distribution within the EE subnetwork $f_{\text{in,out}}$ through

$$
f_K^*(k) = \frac{1}{\langle K\rangle_{EE}} \int_0^\infty y\, f_{\text{in,out}}(k,\, y)\,\mathrm{d}y \tag{2.3.17}
$$

(this is just the result of putting together Eqs. (2.3.13) and (2.3.14)).

E/I network with an arbitrary degree distribution within the EE subnetwork and selective inhibition

In the preceding example, degrees within the EE subnetwork followed an arbitrary distribution, whereas the remaining connectivity was purely random (in the ER sense). But networks in which the EE in-degree distribution is much broader than in ER counterparts can present a clear unbalance in the inputs that E neurons receive. This feature is thought to be unrealistic because in physiological situations excitation and inhibition tend to cancel each other in the mean, leading to what has been called a *balanced state* [Okun and Lampl, 2008; Isaacson and Scanziani, 2011; Xue et al., 2014]. This condition is not only desirable in any model which tries to reproduce the properties of real neuronal networks but is also a necessary requirement for the

asynchronous stationary state to be reached within the context of the present network model. If the excitatory component of the input is much larger than the inhibitory one, some neurons may receive suprathreshold input and this induces regular spiking, which violates one of the assumptions needed for the mean-field theory to work. Here we define connectivity conditions that will compensate for the excess of excitation in the E population.

The structure of connections involving inhibitory neurons has been less studied than that of excitatory-to-excitatory synapses. In very general terms, and despite a wide variety of interneuron types, inhibitory connectivity in the mammalian cortex appears to be denser, less specific and more homogeneous than connectivity between pyramidal neurons [Fino and Yuste, 2011; Hofer et al., 2011; Fino et al., 2013]. This makes ER models a reasonable preliminary framework to represent connections from/to inhibitory neurons. On the other hand, experiments on cortical slices have shown that the magnitude of the inhibitory component of the input adapts to that of the excitatory component through plasticity mechanisms that modulate the strength of inhibitory connections [Xue et al., 2014]. Such adaptations could be responsible for maintaining a proper balance between excitation and inhibition at the single cell level [Landau et al., 2016].

We consider here a situation in which the connectivity within the II subnetwork and from the E to the I population is of ER type but where the connections from I to E appear independently with a probability that depends on the excitatory in-degree of the post-synaptic neuron, so as to compensate for the excess of excitation that some E neurons might receive. It would be desirable that the inhibitory in-degrees in the E population were such that the total mean input received did not depend on the excitatory in-degree. This is equivalent to

$$J_E K_{EE} \bar{\nu}_E^* - J_I K_{EI} \bar{\nu}_I = C, \tag{2.3.18}$$

where C is a constant. We obtain

$$K_{EI} = \frac{J_E K_{EE} \bar{\nu}_E^* - C}{J_I \bar{\nu}_I}. \tag{2.3.19}$$

Condition (2.3.19) states that the total inhibitory in-degree must grow linearly with the total excitatory in-degree K_{EE}, so the same should happen with the probability to receive a single inhibitory connection. The coefficient $\frac{J_E \bar{\nu}_E^*}{J_I \bar{\nu}_I}$, however, is not known a priori because it depends on the stationary firing rates. To proceed we will suppose that the mean firing rates are similar within both populations (and so is the mean biased E rate) so that the probability to generate a

single I→E connection is $p_{EI}(K_{EE}) = aK_{EE} + b$, with $a = \frac{J_E}{N_I J_I}$. The fact that the mean rates $\bar{\nu}_E^*$ and $\bar{\nu}_I$ do not exactly coincide will produce a degree-dependent bias in the total mean input received proportional to $J_E K_{EE}(\bar{\nu}_E^* - \bar{\nu}_I)$, but as long as the difference $(\bar{\nu}_E^* - \bar{\nu}_I)$ is not very large and the degrees remain finite, this is expected to introduce small balance differences between neurons (recall that J_E is assumed to be small).

Let us consider an E neuron with excitatory in-degree K_{EE} in such a network. Its inhibitory in-degree follows, as before, a Binomial distribution whose parameters depend on K_{EE}. This means that the same approximations that we did in the previous section are still valid now. The only new ingredient is the fact that $\langle K \rangle_{EI}$ and ΔK_{EI} are not constant but functions of K_{EE}:

$$\begin{aligned} \langle K \rangle_{EI}(K_{EE}) &= p(K_{EE}) N_I \\ (\Delta K_{EI})^2(K_{EE}) &= p(K_{EE})(1 - p(K_{EE})) N_I \\ p(k) &= a(k) + b. \end{aligned} \tag{2.3.20}$$

2.3.4 Comparison between mean-field theory and computer simulations

We next compare the theory developed in the previous sections with computer simulations of networks of LIF neurons. We have chosen for illustration networks in which the EE degrees follow (integer versions of) Gaussian and Gamma distributions. We explore first the case of distributions with small variance and we move, afterwards, to networks with highly heterogeneous degree distributions.

The mean-field theory presented here assumes that the network under study operates in a regime in which the voltage of individual neurons behaves like a stochastic diffusion process (see Appendix B). This *diffusion approximation* is valid whenever the synaptic weights J_E, J_I, J_{ext} are small compared with the threshold and all the neurons fire approximately as independent Poisson processes. Whereas the first condition is directly controlled by the parameter settings, there is no a priori condition which ensures that the second will be fulfilled.

In order to determine to what extent the dynamics of individual neurons resembles that of Poisson processes, we have included in our analysis the computation of the coefficient of variation (CV) of inter-spike intervals (ISI). The inter-spike interval is just the time lapse between spiking episodes of a given neuron. When the spiking process is stochastic, the ISI becomes a random variable whose moments can be estimated, in the stationary state, by means of temporal averages. The CV of the ISI is the ratio between the standard deviation of the ISI and its mean. It provides a

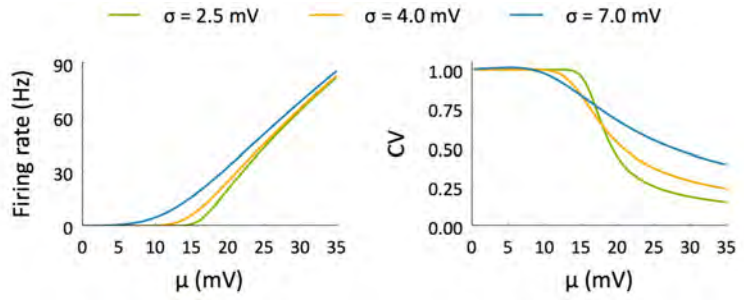

Figure 2.1: Firing rate (left, Eq. (2.2.3)) and CV of ISIs (right, Eq. (2.3.21)) of a neuron as a function of μ for different values of σ when $V_r = 10$ mV, $V_\theta = 20$ mV, $\tau = 20$ ms, $\tau_r = 2$ ms.

measure of the irregularity of the spiking process: highly irregular neurons have larger CVs and vice-versa. If a neuron spikes exactly as a Poisson process, the CV equals 1.

The CV of a neuron whose voltage obeys Eq. (2.2.1) when the input has the same properties that we assumed in this chapter can be analytically computed using the properties of first passage times of stochastic processes, as detailed in Appendices A and B. The final result is that, in the stationary state, the CV of a neuron which receives, within a time window of length τ, a total input with mean μ and standard deviation σ is

$$\mathrm{CV}^2 \;=\; 2\tau^2\pi\nu^2 \int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} e^{t^2} \int_{-\infty}^{t} e^{u^2}\,\mathrm{erfc}(-u)^2\,\mathrm{d}u\,\mathrm{d}t, \tag{2.3.21}$$

where ν is the neuron's stationary firing rate. Figure 2.1 shows how ν and the CV change as a function of μ and σ. The results presented in Sections (2.3.1) and (2.3.2) provide the details to compute the parameters μ and σ for each neuron in a network with arbitrary degree distributions. Therefore, the same tools can be applied to compute the CVs analytically.

When describing the dynamics of a closed neuronal network, this mean-field theory will only be exact when the firing process is close to Poisson. The computation of the CVs in the network provides a way to assess how far the network is from the Poisson assumption. The finding that the CVs are far from 1 will be a signature that the diffusion approximation is not totally correct and therefore the analytical results only describe the system with approximate accuracy.

Degree distributions with a low level of heterogeneity

We include in this category networks in which the degrees of the EE subnetwork follow distributions whose variance is close to the variance in ER networks of the same density (see Fig. 2.2 A).

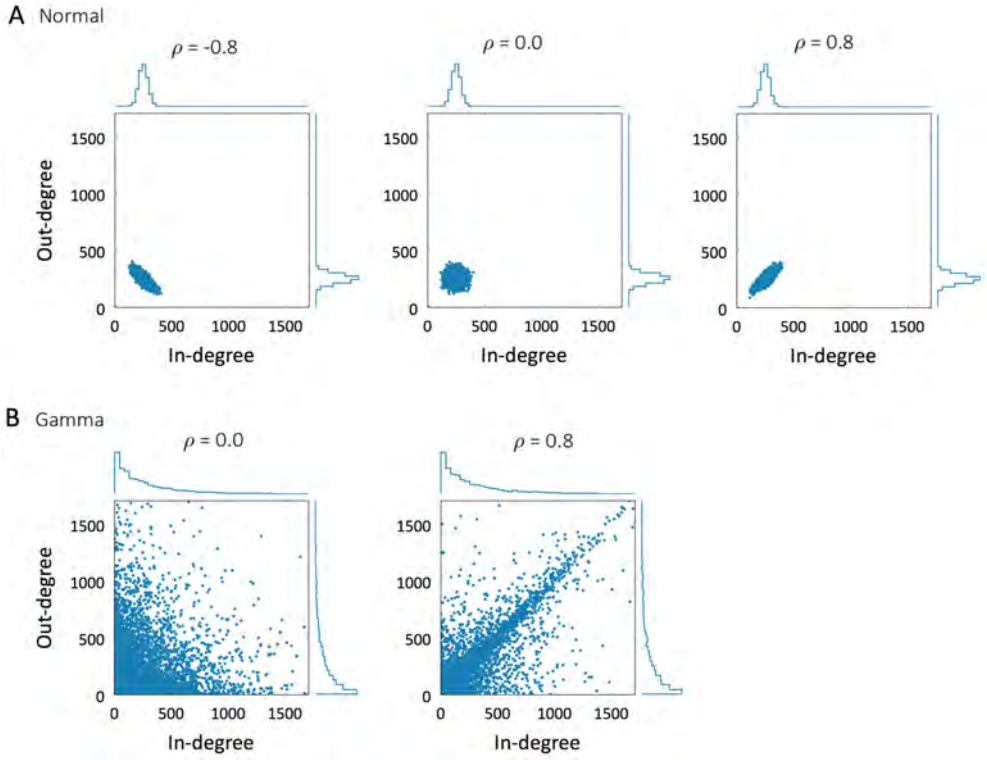

Figure 2.2: In/out-degree distributions within the excitatory subnetwork. **A** Networks whose degrees follow Normal(μ, σ) degree distributions with $\mu = 250$, $\sigma = 40$ and correlation coefficient $\rho = -0.8$ (left), $\rho = 0$ (middle) and $\rho = 0.8$ (right). **B** Networks with Gamma(κ, θ) degree distributions, $\kappa = 0.8$, $\theta = 312.5$ and correlation coefficient $\rho = 0$ (left) and $\rho = 0.8$ (right). In all the cases $N_E = 5000$.

Since these networks have a low level of degree heterogeneity within the excitatory population, inhibition has been set to be completely homogeneous: connections from and to I neurons are created with a fixed probability p that is the same for the entire population.

We looked first at networks whose EE structure is totally random (that is, ER-like). We set $J_I = -gJ_E$ with $g = 8$ and by varying the value of J_E we could modify the importance of recurrent connections. When $J_E = 0$, the network is just an assembly of equivalent independent neurons. Therefore, the system is characterized by a single CV and a single stationary firing rate ν, which are both well approximated by the analytical formulas (of course, the empirical results show a variety of CVs and rates, but this heterogeneity would be reduced by increasing the total time of integration). However, the CV may be far below 1, as in the example shown in Fig. 2.3, meaning that the spiking process is not exactly Poisson.

As we increase the synaptic coupling, the heterogeneity of in- and out-degrees starts to show up and this translates into a variability of firing rates and CVs (Fig. 2.3 E, F). Both distributions are well approximated by the theoretical predictions, except when the coupling becomes too strong (Fig. 2.3 H, I). When J_E is large enough, there is a clear disagreement between the theoretical

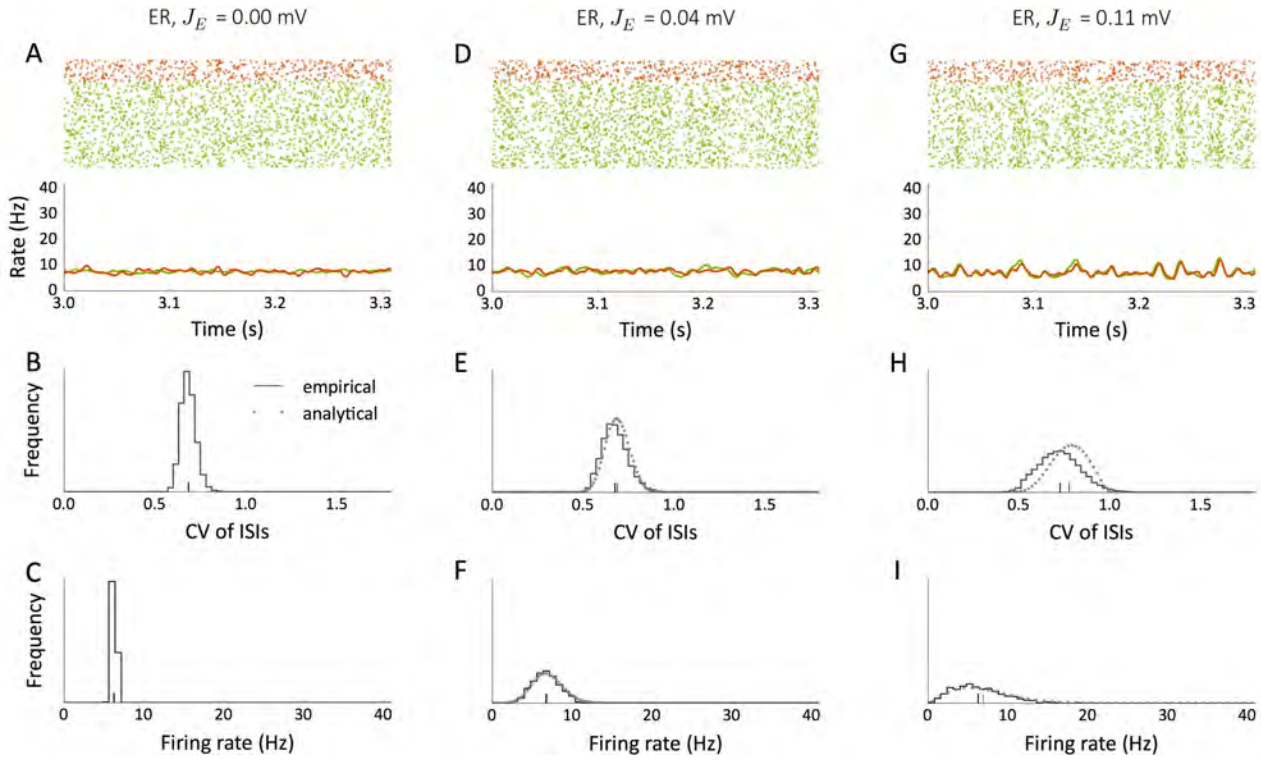

Figure 2.3: Dynamics of ER(p) networks with different degrees of coupling, $p = 0.05$. Continuous lines show the results from simulations and dotted lines correspond to the analytical formulas. **A** Raster plots and population firing rates of inhibitory (red) and excitatory (green) neurons when $J_E = 0$ mV, $\nu_{\text{ext}} = 6.5$ Hz. **B** Stationary distribution of CVs when $J_E = 0$ mV, $\nu_{\text{ext}} = 6.5$ Hz. **C** Stationary firing rate distribution when $J_E = 0$ mV, $\nu_{\text{ext}} = 6.5$ Hz. **D, E, F** Same as A, B, C for $J_E = 0.04$ mV, $\nu_{\text{ext}} = 6.9$ Hz. **G, H, I** Same as A, B, C for $J_E = 0.11$ mV, $\nu_{\text{ext}} = 7.17$ Hz. Vertical lines indicate empirical and analytical averages. In all the cases $N_E = 5000$, $N_I = 1250$, $\tau = 20$ ms, $V_r = 10$ mV, $V_\theta = 20$ mV, $\tau_r = 2$ ms, $J_I = -gJ_E$, $g = 8$, $K_{\text{ext}} = 1000$, $J_{\text{ext}} = 0.14$ mV. ν_{ext} was varied to get comparable mean firing rates.

CV distribution and the empirical one. This indicates that the hypotheses of the mean-field formulation do not exactly hold. Despite this, the analytical prediction for the distribution of firing rates is very close to the results obtained from the simulations (Fig. 2.3 F, I).

This example suggests that even when the Poisson approximation is not exact, the results regarding the distribution of firing rates provided by the mean-field theory can be quite accurate. We next asked ourselves what the effect would be of varying the correlation between EE degrees and if the mean-field extension described earlier would be able to capture it. We started by imposing degrees in the EE subnetwork which are integer versions of Bivariate Normal distributions, keeping a relatively small marginal variance. Figure 2.4 shows three examples of networks that only differ in the correlation coefficient ρ between in- and out- excitatory degrees in the EE subnetwork. The degree distributions of these examples coincide with those shown in Fig. 2.2 A, where $N_E = 5000$, the density of connections is $p = 0.05$ and the standard deviation of the EE degrees is set to 40, which is larger than that of an ER counterpart, $\sigma_{\text{ER}} = \sqrt{p(1-p)(N-1)} \approx 15.4$ but

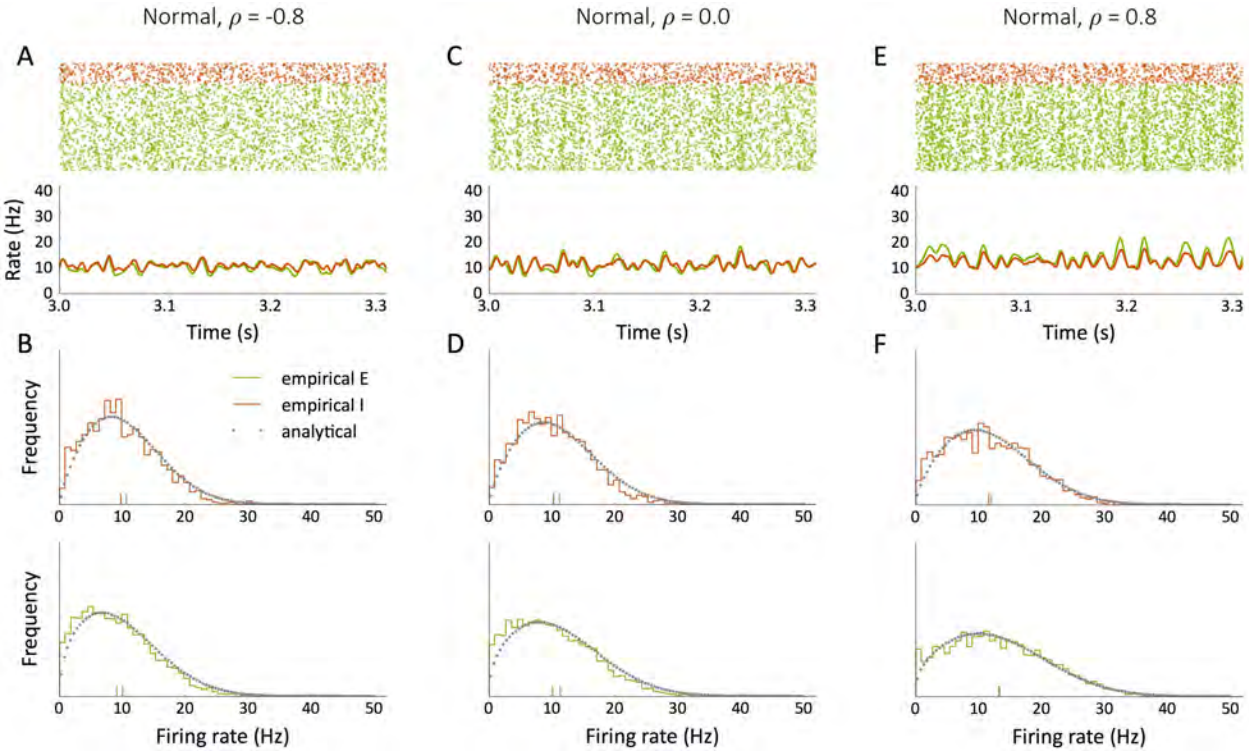

Figure 2.4: Dynamics of networks which only differ in the correlation coefficient ρ between in- and out-degrees within the EE subnetwork. In all the cases, the EE degree distribution is Normal(μ, σ) with $\mu = 250$ and $\sigma = 40$, as in Fig. 2.2 A. All the other connections are created independently with probability $p = 0.05$. **A** Raster plots and population firing rates of inhibitory (red) and excitatory (green) neurons when $\rho = -0.8$. **B** Firing rate distribution in the stationary state for I (red) and E (green) neurons, from simulations (continuous histogram) and from the analytical formula (grey dots) when $\rho = -0.8$. **C** Same as A for $\rho = 0$. **D** Same as B for $\rho = 0$. **E** Same as A for $\rho = 0.8$. **F** Same as B for $\rho = 0.8$. Vertical lines indicate empirical and analytical averages. In all the cases $N_E = 5000$, $N_I = 1250$, $\tau = 20$ ms, $V_r = 10$ mV, $V_\theta = 20$ mV, $\tau_r = 2$ ms, $J_E = 0.11$ mV, $J_I = -gJ_E$, $g = 8$, $K_{\text{ext}} = 1000$, $\nu_{\text{ext}} = 8.1$ Hz, $J_{\text{ext}} = 0.14$ mV.

it is close to it. The results suggest that the theory agrees well with simulations and is able to explain the modulation of the firing rate distribution when the only parameter varied is ρ.

Highly heterogeneous degree distributions

So far, we have seen that the presented theory gives an accurate prediction of the distribution of firing rates when the structure within the EE subnetwork is moderately heterogeneous. The next step is to study if the same equations can explain the behavior of the system when the structure is perturbed even more. We consider extreme cases of heterogeneous networks by using EE degrees which follow (integer versions of) Gamma(κ, θ) distributions with correlation coefficient ρ (generated as in Chapter 1, see p. 40). Figure 2.2 B shows two example EE degree distributions of this type. As noted before, in these cases we need to introduce a compensatory mechanism to

keep a balance between excitatory and inhibitory inputs. We do so by creating I→E connections with a probability that depends on the excitatory in-degree K_{EE}:

$$p(K_{EE}) \quad = \quad p + \frac{J_E}{J_I N_I}(K_{EE} - \langle K_{EE} \rangle), \tag{2.3.22}$$

where p is the probability with which E→I and I→I connections are defined.

Again, the theory predicts the observed firing rate distribution with good accuracy. In the examples of Fig. 2.5 (whose structures correspond to those in Fig. 2.2 B), the variation of the correlation coefficient ρ from 0 to 0.8 has a dramatic effect on the firing rates in the network, especially in the E population (Fig. 2.5 B, D). Notice that the two networks of this example only differ in ρ; their marginal degree distributions are statistically identical. This shows the great importance that the degree correlation has on dynamics. The presence of neurons which *at the same time* receive lots of inputs from the network and have an influence on many other neurons largely influences the dynamical properties of the network as a whole. Moreover, this effect of ρ is almost perfectly captured by the introduction of the moments of the biased rate distribution, $\bar{\nu}_E^*$, $(s_E^*)^2$, in the analytical formalism.

2.4 A possible functional role of degree correlations

We next explore a bit more the effect of ρ on neuronal dynamics. In the previous sections we have studied the role of ρ in shaping the repertoire of firing rates in a stationary, asynchronous state. We asked if the presence of such "in/out-hubs" could modify the response of the network to transient stimuli as well.

To do so, we applied transient pulses of stimulation to a fraction f_{ext} of excitatory neurons, chosen at random. The modified version of the voltage dynamics for a stimulated neuron i is, for t within the stimulation period,

$$\tau \frac{\mathrm{d}V_i(t)}{\mathrm{d}t} = -V_i(t) + \tau I_i(t) + \mu_{\text{ext}}, \tag{2.4.1}$$

where $I_i(t)$ takes into account all the inputs coming from other neurons (including the external ones).

The results are quite surprising. When the fraction of stimulated neurons is small, there is no response apart from a tiny increase in population rates while the stimulus is applied, independently of the network's topology. This behavior persists regardless of f_{ext} as long as the EE degrees are

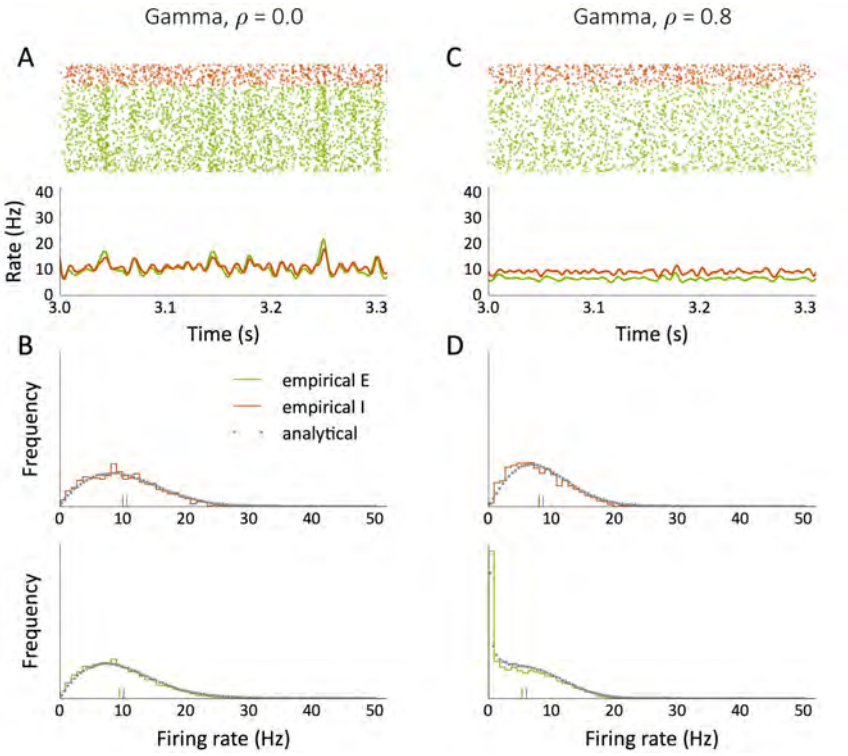

Figure 2.5: Dynamics of networks which only differ in the correlation coefficient ρ between in- and out-degrees within the EE subnetwork. In all the cases, the EE degree distribution is Gamma(κ, θ) with $\kappa = 0.8$ and $\theta = 312.5$, as in Fig. 2.2 B. **A** Raster plots and population firing rates of inhibitory (red) and excitatory (green) neurons when $\rho = 0$. **B** Firing rate distribution in the stationary state for I (red) and E (green) neurons, from simulations (continuous histogram) and from the analytical formula (grey dots) when $\rho = 0$. **C** Same as A for $\rho = 0.8$. **D** Same as B for $\rho = 0.8$. Vertical lines indicate empirical and analytical averages. In all the cases $N_E = 5000$, $N_I = 1250$, $p = 0.05$, $\tau = 20$ ms, $V_r = 10$ mV, $V_\theta = 20$ mV, $\tau_r = 2$ ms, $J_E = 0.11$ mV, $J_I = -gJ_E$, $g = 8$, $K_{\text{ext}} = 1000$, $\nu_{\text{ext}} = 8.1$ Hz, $J_{\text{ext}} = 0.14$ mV.

not correlated. However, for values of ρ significantly positive, the response changes dramatically. In this case there is a critical fraction of stimulated neurons above which the network transiently enters in a qualitatively distinct dynamical regime, characterized by bursting and synchronous activity. Such a "responsive" state lasts even after the stimulus has been removed, although the network eventually relaxes back to its original asynchronous state, as shown in Fig. 2.6.

This observation suggests that ρ is important in defining how external inputs are propagated through the network. The presence of a significant number of in/out-hubs probably enhances the transmission of information because such nodes act as "organizing centers": they receive information from a large fraction of the network and, simultaneously, they have the power to transmit it to a large number of other units.

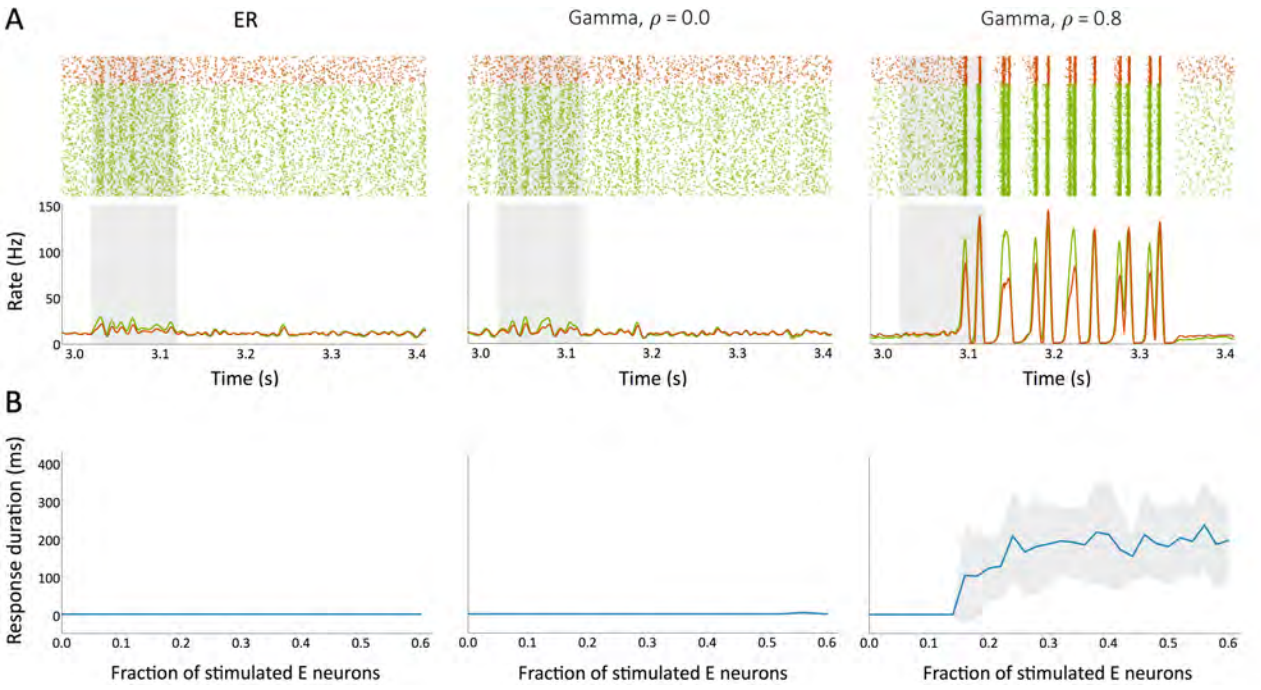

Figure 2.6: Response to transient stimulation of three different kinds of topologies for the EE subnetwork: ER (left), broad Gamma distribution without correlation (middle), and broad Gamma distribution with positive correlation (right). **A** Raster plots and population firing rates for a stimulation of magnitude $\mu_{\text{ext}} = 5$ mV applied to a fraction of $f_{\text{ext}} = 0.2$ E neurons during $t_{\text{ext}} = 100$ ms (shaded region). **B** Total duration of the network's response (average $\pm$ standard deviation) as a function of f_{ext} for $\mu_{ext} = 5$ mV and $t_{\text{ext}} = 100$ ms. The remaining parameters are the same as in Fig. 2.5.

2.5 Conclusion

In this chapter we have explored some of the consequences of imposing a heterogeneous topology in networks of leaky integrate-and-fire neurons. We have focused, in particular, on networks that are defined through a given joint distribution of incoming and outgoing connections. We say that these models are *heterogeneous* because they incorporate a richer variability across neurons than classical homogeneous or purely random models do. The models studied include networks for which degree distributions are broad even in the large N limit and networks with an arbitrary correlation coefficient between the degrees of individual neurons.

We have extended a well-known mean-field formalism [Brunel, 2000] to such a more general family of possible connectivities. The theory presented here does not exclude the common topologies, but can be applied to any network whose structure is determined solely by its degree distribution. Therefore, homogeneous and purely random (ER) models are still included in the formalism. It is important to bear in mind, however, that the theory does not apply to networks whose structure is determined by other principles: despite any network family has a characteristic degree distribution, this does not necessarily imply that the distribution *per se* defines its structure.

At the beginning of the chapter, we exposed the analytical formalism. First, we showed how the presence of in-degrees with a large variance can be introduced in the mean-field equations. It is well known that the presence of degree heterogeneity translates into an heterogeneity of firing rates in the stationary state [Amit and Brunel, 1997a; Roxin et al., 2011]. Therefore, neurons in heterogeneous networks can be parametrized by a (random) variable which defines its "rate identity" in the network. The rate of every neuron is then a function of this variable, which captures the neuron-to-neuron differences in terms of both the number of inputs received and the rates of the input units. In classical ER models, such a variable is a scalar and can be assumed to be normally distributed (because the in-degrees follow approximately Gaussian distributions). The new ingredient when dealing with a network whose in-degrees follow an arbitrary distribution is that the above mentioned neuron-to-neuron variability is no longer captured by a scalar but by a two-component vector. This is so because the two contributions to the neuronal variability −differences in in-degree and differences in the firing rates of input neurons− can no longer be grouped into a single variable, simply because the heterogeneity coming from differences in in-degree follows an arbitrary distribution in the network. The general formalism is analogous to the classical scenarios except for the fact that rate identities depend on two random variables, one which is Gaussian and another whose distribution is dictated by the in-degrees.

We then addressed the issue of having a network where the number of incoming and outgoing connections are correlated. We argued that such a degree correlation ρ plays a role in defining what is the *available* repertoire of firing rates within the set of possible input neurons to a given unit. In general, the collection of out-degrees in the set of pre-synaptic neurons to a given neuron is biased with respect to the out-degrees in the whole network. This is just because, in the set of pre-synaptic neurons, units with large out-degrees are more likely to occur. Therefore, when individual in- and out-degrees are correlated, this bias of out-degrees translates into a bias of in-degrees in the set of pre-synaptic neurons. Since the firing rate is a function of a neuron's in-degree, the final result is that the repertoire of firing rates in the set of possible inputs of a neuron can be biased with respect to the distribution of rates in the network. Such a pre-synaptic rate bias is key in shaping the macroscopic dynamical properties of the network in the stationary state.

The effect of such a rate bias can be analytically computed and introduced in the mean-field equations. The final mean-field formulation takes into account not only the moments of the firing rate distribution but also those of the biased rate. The macroscopic unknowns of the system are, then, the mean and variance of the rate distribution and also the moments of the biased rate distribution, and the equations can be self-consistently closed by using the definitions of these four quantities.

We compared the predictions of this extended formulation with the results of computer simulations. The predicted firing rate distributions are quite close to the empirical results, both when we use almost homogeneous and highly heterogeneous networks. The analytical formulation provides accurate predictions of the effect of varying ρ in these networks as well. Hence, these techniques constitute a powerful toolkit for anticipating some of the macroscopic properties of large LIF networks in the stationary state, even in the presence of great structural heterogeneity.

The role of the above-mentioned biased firing rate goes beyond its technical use in the described formalism. Our observations show that, in fact, the relevant parameter when studying the behavior of the system is precisely this biased distribution. What neurons really "perceive" is the firing rates in their set of neighbors, not in the entire network. A network might contain a large fraction of very active neurons which nevertheless do not project to any others. In terms of function, such neurons presumably do not play any role because they have no influence. Thus, the firing rate distribution restricted to the set of possible pre-synaptic neurons is a much more relevant dynamical magnitude. This suggests that determining the firing rate distribution in real networks

can be not very informative unless it is accompanied by a thorough study of connectivity.

We finally explored the role of ρ in the network's response to transient stimuli. We found that networks with significant degree correlations can experience dynamical states that are not available otherwise. In particular, positive correlations enhance the transmission of information, and can shift the dynamics into alternative regimes even after the stimulus has been removed. This behavior resembles that of short-term memory networks, in which the memory is stored as an activity pattern that is maintained in time even in the absence of the stimulus that originated it [Chaudhuri and Fiete, 2016]. Such observations could link degree correlations −which, as detailed in Chapter 1, are present in real cortical microcircuits− with enhanced abilities to transmit and process external inputs. Although we have not explored in detail these potentialities, we postulate that they can play a significant role in the way real networks respond to a variety of different stimuli.

Chapter 3

Eigenvalues of connectivity matrices with a prescribed degree distribution

Part of the work in this chapter has
been published [Aljadeff et al., 2016].

3.1 Introduction

In the previous chapter we have studied the effect of imposing a certain degree distribution in the repertoire of stationary firing rates of networks of spiking neurons. Another classical approach to the study of neuronal networks is the one provided by firing rate models. These models are less realistic in the sense that they do not describe the spiking events of individual neurons, but rather represent the neuronal dynamics through more macroscopic, or statistical, variables such as the firing rates. Rate models describe the temporal evolution of the firing rates (or associated quantities) of individual units (i. e., neurons or groups of neurons). Such models have been widely used in theoretical neuroscience as macroscopic approximations to the study of neuronal processes, and they tend to be simpler and more tractable than systems of spiking neurons.

Firing rate models are usually formulated by providing the temporal derivative of the firing rate, which is a function of the firing rates of all the units in the system. The synaptic coupling between units is key in defining how the rate of a single unit affects the rates of others. One of the simplest possibilities is that the derivative of the vector of rates is just the product of a matrix and the rate vector itself. In these cases the dynamics is simple and it is shaped by the spectrum, that is,

the set of eigenvalues, of this matrix.

Non-linear models are more realistic but can be extremely difficult to treat analytically. Under certain circumstances, however, the tools from the study of dynamical systems can be used to derive, at least, some of their qualitative properties. A classical example is a model defined through

$$\begin{aligned} \frac{\mathrm{d}h_i}{\mathrm{d}t} &= -h_i(t) + \sum_{j=1}^{N} J_{ij}\, r_j(t) \\ r_j(t) &= \phi(h_j(t)), \end{aligned} \tag{3.1.1}$$

where $r_i(t)$ represents the firing rate of unit i and $h_i(t)$ is a local field associated with it (for example some quantity related to the membrane potential of that unit). The function ϕ describes the relationship between the input (the field) and the output (the rate) of every unit. It is usually set to be a non-linear, sigmoid function, for example

$$\phi(x) = \tanh(x). \tag{3.1.2}$$

It should be noted that, since the hyperbolic tangent can take negative values, so too can firing rates in this model. In this case, $r_i(t)$ should be interpreted as the difference between the actual firing rate and a reference value.

Eq. (3.1.1) does not directly describe the evolution of the firing rates but rather that of the related quantities $\{h_i(t)\}_i$ (and therefore can be considered a firing rate model as well). This particular system was studied by Sompolinsky, Crisanty and Sommers in a seminal paper [Sompolinsky et al., 1988]. The authors assumed that the connectivity among units is such that each synaptic weight J_{ij} is an independent random variable, normally distributed with zero mean and variance g^2/N. They showed that, in the limit $N \to \infty$, the system undergoes a transition from a quiescent state to a chaotic state when g exceeds the critical value $g^* = 1$. The critical value coincides with the value of g for which the connectivity matrix starts to have eigenvalues with real part greater than 1. In fact, the linearization of (3.1.1) around the quiescent state $h_1 = \cdots h_N = 0$ is defined by the matrix $\mathbf{A} := \mathbf{J} - \mathbf{I}$, where $\mathbf{J}$ and $\mathbf{I}$ denote the connectivity and identity matrices, respectively. Therefore, the destabilization of the quiescent state, that occurs as long as $\mathbf{A}$ has eigenvalues whose real part is positive (or, equivalently, as long as $\mathbf{J}$ has eigenvalues with real part greater than 1), leads to a chaotic state. It can be shown that the spectrum of $\mathbf{J}$ is, in the large N limit, densely confined within a disk (in the complex plane) centered at the origin with radius g. This means that when g increases and crosses the critical value, there is a sudden appearance of many

eigenvalues whose real part surpasses 1. Thus, in this particular system, the spectrum of the connectivity matrix contains all the information needed to predict whether the dynamics will be quiescent or chaotic. A similar transition has been identified in networks of spiking neurons as the spectrum of the appropriate matrix includes eigenvalues with real part greater than 1 [Ostojic, 2014].

Taking this system as a starting point, Aljadeff and colleagues [Aljadeff et al., 2015b] asked whether an analogous transition would occur when the connectivity matrix has a richer structure. They again considered random matrices but supposed that the variance of each entry is no longer constant across the network but it can take a fixed number of values (and, again, all these values scale as $1/N$). This model might represent an heterogeneous network in which units belong to different groups and where each pair of pre-synaptic and post-synaptic groups has an associated parameter which defines the magnitude of the variance of the corresponding synaptic weights. Their finding is that the same type of transition occurs. Now the critical condition is that the spectral radius (that is, the larger modulus of all the eigenvalues) of the matrix of variances exceeds 1, where the matrix of variances is the $N \times N$ matrix whose entries are the variances of the elements in the random connectivity matrix. In parallel, they showed [Aljadeff et al., 2015a] that the spectrum of such "heterogeneous" connectivity matrices is also densely confined in a disk centered at zero, but now the radius of the disk is given by the square root of the spectral radius of the variance matrix. This result includes the homogeneous case as well, because in this case the spectral radius of the variance matrix is just g^2.

These examples show to what extent the spectral density of the matrix of synaptic weights can be important in determining dynamical properties of the system, although this is only true for specific types of systems. Outside of the field of neuroscience, the study of the spectrum of large random connectivity matrices is an intense research topic by itself, which has its roots in the works of Ginibre [Ginibre, 1965] and Girko [Girko, 1985] on random Gaussian matrices and nowadays continues to provide new and interesting results [Tao and Vu, 2010; Tao, 2013].

The aim of the work presented in this chapter is to study if it would be possible to further extend this spectral analysis to even more heterogeneous networks. As in Chapter 2, we focus on networks that have been generated to preserve a given degree distribution, which implies that the variance of the matrix elements is a distinct property of every entry, as we will show later on. Notice that the networks for which the previous studies have been carried out do not preserve Dale's law: a given unit can exert positive and negative effects on other units simultaneously.

One possible interpretation is that units are in fact neuronal clusters composed of both excitatory and inhibitory neurons so they can globally both excite and inhibit other units. But in order to represent the connectivity between single neurons other matrices should be used. We thus study the case of networks of excitatory and inhibitory neurons where the structure within the EE subnetwork is highly heterogeneous in the sense described earlier and where the rest of the connections appear randomly with a fixed probability. We show that it is indeed possible to predict some properties of the spectrum of such matrices in the large N limit. These results can shed some light in the study of neuronal systems with similar connectivity matrices for which the spectrum provides information about the possible dynamical regimes.

3.2 Preliminaries: networks with block structure

Aljadeff et al. [2015b] have studied the case of neural networks with a connectivity structure organized in groups. They consider networks of N neurons in which the $N \times N$ connectivity matrix has random and independent entries but where different entries may have different variances. The authors assume the neurons are arranged into distinct groups so that the variance of the random variable that gives the weight of the connection between a neuron i in group c_i and a neuron j in group c_j is determined by c_i and c_j. The number of different groups is kept constant as $N \to \infty$ and all the synaptic efficacies scale as $1/\sqrt{N}$. This can represent a scenario where there is a limited number of distinct neuronal types that is independent of the network size N. The laws controlling how neurons are connected depend on the pre- and post-synaptic neuronal types.

In order to properly review their result we need to make use of some definitions. First, we assume that there is a finite quantity D which gives the number of neuronal groups. The fraction of neurons that belong to group d is given by α_d, so that $\sum_{d=1}^{D} \alpha_d = 1$. The index of the group to which neuron i belongs is the number c_i defined as

$$c_i = \left\{ c \in \{1, \cdots, D\} : \frac{i}{N} \in \left(\sum_{d=1}^{c-1} \alpha_d, \sum_{d=1}^{c} \alpha_d \right] \right\}. \tag{3.2.1}$$

The connection weight from neuron j to neuron i is then given by

$$J_{ij} = \frac{1}{\sqrt{N}} g_{c_i c_j} X_{ij}, \tag{3.2.2}$$

where $\{g_{cd}\}_{c,d} \in \mathbb{R}^+$ are parameters of the model and X_{ij} is a random variable with mean zero

and unit variance. The $\{X_{ij}\}_{i,j}$ variables are i.i.d. (independent and identically distributed), so the $N \times N$ matrix $\mathbf{X} := (X_{ij})_{i,j}$ is an i.i.d. random matrix (see Appendix C). The distribution of X_{ij} is known as the *atom* distribution of $\mathbf{X}$. Therefore, the connectivity matrix has independent random entries, with mean zero and variance that depends on the neuronal type and scales as $1/N$. The authors use the name *block matrix* to denote a matrix of this type because it contains D^2 different "blocks" inside of which the random laws are preserved.

The main result of [Aljadeff et al., 2015b] is that, in the limit $N \to \infty$, the spectrum of the random connectivity matrix $\mathbf{J} := (J_{ij})_{i,j}$ is circularly symmetric and its radius $R_{\mathbf{J}}$ is given by

$$R_{\mathbf{J}} = \sqrt{r(\mathbf{G})}, \tag{3.2.3}$$

where $r()$ denotes the spectral radius of a matrix and $\mathbf{G} = (G_{cd})_{c,d}$ is the $D \times D$ deterministic matrix which gives the (weighted) variances of the different neuronal associations:

$$G_{cd} = \alpha_d \, g_{cd}^2. \tag{3.2.4}$$

This result has been enunciated properly in Appendix C. It essentially states that, in the large N limit, the density of the eigenvalues of $\mathbf{J}$ has as support the circle centered at 0 with radius $R_{\mathbf{J}}$. Not only this, it can also be proven that there are no eigenvalues outside of this disk (almost surely when $N \to \infty$). Notice that in (3.2.3) we can substitute $\mathbf{G}$ by the $N \times N$ matrix of the variances of $\mathbf{J}$, $\left(\frac{1}{N} g_{c_i c_j}^2\right)_{i,j}$, because the latter has the same eigenvalues as $\mathbf{G}$ plus zeroes.

3.3 Networks with an infinite number of blocks

The case of networks with block structure and a finite number of blocks cannot represent more general connectivity matrices, for example those in which the connectivity rules are unique to each neuron. Therefore, it is natural to ask if the previous results could be extended to random matrices whose entries have unique variances. To represent such a situation, let us imagine that we have an $N \times N$ connectivity matrix $\mathbf{J} = (J_{ij})_{i,j}$ whose entries are

$$J_{ij} = \frac{1}{\sqrt{N}} \, g_{ij} \, X_{ij}, \tag{3.3.1}$$

where g_{ij} are parameters unique of each neuronal pair (i, j). In particular, we assume that they are given by a function $g_N : [0,1] \times [0,1] \to \mathbb{R}^+$ through

$$g_{ij} = g_N\left(\frac{i}{N}, \frac{j}{N}\right). \tag{3.3.2}$$

We use the notation g_N to make clear that this function might depend on N.

We now follow the ideas presented in [Aljadeff et al., 2016] to justify that, under certain conditions, it is possible to approximate this sequence of matrices by block matrices. In these block matrices, the number of neuronal groups $D = D(N)$ is no longer constant but increases with N. However, they are defined so that $D(N)$ grows sublinearly with N, which means that the ratio between the number of groups and the number of neurons tends to zero as $N \to \infty$ while the number of neurons in each group still grows arbitrarily with N.

Here we describe how we can do this approximation under conditions more restrictive than those presented in [Aljadeff et al., 2016]. This will make the presentation clearer.

We first suppose that for sufficiently large N, g_N is Lipschitz with constant $K > 0$ (independent of N). Recall that this means that

$$|g_N(x) - g_N(y)| \leq K \|x - y\| \quad \text{for all } x, y \in [0,1] \times [0,1]. \tag{3.3.3}$$

We want to justify that, for N sufficiently large, we can approximate the matrix $\mathbf{J}$ by a block matrix with $D(N) = [N^\alpha]$ neuronal groups with $\frac{1}{2} < \alpha < 1$ and where each group has (almost) the same number of neurons. To do so, let us fix N and define the group to which each neuron belongs,

$$c_i = \left\{ c \in \{1, \cdots, D(N)\} : \frac{i}{N} \in \left(\frac{c-1}{D(N)}, \frac{c}{D(N)}\right] \right\}. \tag{3.3.4}$$

Now we consider the $D(N)^2$-block matrix $\tilde{\mathbf{J}} = (\tilde{J}_{ij})_{i,j}$ defined by

$$\tilde{J}_{ij} = \frac{1}{\sqrt{N}} \tilde{g}_{ij} X_{ij}, \tag{3.3.5}$$

where

$$\tilde{g}_{ij} = g_N\left(\frac{c_i - \frac{1}{2}}{D(N)}, \frac{c_j - \frac{1}{2}}{D(N)}\right). \tag{3.3.6}$$

It is convenient to interpret (3.3.5) as if the randomness coming from X_{ij} were fixed (that is, $\mathbf{J}$ and $\tilde{\mathbf{J}}$ are defined through the same realization of the random matrix $\mathbf{X}$). This clearly defines a block

matrix because all the neural pairs (i,j) that belong to the same group pair $(c,d) \in D(N) \times D(N)$ have the same variance. Moreover, the variance in a given block coincides with the variance of the original matrix in the middle of that block.

In the large N limit, $g_{ij} - \tilde{g}_{ij} = \mathcal{O}\left(\frac{1}{N^\alpha}\right)$:

$$\begin{aligned} |g_{ij} - \tilde{g}_{ij}| &= \left| g_N\left(\frac{i}{N}, \frac{j}{N}\right) - g_N\left(\frac{c_i - \frac{1}{2}}{D(N)}, \frac{c_j - \frac{1}{2}}{D(N)}\right)\right| \\ &\leq K \left\| \left(\frac{i}{N}, \frac{j}{N}\right) - \left(\frac{c_i - \frac{1}{2}}{D(N)}, \frac{c_j - \frac{1}{2}}{D(N)}\right)\right\| \\ &\leq K \frac{1}{D(N)} \\ &\leq C \frac{1}{N^\alpha} \end{aligned} \tag{3.3.7}$$

for all $i,j \in \{1, \cdots, N\}$, where C is a constant. Therefore, the difference between the entries of $\mathbf{J}$ and the entries of $\tilde{\mathbf{J}}$ approaches zero at least as $\mathcal{O}\left(\frac{1}{N^\beta}\right)$, where $\beta := \alpha + \frac{1}{2} > 1$. This element-wise scaling ensures that the norms of our matrices coincide in the limit $N \to \infty$:

$$\begin{aligned} \left\|\mathbf{J} - \tilde{\mathbf{J}}\right\|_1 &:= \max_{1 \leq j \leq N} \sum_{i=1}^{N} |J_{ij} - \tilde{J}_{ij}| \\ &\leq \frac{C}{N^\beta} \max_{1 \leq j \leq N} \sum_{i=1}^{N} |X_{ij}| \\ &= \mathcal{O}\left(\frac{1}{N^{\beta-1}}\right), \end{aligned} \tag{3.3.8}$$

so $\left\|\mathbf{J} - \tilde{\mathbf{J}}\right\|_1 \to 0$ as $N \to \infty$. As a consequence, for N large enough, the operators $\mathbf{J}$ and $\tilde{\mathbf{J}}$ behave essentially in the same way. This applies in particular to eigenvectors, so the spectrum of $\mathbf{J}$ and $\tilde{\mathbf{J}}$ become closer and closer as N grows. We therefore can study the spectrum of the sequence of block matrices $\{\tilde{\mathbf{J}}\}_N$ to approximate the spectrum of $\{\mathbf{J}\}_N$ when N is large enough.

Notice that the blocks of $\tilde{\mathbf{J}}$ have size $\frac{N}{D(N)} = \mathcal{O}(N^{1-\alpha})$ with $1 - \alpha > 0$. Thus, the block size becomes arbitrarily large as $N \to \infty$. We have pointed out that the spectrum of matrices of this type can be analytically computed in the limit in which the block size gets arbitrarily large. We postulate that this result can be applied to matrices like $\tilde{\mathbf{J}}$ *even when* the number of blocks increases with N as long as the number of elements in each group goes to infinity too.

We conclude that, in the limit $N \to \infty$, the spectrum of $\mathbf{J}$ is circularly symmetric and its spectral radius $R_{\mathbf{J}}$ is

$$R_{\mathbf{J}} = \sqrt{r(\tilde{\mathbf{G}})}, \tag{3.3.9}$$

where $\tilde{\mathbf{G}} = (\tilde{G}_{cd})_{c,d}$ is the $D(N) \times D(N)$ deterministic matrix defined by

$$\tilde{G}_{cd} = \frac{1}{D(N)}\, \tilde{g}_{cd}^2. \tag{3.3.10}$$

As we have mentioned before and in Appendix C, the spectral radius of $\tilde{\mathbf{G}}$ coincides with the spectral radius of the corresponding $N \times N$ matrix of variances $\tilde{\mathbf{G}}^{\mathbf{bl}} = (\tilde{G}^{\mathrm{bl}}_{ij})_{i,j}$, which is a deterministic block matrix:

$$\tilde{G}^{\mathrm{bl}}_{ij} := \frac{1}{N}\, \tilde{g}_{ij}^2. \tag{3.3.11}$$

If the function g_N is bounded for N large enough, that is, if for $N > N_0$

$$g_N(x,y) \leq g_{\max} \quad \text{for all } (x,y) \in [0,1] \times [0,1], \tag{3.3.12}$$

then

$$\left| g_{ij}^2 - \tilde{g}_{ij}^2 \right| \;\leq\; 2 g_{\max} \left| g_{ij} - \tilde{g}_{ij} \right| \;=\; \mathcal{O}\left(\tfrac{1}{N^{\alpha}}\right) \tag{3.3.13}$$

by virtue of (3.3.7). We can reason as before to conclude that the spectral properties of the block matrix of variances $\tilde{\mathbf{G}}^{\mathbf{bl}}$ and the true matrix of variances $\mathbf{G} = (G_{ij})_{i,j}$ defined by

$$G_{ij} := \frac{1}{N}\, g_{ij}^2 \tag{3.3.14}$$

coincide in the large N limit. This finally allows us to conclude that the spectral radius of $\mathbf{J}$ can be computed as

$$R_{\mathbf{J}} = \sqrt{r(\mathbf{G})}. \tag{3.3.15}$$

As a final remark, if all the variances are strictly positive, then $\mathbf{G}$ has positive entries and, due to the Perron-Frobenius Theorem, it has a positive eigenvalue that coincides with its spectral radius. Thus, in this case finding $R_{\mathbf{J}}$ involves computing the largest eigenvalue of the matrix of variances $\mathbf{G}$.

3.4 Application to networks with a prescribed degree distribution

We have seen that it is possible to predict, in the limit $N \to \infty$, the spectral radius $R_{\mathbf{J}}$ of a random matrix whose entries have mean zero and unique variance under certain assumptions.

The results we have reviewed also state that the spectral density of such a matrix converges to a measure in the complex plane whose support is a disk centered at zero with radius $R_{\mathbf{J}}$. We would like to exploit this result in order to predict properties of the limiting spectral measure in the case of connectivity matrices for networks with a given joint degree distribution.

Let us consider N_E excitatory and N_I inhibitory neurons, $N = N_E + N_I$. We construct a network where connections within the EE subpopulation come from a prescribed in/out-degree distribution and where all the other connections simply appear independently with probability p. As we described in the previous chapters, this is done by defining a sequence of in- and out-degrees $S = \{(K_j^{\text{in}}, K_j^{\text{out}})\}_{j=1}^{N_E}$ and taking, uniformly at random, a pairing of incoming and outgoing connections. We assume that the degrees are large and the networks are sparse, which means that the degree distribution is kept constant as N increases. In Chapter 1 (p. 34) we saw that, in such a network,

$$P(i \to j \,|\, K_j^{\text{in}} = k, K_i^{\text{out}} = k') \;\; = \;\; \frac{kk'}{N_E \langle K \rangle} \tag{3.4.1}$$

for N_E large enough, where $\langle K \rangle$ is the mean in/out-degree.

Now we use this result to construct a network in the following way: we generate a degree sequence $\tilde{S} = \{(\tilde{K}_j^{\text{in}}, \tilde{K}_j^{\text{out}})\}_{j=1}^{N_E}$ from a joint degree distribution with probability density function $\tilde{f}_{(\text{in, out})}$. We call these degrees *auxiliary*. For every ordered pair of neurons (i, j), we draw a connection from i to j with probability

$$P(i \to j \,|\, \tilde{K}_j^{\text{in}} = k, \tilde{K}_i^{\text{out}} = k') \;\; = \;\; \frac{kk'}{N_E \langle \tilde{K} \rangle}. \tag{3.4.2}$$

This procedure generates a network whose degree sequence is no longer $\tilde{S}$ but it is close to it. In fact, once conditioned to the auxiliary degree, the *real* degree follows a Binomial distribution whose expectation and variance are

$$\begin{aligned} &\mathbb{E}[K_i^{\text{in}} \,|\, \tilde{K}_i^{\text{in}}] = \tilde{K}_i^{\text{in}} \qquad && \text{Var}(K_i^{\text{in}} \,|\, \tilde{K}_i^{\text{in}}) = \tilde{K}_i^{\text{in}} \left(1 - \tfrac{\tilde{K}_i^{\text{in}}}{N_E}\right) \\ &\mathbb{E}[K_i^{\text{out}} \,|\, \tilde{K}_i^{\text{out}}] = \tilde{K}_i^{\text{out}} \qquad && \text{Var}(K_i^{\text{out}} \,|\, \tilde{K}_i^{\text{out}}) = \tilde{K}_i^{\text{out}} \left(1 - \tfrac{\tilde{K}_i^{\text{out}}}{N_E}\right), \end{aligned} \tag{3.4.3}$$

which means not only that the degree sequence is preserved *on average* but also that in the limit of large degrees the differences between the real and the auxiliary sequences become negligible. It can be easily seen that the expectation of the degrees and their covariance are preserved, whereas

the variance becomes larger:

$$\begin{aligned} \mathrm{Var}(K_i^{\mathrm{in}}) &= \mathrm{Var}(\tilde{K}_i^{\mathrm{in}}) + \langle \tilde{K} \rangle \\ \mathrm{Var}(K_i^{\mathrm{out}}) &= \mathrm{Var}(\tilde{K}_i^{\mathrm{out}}) + \langle \tilde{K} \rangle. \end{aligned} \tag{3.4.4}$$

We use this procedure as an alternative way of defining networks with a given degree distribution, keeping in mind that the final degree sequence will be perturbed in the ways described earlier. For the sake of simplicity, we will use the notation

$$x_i := \frac{\tilde{K}_i^{\mathrm{in}}}{\sqrt{N_E \langle K \rangle}} \qquad y_i := \frac{\tilde{K}_i^{\mathrm{out}}}{\sqrt{N_E \langle K \rangle}}. \tag{3.4.5}$$

Once the auxiliary degree sequence is defined, the complete $N \times N$ connectivity matrix $\mathbf{J} = (J_{ij})_{i,j}$ (including synaptic weights) is defined by

$$J_{ij} = J_E \, w_{ij} Z_{ij}, \text{ where}$$

$$w_{ij} := \begin{cases} 1 & \text{if } j \leq N_E \\ -w_0 & \text{otherwise} \end{cases}, \qquad Z_{ij} \sim \text{Bernoulli}(p_{ij}), \qquad p_{ij} := \begin{cases} x_i y_j & \text{if } i, j \leq N_E \\ p & \text{otherwise} \end{cases} \tag{3.4.6}$$

and the $\{Z_{ij}\}_{i,j}$ random variables are all to all independent. The parameter $J_E > 0$ gives the strength of excitatory connections and $w_0 > 0$ represents the ratio between the absolute values of inhibitory and excitatory synaptic efficacies ($J_I = -w_0 J_E$). We assume that the distribution of all the degrees is kept constant as N increases, which means $p = \mathcal{O}\left(\frac{1}{N}\right)$. Notice that according to definition (3.4.6), the diagonal entries of the connectivity matrix are different from zero with non-zero probability. One possible interpretation is that the units do not represent individual neurons but groups composed of many neurons, so that a given unit can influence itself. But even if the matrix is supposed to reflect the connectivity between individual neurons, the diagonal elements will have a negligible role in defining the spectrum as $N \to \infty$. Assuming that the connectivity is the one defined by (3.4.6) makes computations easier, as we will see in the next sections.

$\mathbf{J}$ is a random matrix whose entries are independent but not identically distributed: each entry has its own mean and variance. Now we analyze these two components (the mean and the variance) separately. Since J_E is a common factor to all the entries of the connectivity matrix, all the eigenvalues of $\mathbf{J}$ are obtained by finding the eigenvalues of the matrix with $J_E = 1$ and then multiplying by J_E. In the following we will assume $J_E = 1$ to simplify computations and we will

remind how the spectrum changes for arbitrary J_E at the end.

3.4.1 Variances

The (deterministic) matrix of variances is $\mathbf{G} = (G_{ij})_{i,j}$ with

$$G_{ij} \;=\; w_{ij}^2\, p_{ij}(1-p_{ij}) \;=\; \mathcal{O}\left(\tfrac{1}{N}\right). \tag{3.4.7}$$

Its elements are therefore defined as

$$\begin{aligned} G_{ij} &= \tfrac{1}{N} g_{ij}^2 \\ g_{ij} &= w_{ij}\sqrt{N p_{ij}(1-p_{ij})}. \end{aligned} \tag{3.4.8}$$

Our connectivity matrix has four main blocks which represent the connections E→E, E→I, I→E and I→I. The last three blocks have a uniform structure, so, in terms of their variance, they can be treated as in the case of block matrices with a constant number of blocks. The problematic block is that of EE connectivity. In order to properly make an analogy between the variances of the EE subnetwork and those of the previous section, we should be able to express g_{ij} as in (3.3.2):

$$g_{ij} = g_N\left(\frac{i}{N}, \frac{j}{N}\right), \tag{3.4.9}$$

where the function g_N is Lipschitz with constant $K > 0$ for every N.

As we will justify now, this cannot be done in general for the presented model. Being able to express $\{g_{ij}\}_{i,j}$ as in (3.4.9) through a Lipschitz function implies that these parameters behave "regularly" with respect to the normalized ordering of neuronal indices. Let us imagine that we order neurons according to their auxiliary in-degree $\{\tilde{K}_i^{\text{in}}\}_i$. Then, for $i, j \leq N_E$, we can write

$$\begin{aligned} g_{ij} &= \sqrt{N p_{ij}(1-p_{ij})} \\ p_{ij} &= \tfrac{1}{N_E\langle K\rangle} f_{\text{in}}\left(\tfrac{i}{N_E}\right) f_{\text{out}}\left(\tfrac{j}{N_E}\right), \end{aligned} \tag{3.4.10}$$

where f_{in} and f_{out} are functions which return the auxiliary degrees once the normalized neuronal index is given. The auxiliary degrees are random numbers that come from a given distribution, so there is no deterministic function which gives these degrees once the neuronal index is known. In the large N limit, however, the empirical distribution of auxiliary degrees approaches the theoretical one. If the neurons are ordered according to their in-degree and we denote by F_{in} the

distribution function of the auxiliary in-degree, in this limit we have

$$f_{\text{in}}(x) = F_{\text{in}}^{-1}(x), \quad x \in [0,1]. \tag{3.4.11}$$

This means that, when the neurons are ordered according to their auxiliary in-degree, this degree is obtained by applying the inverse of its distribution function to the corresponding index. The same cannot be done with out-degrees, though. To obtain the auxiliary out-degree from the in-degree index, there should be a mapping from the in-degree index to the out-degree index, and then we should apply the inverse of the out-degree distribution function. Such a mapping cannot be properly defined for a general in/out-degree distribution. For example, when in- and out-degrees are independent, there is no relationship between the degrees of individual neurons, and this is the reason why this mapping does not exist. Another way to express this difficulty is by noticing that neurons which have very similar in-degrees can have very different out-degrees, so a putative function g_N which maps the normalized indices into the connection probability would not be Lipschitz.

This observation poses an important problem when trying to apply the results of the previous section to this particular network structure. Despite of this, we asked if the previous theory could still be used to predict the spectrum under such "new" circumstances. To this end, we created matrices whose elements have mean zero and variance given by (3.4.7). We chose binary atom distributions and Gamma(κ, θ) distributions for the auxiliary degrees, with correlation coefficient ρ. If the theoretical predictions of the previous section still apply to this case, in the large N limit the spectrum should be confined within a disk centered at 0 with radius $R_{\mathbf{J}} = \sqrt{r(\mathbf{G})}$, where $r(\mathbf{G})$ is the spectral radius of the matrix of variances $\mathbf{G}$.

In the next sections we show a way to analytically compute this radius as a function of the system's parameters. Using these predictions, we computed the radii for different network sizes and compared the results with direct computation of the spectra from network realizations. Figure 3.1 shows this comparison. It seems that the limit spectral measure is close to a measure with support given by a disk centered at 0 with radius $R_{\mathbf{J}}$. Moreover, there are no eigenvalues outside of this disk. We conclude from these examples that the prediction also works in the case of networks where the variances vary from element to element even when they are not the result of applying a Lipschitz function to the pair of normalized neuronal indices.

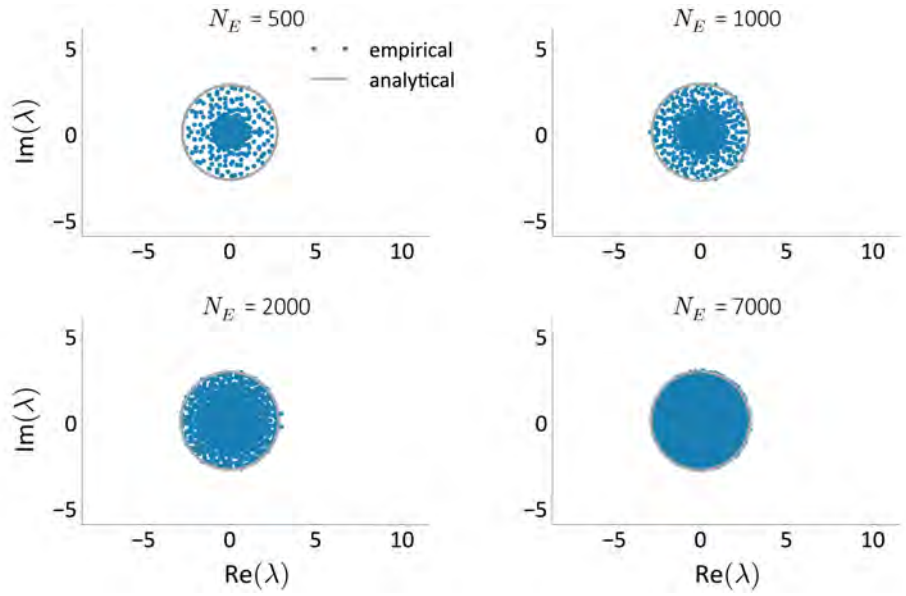

Figure 3.1: Spectra of random connectivity matrices whose atom distribution has zero mean and where each entry has its own variance, which coincides with the variance of matrices constructed according to Eq. (3.4.6). The plots only differ in the size of the network (and in the parameter p, which scales inversely with it). Every plot shows the analytical prediction of the spectrum (grey) and the result of empirically computing the spectrum of a single realization of the matrix (blue). In all the cases $N_I = \frac{N_E}{4}$, $p = \frac{50}{N_E+N_I}$, $J_E = 0.11$, $w_0 = 8$. The EE degrees come from Gamma(κ, θ) distributions, $\kappa = 0.8$, $\theta = \frac{pN_E}{\kappa}$, with correlation coefficient $\rho = 0.8$.

3.4.2 Means

Another important difference between the networks described in the previous section and the connectivity matrices that result from imposing a given in/out-degree distribution is that in the latter case the matrix entries have positive mean. The connectivity matrix $\mathbf{J}$ in this case can be interpreted as a sum of a random matrix $\mathbf{Y} = (Y_{ij})_{i,j}$ whose entries have mean zero and the deterministic matrix of the means $\mathbf{M} = (M_{ij})_{i,j}$:

$$\begin{aligned} &\mathbf{J} = \mathbf{M} + \mathbf{Y} \\ &M_{ij} = p_{ij} \\ &Y_{ij} = \sqrt{G_{ij}}\, X_{ij}, \end{aligned} \tag{3.4.12}$$

where p_{ij} and G_{ij} are the same as before (see Eqs. (3.4.6), (3.4.7)) and $\{X_{ij}\}_{i,j}$ are i.i.d. random variables with mean 0 and unit variance. We have already seen that, for N large enough, the random matrix $\mathbf{Y}$ has a spectrum confined in a disk centered at 0 with radius $R_{\mathbf{J}} = \sqrt{r(\mathbf{G})}$. Now the question is what happens to the spectrum when we apply a deterministic perturbation defined by $\mathbf{M}$.

In Appendix C we have presented a theorem by Tao [Tao, 2013] which defines how the spectrum of random matrices with uniform variance changes after the application of a low rank perturbation. Here a low rank perturbation is an $N \times N$ matrix $\mathbf{C}_N$ whose rank is $\mathcal{O}(1)$, which means that it has also $\mathcal{O}(1)$ eigenvalues different from zero. Another requirement is that $\mathbf{C}_N$ has $\mathcal{O}(1)$ operator norm. The theorem states the following: if the perturbation matrix $\mathbf{C}_N$ has $j_N = \mathcal{O}(1)$ eigenvalues outside of the disk that defines the spectrum of the original random matrix, then the perturbed matrix has exactly j_N eigenvalues outside of this disk and they are arbitrarily close (as $N \to \infty$) to those of $\mathbf{C}_N$.

In the case presented here, the perturbation matrix is the matrix of the means $\mathbf{M}$. We will show that it is indeed a low rank matrix: $\text{rank}(\mathbf{M}) \leq 3$ for any $N \geq 3$. Moreover, its operator norm scales as $\mathcal{O}(1)$ provided that $p = \mathcal{O}\left(\frac{1}{N}\right)$ and the auxiliary degrees are kept constant as $N \to \infty$ (see p. 100 for details). We postulate that these results might be applied to our connectivity matrices as well. In this case, the final spectrum will be a combination of a bulk of eigenvalues in the disk centered at 0 with radius $R_{\mathbf{J}}$ and three (or less) outliers given by the outliers of the matrix of means $\mathbf{M}$. In Materials and Methods at the end of this chapter we prove that $\mathbf{M}$ is a low rank matrix by directly computing its characteristic polynomial.

3.4.3 Computation of the spectrum

As we have already discussed, we hypothesize that the spectrum of connectivity matrices from a prescribed degree distribution can be predicted analytically using results that have been proved to be true for more "regular" matrices. The summary of such results applied to our system is the following: in the large N limit, the spectrum of the connectivity matrix defined by (3.4.6) and $J_E = 1$ is composed of a bulk of eigenvalues in the disk centered at zero with radius $R_{\mathbf{J}} = \sqrt{r(\mathbf{G})}$ and additional eigenvalues (outliers) given by the non-zero eigenvalues of the matrix $\mathbf{M}$. $\mathbf{M}$ and $\mathbf{G}$ are, respectively, the matrices whose entries give the average and the variance of the entries of the connectivity matrix $\mathbf{J}$ when $J_E = 1$. Therefore, in order to predict the properties of the spectrum of $\mathbf{J}$ we should have a way to compute the spectral radius of $\mathbf{G}$ and the non-zero eigenvalues of $\mathbf{M}$. In the case of arbitrary $J_E > 0$, both the radius and the outliers get multiplied by J_E.

We have computed the characteristic polynomials of $\mathbf{G}$ and $\mathbf{M}$, $q_{\mathbf{G}}$ and $q_{\mathbf{M}}$ (the details are shown Materials and Methods). The results are the following: if we conceive the auxiliary degree sequence as a set of fixed parameters, these polynomials are functions of what we call "empirical

statistics" of the normalized degree sequence:

$$\begin{array}{lll} \mathcal{S}_x := \sum_{i=1}^{N_E} x_i & \mathcal{S}_y := \sum_{i=1}^{N_E} y_i & \mathcal{T} := \sum_{i=1}^{N_E} x_i y_i \\ \mathcal{U}_x := \sum_{i=1}^{N_E} x_i^2 & \mathcal{U}_y := \sum_{i=1}^{N_E} y_i^2 & \mathcal{Z} := \sum_{i=1}^{N_E} x_i^2 y_i^2 \\ \mathcal{V}_x := \sum_{i=1}^{N_E} x_i y_i^2 & \mathcal{V}_y := \sum_{i=1}^{N_E} x_i^2 y_i. & \end{array} \tag{3.4.13}$$

The polynomials are

$$\begin{aligned} q_{\mathbf{G}}(\lambda) &= \left(g_0\,\lambda^4 - g_1\,\lambda^3 + g_2\,\lambda^2 - g_3\,\lambda + g_4\right)\lambda^{N-4}, \\ \\ g_0 &= 1 \\ g_1 &= \mathcal{T} - \mathcal{Z} + w_0^2\,p(1-p)N_I \\ g_2 &= \mathcal{V}_x\mathcal{V}_y - \mathcal{T}\mathcal{Z} + w_0^2\,p(1-p)N_I\,(\mathcal{T} - \mathcal{Z} - p(1-p)N_E) \\ g_3 &= w_0^2\,p(1-p)N_I\,\{\mathcal{V}_x\mathcal{V}_y - \mathcal{T}\mathcal{Z} + p(1-p)\,[N_E(\mathcal{Z}-\mathcal{T}) + \mathcal{S}_x\mathcal{S}_y - \mathcal{U}_x\mathcal{U}_y]\} \\ g_4 &= w_0^2\,p^2(1-p)^2N_I\,[N_E(\mathcal{T}\mathcal{Z} - \mathcal{V}_x\mathcal{V}_y) - \mathcal{S}_x\mathcal{S}_y\mathcal{Z} + \mathcal{S}_y\mathcal{U}_x\mathcal{V}_x + \mathcal{S}_x\mathcal{U}_y\mathcal{V}_y - \mathcal{U}_x\mathcal{U}_y\mathcal{T}] \end{aligned} \tag{3.4.14}$$

and

$$\begin{aligned} q_{\mathbf{M}}(\lambda) &= \left(m_0\,\lambda^3 - m_1\,\lambda^2 + m_2\,\lambda - m_3\right)\lambda^{N-3}, \\ \\ m_0 &= 1 \\ m_1 &= \mathcal{T} - w_0\,p\,N_I \\ m_2 &= w_0\,p\,N_I\,(pN_E - \mathcal{T}) \\ m_3 &= w_0\,p^2N_I\,(N_E\mathcal{T} - \mathcal{S}_x\mathcal{S}_y)\,. \end{aligned} \tag{3.4.15}$$

As these expressions directly show, the rank of $\mathbf{G}$ is ≤ 4 and the rank of $\mathbf{M}$ is ≤ 3 for any choice of the model parameters. These expressions are nevertheless not very useful unless the auxiliary degree sequences are previously known.

We are interested in statistically studying the spectrum of random connectivity matrices with a given degree distribution *before* these sequences are specified. This means we have to move a step backwards and suppose that the only information we have about the degree sequences is that each pair of auxiliary degrees appears independently according to a prescribed joint degree distribution. In this scenario, the empirical statistics become random variables. Since all of them are defined through sums over the entire degree sequences, the deviation from their expected values is negligible in the large N limit (as a result of the Central Limit Theorem). Thus, in the

limit of large N it is reasonable to substitute the empirical statistics defined in (3.4.13) by their expectations.

These expected values of course will depend on the chosen degree distribution. As an example, we provide here the results for degrees coming from Gamma(κ, θ) distributions with correlation coefficient ρ. As shown in Materials and Methods, the dependency of the empirical statistics on κ, θ and ρ is as follows:

$$\begin{aligned}
\langle \mathcal{S}_x \rangle = \langle \mathcal{S}_y \rangle &= \sqrt{N_E \kappa \theta} \\
\langle \mathcal{T} \rangle &= (\rho + \kappa)\,\theta \\
\langle \mathcal{U}_x \rangle = \langle \mathcal{U}_y \rangle &= (\kappa + 1)\,\theta \\
\langle \mathcal{V}_x \rangle = \langle \mathcal{V}_y \rangle &= \frac{1}{\sqrt{N_E \kappa \theta}}\,\theta^2(\kappa+1)(\kappa+2\rho) \\
\langle \mathcal{Z} \rangle &= \frac{1}{N_E \kappa}\,\theta^2\left[6\rho + \kappa(1 + 2\kappa + \kappa^2 + 8\rho + 4\kappa\rho + 2\rho^2)\right].
\end{aligned} \tag{3.4.16}$$

Behavior in the large N limit

It is important to bear in mind that for the theory to work we need an appropriate scaling of the parameters with the system's size N. In particular, we assume that $p = \mathcal{O}\left(\frac{1}{N}\right)$ and that the auxiliary degree distribution is held constant. The consequence of this scaling is that the coefficients of the characteristic polynomials of $\mathbf{G}$ and $\mathbf{M}$ are $\mathcal{O}(1)$. It is therefore appropriate to rewrite these coefficients neglecting terms of order smaller than 1. To do so, we first notice that from (3.4.16) we have

$$\begin{aligned}
\langle \mathcal{S}_x \rangle = \langle \mathcal{S}_y \rangle &= \mathcal{O}\left(\sqrt{N}\right) \\
\langle \mathcal{T} \rangle &= \mathcal{O}(1) \\
\langle \mathcal{U}_x \rangle = \langle \mathcal{U}_y \rangle &= \mathcal{O}(1) \\
\langle \mathcal{V}_x \rangle = \langle \mathcal{V}_y \rangle &= \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) \\
\langle \mathcal{Z} \rangle &= \mathcal{O}\left(\frac{1}{N}\right).
\end{aligned} \tag{3.4.17}$$

The polynomial $q_{\mathbf{M}}$ already has all the coefficients composed of terms of $\mathcal{O}(1)$. If we go back to the other polynomial, we get

$$\begin{aligned}
g_1 &\approx \mathcal{T} + w_0^2\, p N_I \\
g_2 &\approx w_0^2\, p N_I\,(\mathcal{T} - p N_E) \\
g_3 &\approx w_0^2\, p^2 N_I\,(\mathcal{S}_x \mathcal{S}_y - N_E \mathcal{T}) \\
g_4 &\approx 0,
\end{aligned} \tag{3.4.18}$$

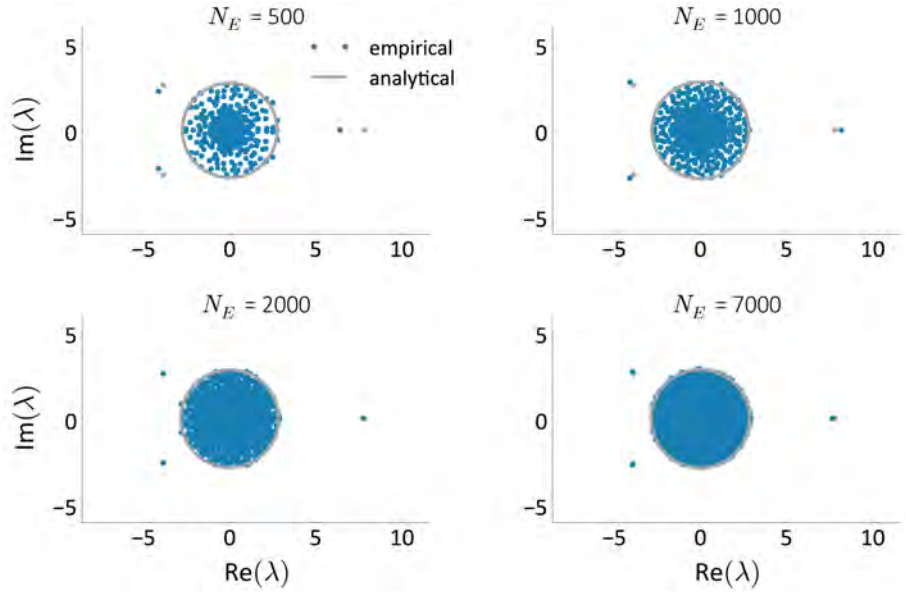

Figure 3.2: Spectra of random connectivity matrices with a broad degree distribution within the EE subnetwork. The plots only differ in the size of the network (and in the parameter p, which scales inversely with it). The elements of the matrices are defined by Eq. (3.4.6). Every plot shows the analytical prediction of the spectrum (with the predicted disk shown in grey and the outliers in orange) and the result of empirically computing the spectrum of a single realization of the matrix (blue). In all the cases $N_I = \frac{N_E}{4}$, $p = \frac{50}{N_E+N_I}$, $J_E = 0.11$, $w_0 = 8$. The EE degrees come from Gamma(κ, θ) distributions, $\kappa = 0.8$, $\theta = \frac{pN_E}{\kappa}$, with correlation coefficient $\rho = 0.8$.

which means that in the large N limit, $q_{\mathbf{G}}$ is in fact a polynomial of degree 3. This also shows that in the limit of a large network the only empirical statistics that matter are the ones given by $\mathcal{S}_x$, $\mathcal{S}_y$ and $\mathcal{T}$, which are related to the average degree and to the covariance between the degrees of individual neurons, respectively.

3.4.4 Comparison between theory and computer simulations

We next studied if the previous results really provide good approximations of the spectral properties of our random matrices. We simulated networks where the EE auxiliary degrees are generated from Gamma(κ, θ) distributions. The joint degree distribution was held independent of N and p was inversely proportional to N so as to have the appropriate scaling. We computed the largest positive solution to $q_{\mathbf{G}}(\lambda) = 0$ (recall that since $\mathbf{G}$ is a matrix of positive entries, it has a positive eigenvalue which gives its spectral radius) and the three non-trivial solutions to $q_{\mathbf{M}}(\lambda) = 0$. From the former we obtain the radius of the disk and from the latter, the outliers. We used the expressions for $q_{\mathbf{G}}(\lambda)$ and $q_{\mathbf{M}}(\lambda)$ given by Eqs. (3.4.14) and (3.4.15) where the empirical statistics are replaced by their average values, defined by (3.4.16). Figure 3.2 shows the spectrum of networks with different size N and fixed joint degree distribution. It is clear that the theory predicts well

the properties of the spectrum and that the prediction becomes more accurate as N increases. In these particular examples, the spectrum is composed of a highly dense bulk of eigenvalues inside a disk centered at zero with radius $R_{\mathbf{J}}$ and three outliers.

The formalism presented here can help to answer questions like how the spectrum changes as we vary the degree correlation within the EE network. To address this issue, we simulated networks with the same size and the same marginal degree distribution which nevertheless differ their correlation coefficient ρ. The results show that whereas the radius of the bulk is almost independent of ρ, the positive outlier is very sensitive to this parameter (Fig. 3.3 A, C). In particular, the higher ρ, the larger the outlier. This finding could help to explain dynamical differences in systems whose linear approximation around a fixed point is defined by matrices as the ones described here. From these results we expect that the appearance of a large positive eigenvalue as ρ exceeds a critical value will be responsible for destabilizing the dynamics and, therefore, for giving rise to qualitatively different dynamical states as compared to cases with lower ρ.

A similar effect is found when the synaptic strengths of the excitatory connections are multiplied by a common factor f. Figure 3.3 shows the spectrum of networks with $\rho = 0$ in which the following parameter modification has been done: $J_E \to fJ_E$, $w_0 \to w_0/f$. This ensures that the strengths of E connections are multiplied by f whereas that of I connections are kept constant as we vary f. The modulation of f produces changes both in the radius of the bulk and in the positive outlier (Fig. 3.3 B, D), although the salient consequence is again the displacement of the outlier "to the right". This indicates that, even though the precise effect of varying the correlation coefficient ρ and increasing the coupling of excitatory synapses is not the same, both modifications of the structure lead to qualitatively similar outcomes, as has also been sugested by Nykamp et al. [2017] in their study of rate networks with arbitrary degree distribution. Notice, on the contrary, that a global increase of the synaptic couplings exerts a multiplicative effect both in the bulk and in the outliers, which can be responsible for a sudden appearance of almost infinitely many eigenvalues with positive real part, whereas the increase of ρ or the excitatory weights alone appear as mechanisms to almost selectively increase a single positive eigenvalue. Therefore, the dynamical consequences of both scenarios can be quite distinct.

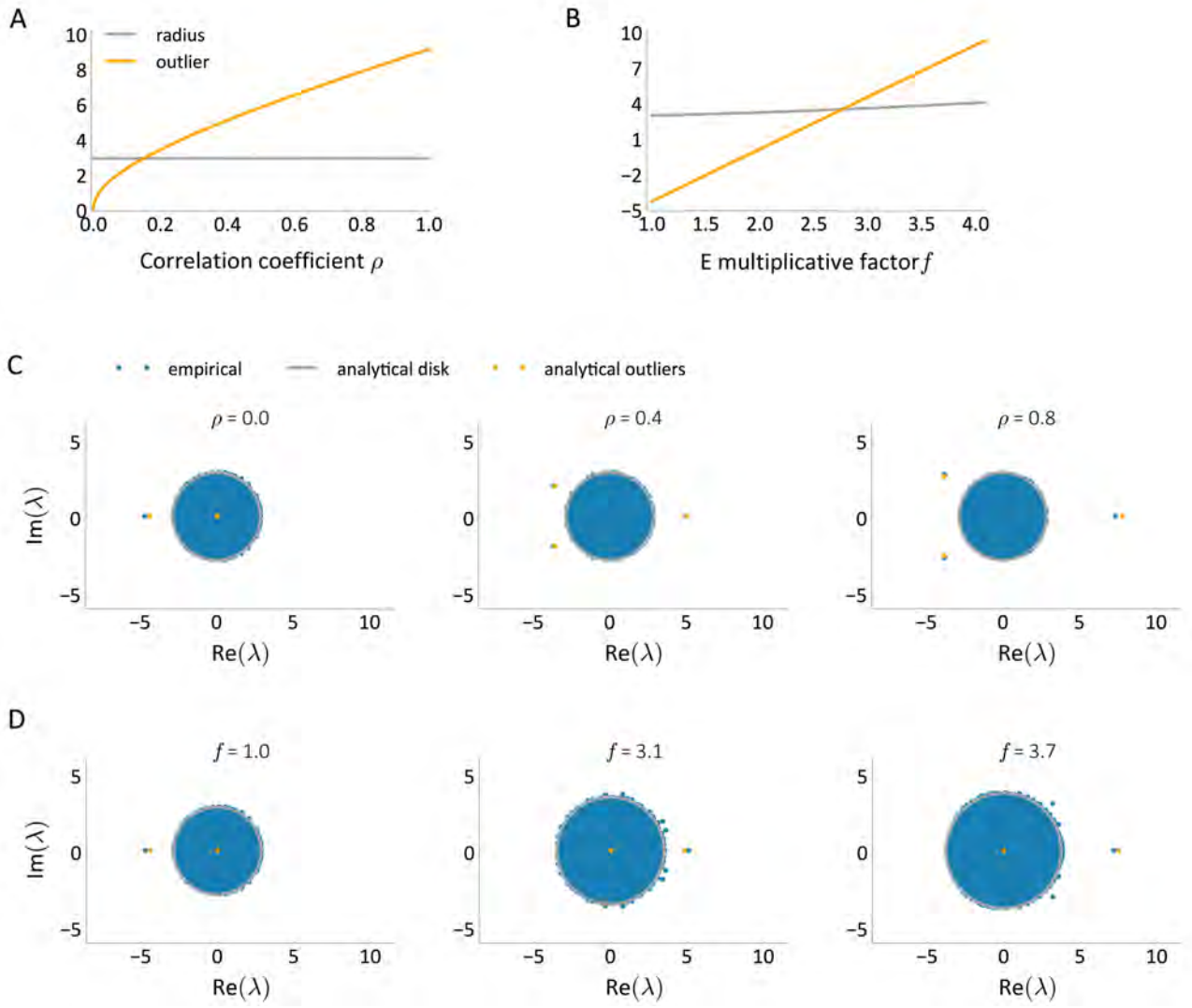

Figure 3.3: Spectra of random connectivity matrices with a broad degree distribution within the EE subnetwork as we vary the degree correlation ρ or the excitatory multiplicative factor f. **A** Radius of the disk (grey) and maximal positive outlier (orange) predicted analytically as a function of ρ when $f = 1$. **B** Same as A for $\rho = 0$ and variable f. **C** Three examples of the whole spectrum for networks with $f = 1$ and which only differ in the parameter ρ. Every plot shows the analytical prediction of the spectrum (with the predicted disk shown in grey and the outliers in orange) and the result of empirically computing the spectrum of a single realization of the matrix (blue). **D** Three examples of the whole spectrum for networks with $\rho = 0$ and which only differ in the parameter f. In all the plots, $N_E = 5000$, $N_I = \frac{N_E}{4}$, $p = \frac{50}{N_E+N_I}$, $J_E = 0.11f$, $w_0 = 8/f$. The EE degrees come from Gamma(κ, θ) distributions, $\kappa = 0.8$, $\theta = \frac{pN_E}{\kappa}$.

3.5 Conclusion

We have described an analytical method for predicting the spectral properties of highly heterogeneous random connectivity matrices, where the degrees within the excitatory population follow a prescribed distribution of in- and out-degrees. Previous work had provided the means of predicting the spectra of random matrices with regular structure [Sompolinsky et al., 1988; Tao, 2013; Aljadeff et al., 2015a,b]. Our goal here was to explore if those techniques can be applied to even more general classes of matrices.

Some of this previous work [Sompolinsky et al., 1988; Aljadeff et al., 2015a,b] addressed the problem of inferring the spectral properties of random connectivity matrices in order to study firing rate models of neural networks. But in the matrices considered in such studies the entries have zero mean and their variances can only take a finite number of possible values, which imposes important limitations to the study of more realistic scenarios. On the one hand, according to Dale's law, one of the paradigms of current neural science, neurons exert either an excitatory or an inhibitory effect to their targets. Thus, a random connectivity matrix whose entries are independent and have zero mean will contradict, with high probability, Dale's law. On the other hand, limiting the number of possible variances excludes networks where connectivity rules might vary continuously across neuronal pairs.

Rajan and Abbott have analytically computed the spectrum of random connectivity matrices with a separation of excitation and inhibition under balance conditions (that is, when the sum of all the synaptic weights to a neuron is zero on average) and assuming that the variances take two values (one for each type of synapse) [Rajan and Abbott, 2006]. Here we have analyzed matrices of excitatory and inhibitory neurons with continuous modulation of variances.

We have found a way to apply the results described in this previous work to our connectivity matrices. To do so, we first assume that the degree sequences are fixed and that connections appear randomly but responding to a probability profile dictated by these degrees. This is what allows us to interpret our matrix as a random matrix with independent entries. Then we let the degrees be random variables from a prescribed distribution, thus introducing a second level of stochasticity which transforms the predicted quantities into random variables as well. We have shown that these relevant variables are in fact empirical moments of the degree distribution, so they can be substituted by the real moments in the limit of large networks.

To apply the previously known results for random matrices, we split our connectivity matrix $\mathbf{J}$

into a sum of a matrix of means $\mathbf{M}$ and a random matrix $\mathbf{Y}$ whose entries have zero mean. Our results suggest that the spectrum of $\mathbf{Y}$ is circularly symmetric, as is the spectrum of random block matrices [Aljadeff et al., 2015b,a]. Therefore, the known results for random block matrices can be extended to matrices with an infinite number of blocks. Although the applicability of this result can be justified when the matrix variances maintain certain regularities [Aljadeff et al., 2016], in our heterogeneous networks such regularities cannot be ensured in general. Nevertheless, the success of the theory in predicting the spectra of these matrices (as shown by our computer simulations) suggests that such regularity requirements could be relaxed even more.

We have proven that $\mathbf{M}$ is a low-rank matrix regardless of the model parameters. We therefore postulated that the effects that low-rank perturbation matrices exert on the spectrum of random matrices could hold in our case too. Such perturbations essentially add a finite number of outliers to the bulk of the circular spectrum, and these outliers are well approximated by the non-trivial eigenvalues of the perturbation matrix [Tao, 2013]. Direct comparisons with simulations indicate that this is true for our matrices as well.

In summary, the results presented in this chapter show, first, that previous mathematical formalisms on the spectrum of random matrices are extendible to more general classes of matrices. On the other hand, we have been able to provide analytical formulas for such spectral properties, which allow for a computationally non-expensive way to calculate them from the set of model parameters. A thorough proof of why those results hold in general is still needed, and it should be the object of further research.

We applied the derived predictions to the study of connectivity matrices where the degrees within the excitatory subnetwork are possibly correlated. The results indicate that, under the studied circumstances, the radius of the spectral bulk is almost insensitive to the correlation coefficient ρ, whereas the outliers vary dramatically with it. In particular, there is a positive outlier which becomes larger as ρ increases. This finding could explain the destabilization of dynamics induced by large degree correlations in appropriate firing rate models. The effect of increasing ρ in these networks is similar to that of multiplying the excitatory synaptic weights by a constant factor f, but different from a general multiplicative increase of all the weights, which can induce the appearance of many eigenvalues with real part above the critical value. We therefore anticipate that these two types of perturbation of the connectivity structure can have very different signatures on neuronal dynamics.

The fact that a single positive outlier can appear both when in- and out-degrees are correlated

within the EE subnetwork and when the excitatory synaptic weights are strengthened suggests that the dynamical state of the system can be perturbed in a similar way by means of two strategies which have very different biological meanings. Increasing the strength of synapses is translated into using more resources (receptors, neurotransmitters, etc.) at every excitatory synapse, whereas introducing correlations between in- and out-degrees involves a rearrangement of the synapses, without an extra investment in chemical resources. Therefore, topologies which exhibit positive degree correlations could have been favored by natural selection due to their efficiency at enhancing the responsiveness of the network without additional biochemical costs.

3.6 Materials and Methods

3.6.1 Computation of the characteristic polynomials of G and M

Here we compute the spectral radius of the matrix of variances $\mathbf{G}$ and the spectrum of the matrix of means $\mathbf{M}$ using the following property: given an $N \times N$ matrix $\mathbf{A}$, the characteristic polynomial of $\mathbf{A}$ is

$$q_{\mathbf{A}}(\lambda) \;=\; \sum_{k=0}^{N} (-1)^k a_k \, \lambda^{N-k}, \tag{3.6.1}$$

where $a_0 = 1$ and a_k is the sum of the k-rowed diagonal minors of $\mathbf{A}$.

We first assume that the auxiliary degree sequence is fixed. This will allow us to give an analytical expression for the characteristic polynomials of $\mathbf{G}$ and $\mathbf{M}$ as a function of this sequence. As we will see, the effect of the sequence on the coefficients of the polynomials goes through some "empirical statistics" on the sequence. These statistics are:

$$\begin{array}{lll} \mathcal{S}_x := \sum_{i=1}^{N_E} x_i & \mathcal{S}_y := \sum_{i=1}^{N_E} y_i & \mathcal{T} := \sum_{i=1}^{N_E} x_i y_i \\ \mathcal{U}_x := \sum_{i=1}^{N_E} x_i^2 & \mathcal{U}_y := \sum_{i=1}^{N_E} y_i^2 & \mathcal{Z} := \sum_{i=1}^{N_E} x_i^2 y_i^2 \\ \mathcal{V}_x := \sum_{i=1}^{N_E} x_i y_i^2 & \mathcal{V}_y := \sum_{i=1}^{N_E} x_i^2 y_i. & \end{array} \tag{3.6.2}$$

The next step is to consider that the auxiliary degrees are not fixed but are independent realizations of a common random vector $(\tilde{K}^{\text{in}}, \tilde{K}^{\text{out}})$, which transforms the empirical statistics into the real statistics in the large N limit.

To simplify the computations we will use the notation $v = p(1-p)$, $w = w_0^2 p(1-p)$.

Characteristic polynomial of G

Let $q_{\mathbf{G}}$ be the characteristic polynomial of $\mathbf{G}$. Using the previous property,

$$q_{\mathbf{G}}(\lambda) \;=\; \sum_{k=0}^{N} (-1)^k g_k \, \lambda^{N-k}, \tag{3.6.3}$$

where $g_0 = 1$ and g_k is the sum of the k-rowed diagonal minors of $\mathbf{G}$. We will compute $g_1, \cdots, g_4$ explicitly and show that $g_k = 0$ for $k \geq 5$.

Recall that $\mathbf{G}$ has the form

$$\mathbf{G} \;=\; \begin{pmatrix} x_1y_1(1-x_1y_1) & \cdots & x_1y_{N_E}(1-x_1y_{N_E}) & w & \cdots & w \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x_{N_E}y_1(1-x_{N_E}y_1) & \cdots & x_{N_E}y_{N_E}(1-x_{N_E}y_{N_E}) & w & \cdots & w \\ v & \cdots & v & w & \cdots & w \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ v & \cdots & v & w & \cdots & w \end{pmatrix}. \tag{3.6.4}$$

We denote by $\mathbf{G}_{i_1 \cdots i_k, m}$ the $(k+m) \times (k+m)$ matrix associated to the diagonal minor of $\mathbf{G}$ that results from selecting the first $i_1, \cdots, i_k$ diagonal elements of $\mathbf{G}$ and m elements among the last N_I diagonal elements of $\mathbf{G}$. From this definition,

$$\det \mathbf{G}_{i_1 \cdots i_k, m} = 0 \qquad \text{for } m \geq 2 \tag{3.6.5}$$

because $\mathbf{G}_{i_1 \cdots i_k, m}$ has, at least, two repeated columns. We also have

$$\begin{aligned} \mathbf{G}_{ijk,0} \;&=\; \begin{pmatrix} x_iy_i(1-x_iy_i) & x_iy_j(1-x_iy_j) & x_iy_k(1-x_iy_k) \\ x_jy_i(1-x_jy_i) & x_jy_j(1-x_jy_j) & x_jy_k(1-x_jy_k) \\ x_ky_i(1-x_ky_i) & x_ky_j(1-x_ky_j) & x_ky_k(1-x_ky_k) \end{pmatrix} \\ &=\; \operatorname{diag}(x_i, x_j, x_k) \begin{pmatrix} 1-x_iy_i & 1-x_iy_j & 1-x_iy_k \\ 1-x_jy_i & 1-x_jy_j & 1-x_jy_k \\ 1-x_ky_i & 1-x_ky_j & 1-x_ky_k \end{pmatrix} \operatorname{diag}(y_i, y_j, y_k), \end{aligned} \tag{3.6.6}$$

where $\operatorname{diag}(a, b, c)$ denotes the diagonal matrix with diagonal elements (a, b, c). The determinant of the matrix in the middle is always zero, so $\det \mathbf{G}_{ijk,0} = 0$. This implies

$$\det \mathbf{G}_{i_1 \cdots i_k, 0} = 0 \qquad \text{for } k \geq 3. \tag{3.6.7}$$

Coefficient g_1

The sum of the 1-rowed diagonal minors of $\mathbf{G}$ is just the trace of $\mathbf{G}$, therefore

$$\begin{aligned} g_1 &= \mathrm{Tr}(\mathbf{G}) \\ &= \sum_{i=1}^{N_E} x_i y_i (1 - x_i y_i) + \sum_{i=1}^{N_I} w \\ &= \mathcal{T} - \mathcal{Z} + w N_I . \end{aligned} \tag{3.6.8}$$

Coefficient g_2

We have to sum all the 2-rowed diagonal minors of $\mathbf{G}$. There are three types of such minors:

- First type

$$\det \mathbf{G}_{ij,0} = \begin{vmatrix} x_i y_i (1 - x_i y_i) & x_i y_j (1 - x_i y_j) \\ x_j y_i (1 - x_j y_i) & x_j y_j (1 - x_j y_j) \end{vmatrix} = x_i x_j \, y_i y_j \, (-x_i y_i + x_j y_i + x_i y_j - x_j y_j) \tag{3.6.9}$$

  and the sum of all the determinants of this type is

$$\begin{aligned} S_{2,0} &= \sum_{i=1}^{N_E} \sum_{j=i+1}^{N_E} \det \mathbf{G}_{ij,0} \\ &= \frac{1}{2} \left(\sum_{i=1}^{N_E} \sum_{j=1}^{N_E} \det \mathbf{G}_{ij,0} - \sum_{i=1}^{N_E} \det \mathbf{G}_{ii,0} \right) \\ &= \frac{1}{2} \sum_{i=1}^{N_E} \sum_{j=1}^{N_E} x_i x_j \, y_i y_j \, (-x_i y_i + x_j y_i + x_i y_j - x_j y_j) \\ &= \mathcal{V}_x \mathcal{V}_y - \mathcal{T} \mathcal{Z} . \end{aligned} \tag{3.6.10}$$

- Second type

$$\det \mathbf{G}_{i,1} = \begin{vmatrix} x_i y_i (1 - x_i y_i) & w \\ v & w \end{vmatrix} = w \left[x_i y_i (1 - x_i y_i) - v \right] \tag{3.6.11}$$

  and the sum of all the determinants of this type is

$$\begin{aligned} S_{1,1} &= \sum_{i=1}^{N_E} \sum_{j=1}^{N_I} \det \mathbf{G}_{i,1} \\ &= w N_I \sum_{i=1}^{N_E} \left[x_i y_i (1 - x_i y_i) - v \right] \\ &= w N_I \left(\mathcal{T} - \mathcal{Z} - v N_E \right) . \end{aligned} \tag{3.6.12}$$

- Third type

$$\det \mathbf{G}_{0,2} \;=\; 0 \tag{3.6.13}$$

(property (3.6.5)).

We obtain

$$g_2 \;=\; \mathcal{V}_x\mathcal{V}_y - \mathcal{T}\mathcal{Z} + wN_I\left(\mathcal{T} - \mathcal{Z} - vN_E\right). \tag{3.6.14}$$

Coefficient g_3

g_3 is the sum all the 3-rowed diagonal minors of $\mathbf{G}$. There are four types of such minors:

- First type

$$\det \mathbf{G}_{ijk,0} = 0 \tag{3.6.15}$$

(see (3.6.7)).

- Second type

$$\begin{aligned}
\det \mathbf{G}_{ij,1} \;&=\; \begin{vmatrix} x_iy_i(1-x_iy_i) & x_iy_j(1-x_iy_j) & w \\ x_jy_i(1-x_jy_i) & x_jy_j(1-x_jy_j) & w \\ v & v & w \end{vmatrix} \\
&=\; w\det \mathbf{G}_{ij,0} \\
&\quad +vw\left[-x_iy_i(1-x_iy_i) + x_jy_i(1-x_jy_i) + x_iy_j(1-x_iy_j) - x_jy_j(1-x_jy_j)\right].
\end{aligned} \tag{3.6.16}$$

The sum of all the determinants of this kind gives

$$\begin{aligned}
S_{2,1} \;&=\; \sum_{i=1}^{N_E}\sum_{j=i+1}^{N_E}\sum_{k=1}^{N_I} \det \mathbf{G}_{ij,1} \\
&=\; wN_IS_{2,0} \\
&\quad +vwN_I\sum_{i=1}^{N_E}\sum_{j=i+1}^{N_E}\left[-x_iy_i(1-x_iy_i) + x_jy_i(1-x_jy_i) + x_iy_j(1-x_iy_j) - x_jy_j(1-x_jy_j)\right] \\
&=\; wN_I\left\{\mathcal{V}_x\mathcal{V}_y - \mathcal{T}\mathcal{Z} + v\left[N_E(\mathcal{Z}-\mathcal{T}) + \mathcal{S}_x\mathcal{S}_y - \mathcal{U}_x\mathcal{U}_y\right]\right\}.
\end{aligned} \tag{3.6.17}$$

- Third and fourth types

$$\begin{aligned}
\det \mathbf{G}_{i,2} \;&=\; 0 \\
\det \mathbf{G}_{0,3} \;&=\; 0
\end{aligned} \tag{3.6.18}$$

(property (3.6.5)).

We obtain

$$g_3 \;=\; wN_I\left\{\mathcal{V}_x\mathcal{V}_y - \mathcal{T}\mathcal{Z} + v\left[N_E(\mathcal{Z}-\mathcal{T}) + \mathcal{S}_x\mathcal{S}_y - \mathcal{U}_x\mathcal{U}_y\right]\right\}. \tag{3.6.19}$$

Coefficient g_4

g_4 is the sum all the 4-rowed diagonal minors of $\mathbf{G}$. There are five types of such minors:

- First type

$$\det \mathbf{G}_{ijkl,0} \;=\; 0 \tag{3.6.20}$$

  (property (3.6.7)).

- Second type

$$\begin{aligned}\det \mathbf{G}_{ijk,1} \;&=\; \begin{vmatrix} x_iy_i(1-x_iy_i) & x_iy_j(1-x_iy_j) & x_iy_k(1-x_iy_k) & w \\ x_jy_i(1-x_jy_i) & x_jy_j(1-x_jy_j) & x_jy_k(1-x_jy_k) & w \\ x_ky_i(1-x_ky_i) & x_ky_j(1-x_ky_j) & x_ky_k(1-x_ky_k) & w \\ v & v & v & w \end{vmatrix} \\ &=\; vw(x_i-x_j)(x_i-x_k)(x_j-x_k)(y_i-y_j)(y_i-y_k)(y_j-y_k).\end{aligned} \tag{3.6.21}$$

  The sum of all the minors of this type is

$$\begin{aligned} S_{3,1} \;&=\; \sum_{i=1}^{N_E}\sum_{j=i+1}^{N_E}\sum_{k=j+1}^{N_E}\sum_{l=1}^{N_I} vw(x_i-x_j)(x_i-x_k)(x_j-x_k)(y_i-y_j)(y_i-y_k)(y_j-y_k) \\ &=\; \tfrac{1}{6}vwN_I\sum_{i=1}^{N_E}\sum_{j=1}^{N_E}\sum_{k=1}^{N_E}\begin{bmatrix} (x_i^2x_j - x_ix_j^2 - x_i^2x_k + x_j^2x_k + x_ix_k^2 - x_jx_k^2) \\ \times(y_i^2y_j - y_iy_j^2 - y_i^2y_k + y_j^2y_k + y_iy_k^2 - y_jy_k^2)\end{bmatrix} \\ &=\; vwN_I\left[N_E(\mathcal{T}\mathcal{Z}-\mathcal{V}_x\mathcal{V}_y) - \mathcal{S}_x\mathcal{S}_y\mathcal{Z} + \mathcal{S}_y\mathcal{U}_x\mathcal{V}_x + \mathcal{S}_x\mathcal{U}_y\mathcal{V}_y - \mathcal{U}_x\mathcal{U}_y\mathcal{T}\right]. \end{aligned} \tag{3.6.22}$$

- Third, fourth and fifth types

$$\begin{aligned}\det \mathbf{G}_{ij,2} \;&=\; 0 \\ \det \mathbf{G}_{i,3} \;&=\; 0 \\ \det \mathbf{G}_{0,4} \;&=\; 0\end{aligned} \tag{3.6.23}$$

  because of property (3.6.5).

We conclude that

$$g_4 \;=\; vwN_I\left[N_E(\mathcal{T}\mathcal{Z}-\mathcal{V}_x\mathcal{V}_y) - \mathcal{S}_x\mathcal{S}_y\mathcal{Z} + \mathcal{S}_y\mathcal{U}_x\mathcal{V}_x + \mathcal{S}_x\mathcal{U}_y\mathcal{V}_y - \mathcal{U}_x\mathcal{U}_y\mathcal{T}\right]. \tag{3.6.24}$$

Coefficient g_n for $n \geq 5$

The coefficient g_n for $n \geq 5$ is the sum of all the n-rowed diagonal minors of $\mathbf{G}$. These minors have the form $\det \mathbf{G}_{i_1 \cdots i_k, m}$ with $m + k = n$.

- If $m = 0$, then $k = n \geq 5$ so the minor is zero by virtue of property (3.6.7).
- If $m = 1$, then $k = n - 1 \geq 4$. This minor has the form

$$\det \mathbf{G}_{i_1 \cdots i_k, 1} = \begin{vmatrix} x_{i_1} y_{i_1}(1 - x_{i_1} y_{i_1}) & \cdots & x_{i_1} y_{i_k}(1 - x_{i_1} y_{i_k}) & w \\ \vdots & \ddots & \vdots & \vdots \\ x_{i_k} y_{i_1}(1 - x_{i_k} y_{i_1}) & \cdots & x_{i_k} y_{i_k}(1 - x_{i_k} y_{i_k}) & w \\ v & \cdots & v & w \end{vmatrix}. \tag{3.6.25}$$

  If we develop the determinant around the last column, we obtain a weighted sum of $\det \mathbf{G}_{i_1 \cdots i_k, 0} = 0$ and determinants of the form

$$\begin{vmatrix} x_{i_{s_1}} y_{i_1}(1 - x_{i_{s_1}} y_{i_1}) & \cdots & x_{i_{s_1}} y_{i_k}(1 - x_{i_{s_1}} y_{i_k}) \\ \vdots & \ddots & \vdots \\ x_{i_{s_{k-1}}} y_{i_1}(1 - x_{i_{s_{k-1}}} y_{i_1}) & \cdots & x_{i_{s_{k-1}}} y_{i_k}(1 - x_{i_{s_{k-1}}} y_{i_k}) \\ v & \cdots & v \end{vmatrix}. \tag{3.6.26}$$

  The last determinant can be developed around the last row and this produces a weighted sum of determinants of the form $\det \mathbf{G}_{j_1 \cdots j_{k-1}, 0}$, which are all zero because $k - 1 \geq 3$ (property (3.6.7)).

- If $m \geq 2$, property (3.6.5) ensures that the minor is zero.

We conclude that

$$g_n = 0 \qquad \text{for } n \geq 5. \tag{3.6.27}$$

Summarizing, the characteristic polynomial of the matrix $\mathbf{G}$ is

$$q_{\mathbf{G}}(\lambda) = \left(g_0\,\lambda^4 - g_1\,\lambda^3 + g_2\,\lambda^2 - g_3\,\lambda + g_4\right)\lambda^{N-4},$$

$$\begin{aligned} g_0 &= 1 \\ g_1 &= \mathcal{T} - \mathcal{Z} + wN_I \\ g_2 &= \mathcal{V}_x\mathcal{V}_y - \mathcal{T}\mathcal{Z} + wN_I\,(\mathcal{T} - \mathcal{Z} - vN_E) \\ g_3 &= wN_I\,\{\mathcal{V}_x\mathcal{V}_y - \mathcal{T}\mathcal{Z} + v\,[N_E(\mathcal{Z} - \mathcal{T}) + \mathcal{S}_x\mathcal{S}_y - \mathcal{U}_x\mathcal{U}_y]\} \\ g_4 &= vwN_I\,[N_E(\mathcal{T}\mathcal{Z} - \mathcal{V}_x\mathcal{V}_y) - \mathcal{S}_x\mathcal{S}_y\mathcal{Z} + \mathcal{S}_y\mathcal{U}_x\mathcal{V}_x + \mathcal{S}_x\mathcal{U}_y\mathcal{V}_y - \mathcal{U}_x\mathcal{U}_y\mathcal{T}]\,. \end{aligned} \tag{3.6.28}$$

Frobenius norm of M

The matrix $\mathbf{M}$ has the form

$$\mathbf{M} = \begin{pmatrix} x_1y_1 & \cdots & x_1y_{N_E} & -w_0p & \cdots & -w_0p \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x_{N_E}y_1 & \cdots & x_{N_E}y_{N_E} & -w_0p & \cdots & -w_0p \\ p & \cdots & p & -w_0p & \cdots & -w_0p \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ p & \cdots & p & -w_0p & \cdots & -w_0p \end{pmatrix}. \tag{3.6.29}$$

The Frobenius norm of $\mathbf{M} = (M_{ij})_{i,j}$ is

$$\|\mathbf{M}\|_F := \sqrt{\sum_{i,j=1}^{N_E+N_I} |M_{ij}|^2}.$$

We have

$$\begin{aligned} \sum_{i,j=1}^{N_E+N_I} |M_{ij}|^2 &= \sum_{i,j=1}^{N_E} |M_{ij}|^2 + N_EN_Ip^2 + (N_E+N_I)N_Iw_0^2p^2 \\ &= \left(\sum_{i=1}^{N_E} x_i^2\right)\left(\sum_{j=1}^{N_E} y_j^2\right) + N_EN_Ip^2 + (N_E+N_I)N_Iw_0^2p^2 \\ &= \mathcal{O}(1) \end{aligned} \tag{3.6.30}$$

because, by construction, $x_i^2 = \dfrac{\left(\tilde{K}_i^{\text{in}}\right)^2}{N_E\langle\tilde{K}\rangle} = \mathcal{O}\left(\dfrac{1}{N}\right)$, $y_i = \mathcal{O}\left(\frac{1}{N}\right)$ and $p = \mathcal{O}\left(\frac{1}{N}\right)$. We conclude that

$$\|\mathbf{M}\|_F = \mathcal{O}(1). \tag{3.6.31}$$

Since the operator norm $\|\cdot\|_2$ induced by the euclidean norm satisfies

$$\|\mathbf{M}\|_2 \quad \leq \quad \|\mathbf{M}\|_F \tag{3.6.32}$$

and all the vector norms are equivalent, any operator norm of $\mathbf{M}$ is $\mathcal{O}(1)$ too.

Characteristic polynomial of M

Let $q_{\mathbf{M}}$ be the characteristic polynomial of $\mathbf{M}$. Again, this polynomial has the form

$$q_{\mathbf{M}}(\lambda) \quad = \quad \sum_{k=0}^{N} (-1)^k m_k \, \lambda^{N-k}, \tag{3.6.33}$$

where $m_0 = 1$ and m_k is the sum of the k-rowed diagonal minors of $\mathbf{M}$. We will compute $m_1, \cdots, m_3$ explicitly and show that $m_k = 0$ for $k \geq 4$.

As before, we denote by $\mathbf{M}_{i_1 \cdots i_k, s}$ the $(k+s) \times (k+s)$ matrix corresponding to the diagonal minor of $\mathbf{M}$ that results from selecting the first $i_1, \cdots, i_k$ diagonal elements of $\mathbf{M}$ and s elements among the last N_I diagonal elements of $\mathbf{M}$. We have

$$\det \mathbf{M}_{i_1 \cdots i_k, s} = 0 \qquad \text{for } s \geq 2 \tag{3.6.34}$$

and

$$\det \mathbf{M}_{ij,0} \quad = \quad \begin{vmatrix} x_i y_i & x_i y_j \\ x_j y_i & x_j y_j \end{vmatrix} \quad = \quad 0, \tag{3.6.35}$$

which implies

$$\det \mathbf{M}_{i_1 \cdots i_k, 0} = 0 \qquad \text{for } k \geq 2. \tag{3.6.36}$$

Coefficient m_1

m_1 is the trace of $\mathbf{M}$:

$$\begin{aligned} m_1 &= \mathrm{Tr}(\mathbf{M}) \\ &= \sum_{i=1}^{N_E} x_i y_i + \sum_{i=1}^{N_I} (-w_0 p) \\ &= \mathcal{T} - w_0 p N_I. \end{aligned} \tag{3.6.37}$$

Coefficient m_2

m_2 is the sum all the 2-rowed diagonal minors of $\mathbf{M}$. There are three types of such minors:

- First type

$$\det \mathbf{M}_{ij,0} = 0. \tag{3.6.38}$$

- Second type

$$\det \mathbf{M}_{i,1} = \begin{vmatrix} x_i y_i & -w_0 p \\ p & -w_0 p \end{vmatrix} = w_0 p\,(p - x_i y_i). \tag{3.6.39}$$

  The sum of all the minors of this type is

$$\begin{aligned} S_{1,1} &= \sum_{i=1}^{N_E} \sum_{j=1}^{N_I} \det \mathbf{M}_{i,1} \\ &= w_0 p\, N_I \sum_{i=1}^{N_E} (p - x_i y_i) \\ &= w_0 p\, N_I\, (p N_E - \mathcal{T})\,. \end{aligned} \tag{3.6.40}$$

- Third type

$$\det \mathbf{M}_{0,2} = 0. \tag{3.6.41}$$

We obtain

$$m_2 = w_0 p\, N_I\, (p N_E - \mathcal{T})\,. \tag{3.6.42}$$

Coefficient m_3

m_3 is the sum all the 3-rowed diagonal minors of $\mathbf{M}$. There are four types of such minors:

- First type

$$\det \mathbf{M}_{ijk,0} = 0. \tag{3.6.43}$$

- Second type

$$\det \mathbf{M}_{ij,1} = \begin{vmatrix} x_i y_i & x_i y_j & -w_0 p \\ x_j y_i & x_j y_j & -w_0 p \\ p & p & -w_0 p \end{vmatrix} = w_0 p^2 \left(x_i y_i - x_i y_j - x_j y_i + x_j y_j \right). \tag{3.6.44}$$

The sum of all the minors of this kind is

$$\begin{aligned} S_{2,1} &= \sum_{i=1}^{N_E} \sum_{j=i+1}^{N_E} \sum_{k=1}^{N_I} \det \mathbf{M}_{ij,1} \\ &= \tfrac{1}{2} w_0 \, p^2 N_I \sum_{i=1}^{N_E} \sum_{j=1}^{N_E} \left(x_i y_i - x_i y_j - x_j y_i + x_j y_j \right) \\ &= w_0 \, p^2 N_I \left(N_E \mathcal{T} - \mathcal{S}_x \mathcal{S}_y \right). \end{aligned} \tag{3.6.45}$$

- Third and fourth types

$$\begin{aligned} \det \mathbf{M}_{i,2} &= 0 \\ \det \mathbf{M}_{0,3} &= 0. \end{aligned} \tag{3.6.46}$$

We conclude that

$$m_3 = w_0 \, p^2 N_I \left(N_E \mathcal{T} - \mathcal{S}_x \mathcal{S}_y \right). \tag{3.6.47}$$

Coefficient m_n for $n \geq 4$

The coefficient m_n for $n \geq 4$ is the sum of all the n-rowed diagonal minors of $\mathbf{M}$. These minors have the form $\det \mathbf{M}_{i_1 \cdots i_k, s}$ with $m + s = n$.

- If $s = 0$, then $k = n \geq 4$ so the minor is zero by virtue of property (3.6.36).
- If $s = 1$, then $k = n - 1 \geq 3$. This minor has the form

$$\det \mathbf{M}_{i_1 \cdots i_k, 1} = \begin{vmatrix} x_{i_1} y_{i_1} & \cdots & x_{i_1} y_{i_k} & -w_0 p \\ \vdots & \ddots & \vdots & \vdots \\ x_{i_k} y_{i_1} & \cdots & x_{i_k} y_{i_k} & -w_0 p \\ p & \cdots & p & -w_0 p \end{vmatrix}. \tag{3.6.48}$$

If we develop the determinant around the last column, we obtain a weighted sum of $\det \mathbf{M}_{i_1 \cdots i_k, 0} = 0$

and determinants of the form

$$\begin{vmatrix} x_{i_{s_1}} y_{i_1} & \cdots & x_{i_{s_1}} y_{i_k} \\ \vdots & \ddots & \vdots \\ x_{i_{s_{k-1}}} y_{i_1} & \cdots & x_{i_{s_{k-1}}} y_{i_k} \\ p & \cdots & p \end{vmatrix}. \tag{3.6.49}$$

The last determinant can be developed around the last row and this produces a weighted sum of determinants of the form $\det \mathbf{M}_{j_1 \cdots j_{k-1},0}$, which are all zero because $k-1 \geq 2$ (property (3.6.36)).

- If $s \geq 2$, the minor is zero (property (3.6.34)).

We conclude that

$$m_n = 0 \qquad \text{for } n \geq 4. \tag{3.6.50}$$

Summarizing, the characteristic polynomial of the matrix $\mathbf{M}$ is

$$\begin{aligned} q_{\mathbf{M}}(\lambda) &= \left(m_0\,\lambda^3 - m_1\,\lambda^2 + m_2\,\lambda - m_3\right)\lambda^{N-3}, \\ \\ m_0 &= 1 \\ m_1 &= \mathcal{T} - w_0 p N_I \\ m_2 &= w_0 p\, N_I\,(pN_E - \mathcal{T}) \\ m_3 &= w_0\, p^2 N_I\,(N_E \mathcal{T} - \mathcal{S}_x \mathcal{S}_y)\,. \end{aligned} \tag{3.6.51}$$

3.6.2 Expected values of the empirical statistics when the degrees follow Gamma distributions

Here we compute the expectations of the empirical statistics assuming that the auxiliary degrees come from Gamma(κ, θ) $(\kappa, \theta > 0)$ distributions with correlation coefficient $\rho \in [0, 1]$. In particular, we construct each pair of auxiliary degrees from a trio of independent random variables:

$$X \sim \text{Gamma}(\kappa_1, \theta), \qquad Y \sim \text{Gamma}(\kappa_2, \theta), \qquad Z \sim \text{Gamma}(\kappa_2, \theta), \tag{3.6.52}$$

where $\kappa_1 = \kappa\rho$ and $\kappa_2 = \kappa - \kappa_1 = \kappa(1-\rho)$. The auxiliary degrees are then defined by

$$\tilde{K}^{\text{in}} = X + Y, \qquad \tilde{K}^{\text{out}} = X + Z \tag{3.6.53}$$

(we have omitted the subindex that provides the neuronal index to point out that this applies to all the degree pairs). It can be seen that, under this definition, $\tilde{K}^{\text{in}}$ and $\tilde{K}^{\text{out}}$ follow Gamma(κ, θ) distributions and their correlation coefficient is ρ. Using the properties of Gamma distributions, we have that

$$\langle\left(\tilde{K}^{\text{in}}\right)^n\rangle \;=\; \langle\left(\tilde{K}^{\text{out}}\right)^n\rangle \;=\; \theta^n\,\frac{\Gamma(\kappa+n)}{\Gamma(\kappa)}, \tag{3.6.54}$$

where $\Gamma(z) := \int_0^\infty t^{z-1}e^{-t}\,\mathrm{d}t$. Fixing $n \in \mathbb{N}$, $n \geq 1$, and integrating by parts,

$$\begin{aligned}
\Gamma(\kappa+n) &= \int_0^\infty t^{\kappa+n-1}e^{-t}\,\mathrm{d}t \\
&= -\,t^{\kappa+n-1}e^{-t}\big]_{t=0}^{t=\infty} + (\kappa+n-1)\int_0^\infty t^{\kappa+n-2}e^{-t}\,\mathrm{d}t \\
&= (\kappa+n-1)\,\Gamma(\kappa+n-1) \\
&= \cdots \\
&= (\kappa+n-1)(\kappa+n-2)\cdots\kappa\,\Gamma(\kappa),
\end{aligned} \tag{3.6.55}$$

so

$$\langle\left(\tilde{K}^{\text{in}}\right)^n\rangle \;=\; \langle\left(\tilde{K}^{\text{out}}\right)^n\rangle \;=\; \theta^n \prod_{m=0}^{n-1}(\kappa+m). \tag{3.6.56}$$

Now we proceed to compute the expected values of the empirical statistics.

$\mathcal{S}_x$, $\mathcal{S}_y$

$$\begin{aligned}
\mathcal{S}_x &= \sum_{i=1}^{N_E} x_i \;=\; \frac{1}{\sqrt{N_E\langle\tilde{K}\rangle}}\sum_{i=1}^{N_E}\tilde{K}_i^{\text{in}}, \\
\langle\mathcal{S}_x\rangle &= \frac{1}{\sqrt{N_E\kappa\theta}}N_E\kappa\theta \;=\; \sqrt{N_E\kappa\theta} \\
\langle\mathcal{S}_y\rangle &= \langle\mathcal{S}_x\rangle.
\end{aligned} \tag{3.6.57}$$

$\mathcal{T}$

$$\begin{aligned}
\mathcal{T} &= \sum_{i=1}^{N_E} x_i y_i \;=\; \frac{1}{N_E\langle K\rangle}\sum_{i=1}^{N_E}\tilde{K}_i^{\text{in}}\tilde{K}_i^{\text{out}}, \\
\langle\mathcal{T}\rangle &= \frac{1}{N_E\kappa\theta}N_E\left(\mathrm{Cov}\left(\tilde{K}_i^{\text{in}},\tilde{K}_i^{\text{out}}\right) + \langle\tilde{K}\rangle^2\right) \\
&= \frac{1}{\kappa\theta}\left(\kappa\theta^2\rho + (\kappa\theta)^2\right) \\
&= (\rho+\kappa)\,\theta.
\end{aligned} \tag{3.6.58}$$

$\mathcal{U}_x, \mathcal{U}_y$

$$\begin{aligned} \mathcal{U}_x &= \sum_{i=1}^{N_E} x_i^2 = \frac{1}{N_E\langle\tilde{K}\rangle}\sum_{i=1}^{N_E}\left(\tilde{K}_i^{\text{in}}\right)^2, \\ \langle\mathcal{U}_x\rangle &= \frac{1}{\kappa\theta}\kappa(\kappa+1)\theta^2 = (\kappa+1)\theta \\ \langle\mathcal{U}_y\rangle &= \langle\mathcal{U}_x\rangle. \end{aligned} \tag{3.6.59}$$

$\mathcal{V}_x, \mathcal{V}_y$

$$\mathcal{V}_x = \sum_{i=1}^{N_E} x_i y_i^2 = \frac{1}{\left(N_E\langle\tilde{K}\rangle\right)^{\frac{3}{2}}}\sum_{i=1}^{N_E}\tilde{K}_i^{\text{in}}\left(\tilde{K}_i^{\text{out}}\right)^2. \tag{3.6.60}$$

Let us notice the following:

$$\begin{aligned} \langle\tilde{K}^{\text{in}}\left(\tilde{K}^{\text{out}}\right)^2\rangle &= \langle(X+Y)\,(X+Z)^2\rangle \\ &= \langle X^3+2X^2Z+XZ^2+X^2Y+2XYZ+YZ^2\rangle \\ &= \theta^3\kappa(\kappa+1)(\kappa+2\rho). \end{aligned} \tag{3.6.61}$$

Therefore,

$$\begin{aligned} \langle\mathcal{V}_x\rangle &= \frac{N_E}{(N_E\kappa\theta)^{\frac{3}{2}}}\langle\tilde{K}^{\text{in}}\left(\tilde{K}^{\text{out}}\right)^2\rangle \\ &= \frac{1}{\sqrt{N_E\kappa\theta}}\theta^2(\kappa+1)(\kappa+2\rho) \\ \langle\mathcal{V}_y\rangle &= \langle\mathcal{V}_x\rangle. \end{aligned} \tag{3.6.62}$$

$\mathcal{Z}$

$$\mathcal{Z} = \sum_{i=1}^{N_E} x_i^2 y_i^2 = \frac{1}{(N_E\langle K\rangle)^2}\sum_{i=1}^{N_E}\left(\tilde{K}_i^{\text{in}}\tilde{K}_i^{\text{out}}\right)^2. \tag{3.6.63}$$

We have

$$\begin{aligned} \langle\left(\tilde{K}_i^{\text{in}}\tilde{K}_i^{\text{out}}\right)^2\rangle &= \langle(X+Y)^2(X+Z)^2\rangle \\ &= \langle(X^2+2XY+Y^2)(X^2+2XZ+Z^2)\rangle \\ &= \theta^4\kappa\left[6\rho+\kappa(1+2\kappa+\kappa^2+8\rho+4\kappa\rho+2\rho^2)\right], \end{aligned} \tag{3.6.64}$$

which implies

$$\begin{aligned}
\langle \mathcal{Z} \rangle &= \frac{N_E}{\left(N_E \langle K \rangle\right)^2} \langle \left(\tilde{K}_i^{\text{in}} \tilde{K}_i^{\text{out}} \right)^2 \rangle \\
&= \frac{1}{N_E \left(\kappa\theta\right)^2} \theta^4 \kappa \left[6\rho + \kappa(1 + 2\kappa + \kappa^2 + 8\rho + 4\kappa\rho + 2\rho^2) \right] \qquad (3.6.65) \\
&= \frac{1}{N_E \kappa} \theta^2 \left[6\rho + \kappa(1 + 2\kappa + \kappa^2 + 8\rho + 4\kappa\rho + 2\rho^2) \right].
\end{aligned}$$

Discussion

In this thesis we have addressed some issues concerning the structure of cortical microcircuits and the relationship between structure and function in these networks. First, we showed that the so-called "nonrandom" features observed in cortex −namely, the over-representation of bidirectional connections and the fact that the connection probability increases with the number of common neighbors (the "common-neighbor rule")− provide, *per se*, little information about the underlying structure, although they can be used to rule out the simplest random structural model, in which connections appear independently with a fixed probability. Such properties are consistent with several alternative structures, ranging from clustered networks to networks generated according to a prescribed degree distribution. We also showed that the statistical measure given by the correlation coefficient of in- and out-degrees in small neuronal samples (SDC) can be used to distinguish between different families of networks even when they quantitatively share the above-mentioned nonrandomness. The analysis of such a measure in data from layer 5 pyramidal neurons in rat somatosensory cortex [Perin et al., 2011] suggests that none of the alternative, canonical structures that we have presented is consistent with the results, even though we could provide a compatible candidate model, defined by a combination of a spatial component and a non-symmetrical modulatory component.

These findings indicate that the connectivity structure that is imposed in the majority of dynamical network models might be overly simplistic, and a natural question to address is what the effects are, in terms of dynamics, of assuming more realistic topologies. Our analysis of the SDC has revealed that positive correlations between degrees of pyramidal neurons exist in cortical circuits when studied locally. Therefore, one is tempted to think that such correlations might provide computational advantages to those networks. Although structures defined by a prescribed degree distribution seem implausible under the light of our data analysis, we decided to explore how these topologies −which can be considered the canonical paradigm to introduce degree correlations− shape dynamics. It is possible that the role played by degree correlations in these networks can

be translated to more realistic neuronal architectures.

We then studied the effect of introducing an arbitrary distribution of in/out-degrees in the repertoire of stationary firing rates that models of leaky integrate-and-fire neurons can exhibit. We considered networks of both excitatory (E) and inhibitory (I) neurons, where the excitatory subnetwork is defined by the above-mentioned model and the other connections are generated independently with a probability that is either constant or, in the case of I-to-E connections, increases with the excitatory in-degree of the post-synaptic neuron. We have defined a way to analytically introduce the effect of this structure in the mean-field formulation already developed to describe the statistical properties of the stationary state in networks with homogeneous or random topology [Amit and Brunel, 1997a,b; Brunel, 2000; Roxin et al., 2011]. We finally applied this formalism to the study of networks with heterogeneous distributions and positive in/out-degree correlation.

The role that degree correlation plays in this formulation is of particular interest. In the mean-field equations, the stationary firing rate of a neuron depends on its in-degree and on the firing rates of pre-synaptic neurons. In a network where there is inter-neuronal variability, neurons with distinct connectivity properties fire at different rates. Therefore, the stationary state is characterized by a distribution of firing rates. This implies that the firing rates among the possible pre-synaptic neurons follow also a certain distribution. If there are no degree correlations, this distribution coincides with the rate distribution in the network. When in- and out-degrees are correlated, however, the distribution of rates in the set of possible pre-synaptic neurons is biased with respect to the distribution of rates in the network. This is due to the fact that the distribution of in-degrees of the putative pre-synaptic neurons deviates from the in-degree distribution in the network due to degree correlations, and such a bias is inherited by the pre-synaptic rate distribution. We have explicitly computed this bias as a function of the degree distribution imposed in the network. The results indicate that degree correlations have an important effect on firing rates in the stationary state, so they could play a non-trivial role in the dynamics of these networks in general.

We also performed simulations to study how the network's response to perturbations of the stationary state depends on the amount of correlation imposed. Our results show that positive degree correlations enhance the responsiveness to transient stimuli. Specifically, there is a critical value for the correlation above which the network responds to the stimulus in a completely different manner, characterized by the presence of bursts and high degree of synchronization. This qualitatively distinct state is maintained even after the stimulus has been removed, lasts for a certain amount of time and finally extinguishes. This finding might be interpreted to mean that

correlations are important for a proper processing of stimuli in real cortical networks.

Finally, we provided an analytical prediction of the spectral properties of connectivity matrices with the above-mentioned structure. It is well known that the spectrum (that is, the collection of eigenvalues) of the connectivity matrix can be used to predict the linear behavior of firing rate models around a stationary state. When dealing with networks whose structure is not fixed but has some degree of randomness, the spectrum is also random. A result due to Girko [Girko, 1985] states that when the entries of this matrix are independent and drawn from a common Gaussian distribution whose mean is zero and whose variance scales as $1/N$, where N is the size of the network, in the large N limit the spectrum is (almost surely) densely confined within a disk (in the complex plane) centered at zero whose radius can be computed from the entries' variance. The study of the spectrum of more general random matrices has attracted much attention in recent years, and new results have been obtained since then [Tao and Vu, 2010; Tao, 2013]. We applied some of these known results to our connectivity matrices to analytically derive the spectral properties in the large N limit. Our results show that these matrices have also their eigenvalues densely located in a disk, except for a fixed number (three or less) of "outliers", whose precise location can be predicted in this limit. When the degrees within the excitatory subnetwork are drawn from a Gamma distribution with correlation coefficient ρ, there is a positive outlier which moves "to the right" as ρ increases.

In appropriate firing rate models, such an eigenvalue may destabilize dynamics in a qualitatively distinct way compared with a multiplicative modulation of all the synaptic weights (because the last perturbation can induce a sudden appearance of many eigenvalues with positive real part). Nevertheless, a multiplicative modulation of the excitatory weights only leads to similar perturbations of the spectrum. Therefore, both a selective increase of the excitatory weights and a rearrangement of the actual synapses so as to introduce positive excitatory-to-excitatory degree correlations appear to be alternative means for destabilizing the stationary state in a similar manner. From the biological point of view, the latter option would be more efficient because of its reduced cost in terms of the demands for chemical resources at the synapses.

In summary, our results indicate that cortical microcircuits deviate from "typical" network models in a way that can be studied locally, although the precise underlying structural plan (assuming that such a plan indeed exists) has not been identified yet. It is highly plausible that real circuits are arranged according to complex topological laws, probably based on a combination of spatial and modulatory, non-symmetric, mechanisms. We have also reported evidences that the number

of incoming and outgoing connections in pyramidal neurons are positively correlated, a structural footprint which might provide new and interesting functional capabilities to these circuits.

Open questions and future directions

The structure of cortical microcircuits

Our research, however, leaves many questions behind and opens the doors to new issues for future investigations. By means of the SDC analysis, we have concluded that it is implausible that cortical microcircuits are built from mainly three general classes of models, but it is difficult to infer additional underlying structural principles from this information only. Although it seems that the local networks of the cortex are constrained by the physical distance between neurons and there exist additional connectivity rules, presumably asymmetric, it is impossible, without further information, to anticipate the nature and properties of such additional mechanisms.

In their attempt to realistically reconstruct, with great detail, a certain volume of the microcircuitry of the rat somatosensory cortex, Markram and colleagues used a random algorithm that positions neurons in 3D space preserving physiological densities and neuronal morphology [Reimann et al., 2015; Markram et al., 2015]. Potential synapses then occur only at incidental appositions between dendrites and axons. If all of these appositions became real synapses, the circuit would end up being massively connected. In order to fit actual connectivity data, they implement a 4-step pruning algorithm that selects only a small fraction of appositions to become functional synapses [Reimann et al., 2015]. They showed that the reconstructed circuit exhibits different "emergent" properties (i.e., properties that have not been used as restrictions for the algorithm) such as the over-representation of bidirectional connections and the common neighbor rule. According to our findings, these features are already found in different simple structural models, so they do not tell much about how realistic a model is. Nonetheless, it would be interesting to study to what extent the SDC in samples from this reconstructed circuit quantitatively matches the experimental results. If the connectivity rules used in the reconstruction were enough to explain many features of the real topology, we could conclude that physical and morphological connectivity rules alone can explain the microcircuit architecture. The asymmetric component that we have identified would then stem from the morphological asymmetry of neurons (when comparing axonal and dendritic arbors).

A recent paper provides a systematic analysis of some of the topological features of this detailed

reconstruction [Gal et al., 2017]. The authors include measures related to degree distributions or motif representation. Interestingly, they show that their digital microcircuit exhibits highly heterogeneous in- and out-degree distributions, although the neurons that possess the larger in- and out-degrees scarcely overlap. This suggests that the cortical microcircuitry might indeed contain a significant number of hubs, but only a small fraction of these highly connected neurons are in- and out-hubs simultaneously. Therefore, in- and out-degrees in the whole circuit could be non-significantly correlated, in contrast with the models that we have studied in Chapters 2 and 3. It is important to note, however, that this result corresponds to the entire reconstructed microcircuit, which spans all the cortical layers and contains many different neuronal types. A study focused on a single layer and cell type (as the data that we used in Chapter 1) might produce different results. In terms of two- and three-neuron motifs, the reconstruction presents clear biases with respect to classical Erdös-Rényi (ER) models. The provided analysis also compares the motif counts with those of distance-dependent null networks (equivalent to our Distance model defined in Chapter 1), and the results show a significant deviation from this model as well. This is in fully agreement with our identification of additional structural rules beyond the effect of intersomatic distance.

An important aspect to bear in mind is that the rules which determine cortical connectivity might depend on types of mechanisms different from the ones that we have considered in this thesis. These rules could involve, for example, interactions between groups of neurons that break down the basic hypothesis of our Modulator model (which is the general framework that includes all the models studied here), namely that connections appear independently from pair to pair once the appropriate neuronal properties are known. Other scenarios are in principle possible, for example a case in which the properties of a given neuron have an influence on the connectivity between neighboring cells.

Another limitation of the analysis that we have performed on the cortical microcircuitry is that we have assumed a binary connectivity: neurons are considered to be either connected or disconnected, without taking into account the variability in synaptic weights that without doubt these networks exhibit [Song et al., 2005]. Such variability for sure adds richness and complexity to the scenarios considered in this thesis. A careful study on the interplay between synaptic weights and the underlying binary structure is therefore needed. On the other hand, the analyzed data come from the examination of excitatory-to-excitatory connections only and are restricted to pyramidal neurons within layer 5 [Song et al., 2005; Perin et al., 2011]. Different types of excitatory neu-

rons have been identified in cortex [Harris and Shepherd, 2015], which indicates that our analysis provides just a preliminary approximation to the question of how the excitatory microcircuitry is organized. Furthermore, the organization of cortical columns into morphologically distinct layers suggests that the connectivity within different layers may be modeled by different laws. Optogenetic studies have shown that the activity spread induced by optical stimulation of a fraction of neurons varies substantially from layer to layer [Beltramo et al., 2013]. This has been interpreted as evidence for a diversity of organizational principles when comparing different layers [Setareh et al., 2017]. It is also well known that inhibitory cells constitute an important fraction of neurons in the brain and they play an essential role in cortical dynamics. Therefore, the comprehension of the whole micro-connectome organization −including inhibition and spanning several cortical layers− is essential for a deep understanding of how these local networks operate. In general terms, inhibitory connections seem to be denser and less specific than excitatory ones [Fino and Yuste, 2011; Fino et al., 2013], but it is known that there exist many types of inhibitory neurons, that their synapses can target specific parts of the post-synaptic cell and that they connect to other neurons with different likelihoods [Markram et al., 2004; Jiang et al., 2015]. A precise picture of the laws that govern these connections is still lacking.

Finally, the study of connectivity between brain areas is another important field of research [Sporns et al., 2007; Bullmore and Sporns, 2012; Markov et al., 2013, 2014; Wang and Kennedy, 2016]. A recent work revealed that interareal connections can be explained to a great extent by models that only depend on physical distance [Ercsey-Ravasz et al., 2013], a finding which suggests that local and long-range connectivity are shaped by both common and unique forces. A complete understanding of how brains are organized will necessarily require a combined research at multiple scales, ranging from the micro-architecture to the connectivity between different areas.

The role of synaptic plasticity

Another issue that our work has not addressed so far is the role of synaptic plasticity in the architecture of the neocortex. The extent to which the acquisition of a proper cortical organization depends on activity and whether this process requires external stimulation is a subject of intense debate. In the visual system, for example, the early formation of ocular dominance columns has been shown to occur even in the absence of visual stimulation, in a process that nevertheless requires spontaneous activity driven by activity waves at the retina [Katz and Crowley, 2002]. In cats, once the columns are formed, there is a critical period during which they are refined in

response to visual experience [Crair et al., 1998]. But recent experiments on the formation of orientation selectivity in mice suggest that it occurs even in the absence of spontaneous cortical activity, although this type of activity is needed in the subsequent reorganization of selectivities [Hagihara et al., 2015].

Our approach when seeking a set of organizational principles that define cortical connectivity implicitly assumed that either plasticity does not shape structure beyond the variability inherent in the proposed models or the effects of plasticity are already included in these structural laws. In the first scenario, the general organization is defined by mechanisms independent of activity, presumably mediated by molecular cues that ultimately rely on a predefined genetic program. In the second scenario, the statistical structure is shaped by plastic modifications of synapses as a result of ongoing activity. In the last case, the general structural scheme that we observe in these circuits can be partially specified genetically but it is also modeled by neuronal activity, presumably at early stages of development. A study that linked realistic activity and plasticity rules with the emergence of particular architectures would therefore be of great interest in this context. But even if the formation of the general microcircuitry plan were activity-independent, a proper candidate model to explain the structure of cortical microcircuits should be such that the introduction of realistic plasticity mechanisms preserves its main statistical features. To what extent the models that we have presented are stable under plausible plastic modifications is a wide question that should be investigated in the future as well, and that could provide additional, indirect criteria for assessing the reliability of different structural candidates.

The interplay between structure and function

The ultimate goal of neural science is to understand the mechanisms by which brains perform different tasks and computations. We started this thesis by stressing that brain functioning is tightly related to brain wiring, so an important step is the understanding of how local neuronal networks are arranged. An extensive set of experimental studies has revealed that the local circuits of the cortex share some regularities, namely that the connectivity tends to be sparse, that some neuronal motifs are over- or underrepresented and that connection probability increases with the number of common neighbors. We have also seen that networks with such features exhibit positive correlations between in- and out-degrees, at least locally. An important (and challenging) next step to make is to study whether there is a link between any of these features and specific functional capabilities.

It is plausible that some of these properties play a functional role in cortical networks, whereas others might be mere by-products of either synaptic plasticity or higher-order organizational principles (when we say that a property plays a given functional role we mean that if we could group all the possible networks into those that exhibit the property and those that do not, the first family would perform better the function in question). The over-representation of reciprocal connections, for example, is an almost omnipresent feature in cortical microcircuits [Mason et al., 1991; Markram et al., 1997; Song et al., 2005; Le Bé and Markram, 2006; Wang et al., 2006]. One possibility is that the excess of bidirectionality facilitates neuronal computation in some way. In a recent study, this property was found in networks that had been optimized to store a large number of patterns [Brunel, 2016], which could be interpreted as evidence for a role of such motif in memory storage. On the other hand, our work shows that such a feature is present in many different types of network organization, so what facilitates memory storage might be not the presence of reciprocal motifs *per se* but a precise structural plan that exhibits −as many models do− this attribute.

Thus, a study of how distinct topological features modulate functional or dynamical properties will help to clarify these issues and will shed some light on the relationship between structure and function in brain networks. Zhao and colleagues studied the influence of second-order connectivity motifs on network synchronization in models of coupled oscillators whose structure is defined by the occurrence of reciprocal, convergent, divergent and chain motifs [Zhao et al., 2011]. They showed that convergent motifs prevent synchrony whereas chain motifs tend to enhance it. Recall that, in any network, an increase of chain motifs implies larger degree correlations (see Eq. (1.6.19)), so positive degree correlations would in principle facilitate synchronization, a prediction consistent with our findings in networks that receive transient stimulation.

Our contribution to this problem was to study the effect of degree distributions and degree correlation in some macroscopic properties of the stationary asynchronous state (Chapter 2) and in the distribution of the connectivity matrix's eigenvalues (Chapter 3). Again, these analyses cannot dissect the effect of degree correlations alone, for we have considered a very particular class of networks: those that are created according to a prescribed degree distribution. Our results are thus constrained to this network family and one should be careful when inferring from them a role of degree correlations in general. For example, all the alternative models presented in Chapter 1 exhibit such correlations locally, but whether they exert a similar influence on dynamics is something that has to be further investigated.

The results presented in Chapter 1 suggest that cortical microcircuits are far from simple clustered, distance-dependent networks and networks generated from a given degree distribution. What we observe, in contrast, is a structure that has a clear dependency on distance and that is shaped by an additional component which is necessarily asymmetric. We have provided a possible interpretation of this asymmetry in terms of hierarchical clustering: the data seem to be compatible with a network that has a clustered organization at different levels, in the sense that connection density is high within clusters and the influences between clusters are asymmetric. The effect of clustering in neuronal dynamics has been studied before [Deco and Hugues, 2012; Litwin-Kumar and Doiron, 2014; Mazzucato et al., 2015]. These studies have shown that networks of excitatory and inhibitory neurons with clustered structure within the excitatory subnetwork can explain some statistical features of the neuronal response to stimuli observed *in vivo*. For example, in general situations there is a large variability in the activity of individual neurons when measured across different trials. This variability is nonetheless reduced when a stimulus is presented [Churchland et al., 2010]. Such a reduction of trial-to-trial variability induced by stimuli can be reproduced in classical clustered models, mainly because clustering promotes the appearance of multiple network attractors, which are characterized by the activation of the neurons that compose the different clusters. These attractors are explored randomly in the absence of specific stimulation but reliably otherwise. Although we have not investigated the effect of a hierarchical organization in this kind of neuronal dynamics, hierarchical clustering may constitute a structural basis for the presence of different attractors whose sequential activation could be to some extent shaped by the hierarchy. This might allow for a partially ordered activation of functional modules during complex computations. The putative hierarchical architecture in the cortical microcircuit could be the structural bedrock for a systematic organization of the fundamental computations that are needed for producing higher order operations in the brain.

Epíleg

> "La experiencia más hermosa que tenemos a nuestro alcance es el misterio. Es la emoción fundamental que está en la cuna del verdadero arte y de la verdadera ciencia."
>
> Albert Einstein, *El mundo tal como yo lo veo* [1]

El cervell és la nostra finestra al món. En l'ésser humà, la necessitat d'adaptació i, per tant, d'interpretació d'allò que ens envolta ha portat a invencions extraordinàries com el llenguatge, el raonament abstracte, la ciència o l'ètica. L'afany per entendre com funciona el cervell es nodreix, en part, d'un interès primigeni per comprendre els mecanismes que han donat lloc a un univers mental tan ric en possibilitats.

Vaig voler estudiar neurociència fruit de l'admiració pel fet que un tros de matèria com és l'òrgan cerebral, esculpit per milions d'anys d'evolució però al cap i a la fi abandonat a la deriva de la prova i de l'error, hagi engendrat totes aquestes habilitats. L'estudi del cervell amaga algunes de les qüestions més intrigants, i a l'inici estava convençuda que l'anàlisi minuciosa dels elements que conformen el sistema nerviós i les seves relacions mútues acabaria portant a una comprensió acurada de tots els processos cerebrals. Després d'uns anys d'estudi de la biologia i de la ciència del sistema nerviós, la meva visió ha canviat en certa manera. Aquí intentaré explicar per què.

La neurociència moderna ha estat capaç de revelar molts detalls sobre la fisiologia del sistema nerviós. Sabem que està constituït per cèl·lules especialitzades −les neurones−, que es comuniquen entre elles a través d'una combinació de senyals elèctrics i químics. Coneixem a grans trets com funcionen els sistemes sensorials, en els quals una cadena d'esdeveniments físico-químics transformen la recepció inicial d'un estímul per part de sensors especialitzats en activitat neuronal localitzada en àrees concretes de l'escorça cerebral. També s'han identificat els mecanismes bàsics

[1] A. Einstein. *Mis ideas y opiniones*. Editorial Bosch, Barcelona, 1985.

de l'aprenentatge, segons els quals la memòria es fonamenta en una reorganització de les connexions entre neurones, que en determina els patrons d'activitat i, en última instància, els tipus d'operacions que es poden dur a terme. Recordar és reactivar seqüències d'activitat neuronal que el cervell va reproduir en el passat. Recentment hem après també que la memòria espacial podria estar lligada a la formació de "mapes" d'activitat que constitueixen una representació de l'espai que ens envolta.

Tots aquests avenços, que s'han assolit després de molts anys de recerca i pacient dedicació, ens diuen molt sobre el cervell com a mecanisme complex que recull, processa i transmet informació. És raonable imaginar que en el futur podríem arribar a explicar els esdeveniments fisiològics que vinculen una certa recepció sensorial amb una resposta concreta, per complicada que pugui arribar a ser. Entendre el cervell és, des d'aquest punt de vista, donar una descripció acurada de l'entramat d'engranatges que constitueixen la seva maquinària interna.

Aquesta descripció, però, no és completa. Hi manca un ingredient, subtil però indispensable, que la ciència del cervell fins ara ha deixat de banda. Em refereixo a l'experiència subjectiva o *consciència*. Per no crear confusió, explicaré millor què entenc per consciència: la consciència és la capacitat de *sentir*, d'experimentar sensacions. La sensació de fred, la percepció del vermell, l'experiència del dolor són atributs de la consciència, i totes elles (i moltes altres) encaixen en una experiència global i unificada que configura el nostre univers particular.

Vull posar èmfasi en el fet que no m'estic referint a una qualitat "elaborada", de les que solem lligar a l'intel·lecte humà. L'experiència subjectiva pot admetre diferents nivells, i és possible que en nivells "superiors" trobem, per exemple, l'autoconsciència, propietat potser només present en determinades espècies com la nostra. Si bé la consciència del jo o de l'altre, la capacitat per representar un món imaginat a través del llenguatge, etc., són fenòmens altament interessants i encara del tot desconeguts, em vull centrar en la consciència en el seu sentit més bàsic.

La consciència marca una diferència fonamental; tenir consciència significa ser capaç d'experimentar algun tipus de sensació subjectiva (per "elemental" que pugui semblar): caminar de les tenebres a la llum. I aquí és on rau la major dificultat a la qual s'enfronta la ciència del cervell: explicar com determinats processos neuronals, presumiblement mecànics en darrer terme, donen lloc a la sensació. És exactament això el que el matemàtic i filòsof David Chalmers ha anomenat "the hard problem" (el problema difícil), en contraposició amb els problemes "fàcils", que són l'objecte d'estudi de la ciència del cervell actual. En paraules de Chalmers:

> "The really hard problem of consciousness is the problem of *experience*. [...]
>
> It is undeniable that some organisms are subjects of experience, but the question of why it is that these systems are subjects of experience is perplexing. [...] It is widely agreed that experience arises from a physical basis, but we have no good explanation of why and how it so arises. Why should physical processing give rise to a rich inner life at all? It seems objectively unreasonable that it should, and yet it does.
>
> If any problem qualifies as *the* problem of consciousness, it is this one."
>
> David Chalmers, *Facing up to the problem of consciousness* [2]

La perplexitat ve de pensar que l'execució d'un conjunt de processos en última instància físics doni lloc a algun tipus d'experiència. En quin punt de la cadena de reaccions elèctriques i químiques entre neurones es produeix la sensació? Que una sèrie d'esdeveniments cerebrals acabi donant com a resultat una resposta adequada o un comportament complex, per difícil d'explicar que sigui, no té res a veure amb el fet que l'execució d'aquests esdeveniments vagi acompanyada d'una experiència. Leibniz va expressar el problema d'aquesta manera:

> "Imaginémonos que haya una máquina cuya estructura la haga pensar, sentir y tener percepción; podremos concebirla agrandada, conservando las mismas proporciones, de tal manera que podamos entrar en ella como en un molino. Esto supuesto, si la inspeccionamos por dentro, no hallaremos más que piezas que se impelen unas a otras, pero nunca nada con que explicar una percepción."
>
> Gottfried Wilhelm Leibniz, *Monadología* [3]

Si crec que hi ha un buit en la nostra aproximació actual al cervell no és perquè pensi que la concepció mecanicista no pot donar lloc a una descripció acurada dels processos cerebrals. No hi veig, en principi, cap obstacle fonamental. De la mateixa manera que els ordinadors són cada vegada més potents i són capaços de superar en certs aspectes la intel·ligència humana, concebo la possibilitat que el cervell es pugui reduir a un seguit de processos computacionals. Però el que m'intriga profundament és per què aquests processos haurien d'anar acompanyats d'una experiència de cap mena. La majoria de les persones creu, encara que sigui indemostrable, que els ordinadors no tenen consciència, però això no és a priori cap obstacle perquè no puguin efectuar un gran nombre d'operacions complicades. I això ens porta a una de les qüestions més fascinants pel que fa al problema de la consciència: el seu rol en l'evolució.

En aquest punt és fàcil caure en un parany conceptual. Sembla evident que tenir la capacitat de sentir dolor, per posar un exemple senzill, és adaptatiu. L'animal que experimenta una sensació

[2] D. J. Chalmers. *The Character of Consciousness*. Oxford Scholarship Online, 2010.
[3] G. W. Leibniz. *Monadología*. Biblioteca Nueva, Madrid, 2001.

desagradable produïda per un agent nociu evitarà tornar a interaccionar-hi. Però si analitzem amb més profunditat aquest argument veiem que no és del tot consistent, almenys dins la concepció que la majoria de nosaltres tenim del cervell. Una relació causa-efecte com aquesta s'hauria de poder explicar com una cadena d'esdeveniments neuronals que lliguen la recepció de l'estímul nociu amb el comportament d'evitació. A efectes pràctics, el rellevant és que hi hagi una successió de processos materials que produeixin el comportament, processos que podrien ser tant innats com apresos (la distinció aquí no és important). Però això, un cop més, sembla no deixar lloc per a l'experiència, que passaria a ser una propietat supèrflua.

Es podria argumentar que el paper de la consciència no roman en l'execució del comportament en si sinó en el seu aprenentatge. Tots sabem, per la nostra pròpia experiència, que hi ha processos cerebrals que es poden produir de manera inconscient. No només això, l'aprenentatge, per repetició, d'un comportament pot acabar convertint-lo en quasi inconscient. El gran físic Erwin Schrödinger, en el seu assaig *Mind and Matter*, ho va expressar així:

> "Any succession of events in which we take part with sensations, perceptions and possibly with actions gradually drops out of the domain of consciousness when the same string of events repeats itself in the same way very often. [...]
>
> On frequent repetition the whole string of events becomes more and more of a routine, it becomes more and more uninteresting, the responses become ever more reliable according as they fade from consciousness. The boy recites his poem, the girl plays her piano sonata 'well-nigh in their sleep'. We follow the habitual path to our workshop, cross the road at the customary places, turn into side-streets, etc., whilst our thoughts are occupied with entirely different things. But whenever the situation exhibits a relevant differential −let us say the road is up at the place where we used to cross it, so that we have to make a detour− this differential and our response to it intrude into consciousness, from which, however, they soon fade below the threshold, if the differential becomes a constantly repeated feature. [...]
>
> One might say, metaphorically, that consciousness is the tutor who supervises the education of the living substance, but leaves his pupil alone to deal with all those tasks for which he is already sufficiently trained."
>
> Erwin Schrödinger, *Mind and Matter* [4]

Podria ser que la consciència tingués un paper en la realització de conductes no estereotipades, per a les quals el cervell no disposa d'un pla d'acció definit prèviament? Un dels avenços de la neurociència moderna ha consistit en la identificació dels mecanismes fisiològics de la memòria.

[4] E. Schrödinger. *What is Life? With Mind and Matter and Autobiographical Sketches*. Cambridge University Press, Cambridge, 2010.

La memòria ja no sembla ser un procés misteriós i inaccessible a l'anàlisi sistemàtica de les ciències empíriques, sinó el resultat, un cop més, d'esdeveniments moleculars ben definits. Novament, la memòria és mecànica, i l'experiència lligada a l'aprenentatge seria un afegit prescindible.

Una manera d'evitar aquest obstacle és suposar que la consciència sorgeix com a epifenomen en sistemes amb un cert grau de complexitat. El terme epifenomen es fa servir per indicar que no hi ha una funció associada al fenomen en si mateix, el qual apareix simplement com a efecte col·lateral. Aquesta idea resulta estranya; la clara coherència de la consciència la fa difícil d'acceptar: si la consciència és tan sols epifenomen, per què hauria de ser coherent amb els interessos de supervivència de l'individu? Per què els estímuls nocius serien dolorosos i les accions que aporten beneficis, plaents?

Sospito que la consciència ha de tenir alguna rellevància per a la vida, una significació que encara no hem entès. Malgrat que és impossible accedir a l'experiència subjectiva aliena, és raonable pensar que moltes espècies estan dotades d'un cert grau de consciència, almenys entre els animals que posseeixen un sistema nerviós (és suficient observar les analogies entre la fisiologia del sistema nerviós dels mamífers, per exemple). Afirmar el contrari seria tan absurd com pensar que els éssers humans que m'envolten no són més que sofisticats autòmats. I estirant d'aquest fil arribem a una altra pregunta: si és cert que la consciència va ser un afortunat "invent" de l'evolució, quan va aparèixer i què va determinar la seva aparició? Podem concebre la consciència simplement com un accident evolutiu més? Com seria un món sense ningú que el pogués observar? De nou, les paraules de Schrödinger defineixen molt bé l'abisme:

> "Are we prepared to believe that this very special turn in the development of the higher animals, a turn that might after all have failed to appear, was a necessary condition for the world to flash up to itself in the light of consciousness? Would it otherwise have remained a play before empty benches, not existing for anybody, thus quite properly speaking not existing? This would seem to me the bankruptcy of a world picture. The urge to find a way out of this impasse ought not to be damped by the fear of incurring the wise rationalists' mockery."
>
> Erwin Schrödinger, *Mind and Matter* [5]

Aquest és un dilema actual que ve d'antic. Alguns filòsofs creuen que mai es resoldrà, mentre que d'altres argumenten que cal ampliar els principis elementals de la física actual per donar-li cabuda. També s'ha suggerit que la consciència podria estar lligada a certs fenòmens de la física quàntica, que sembla donar un paper important a l'acte d'observar. Alguns fins i tot creuen que

[5] E. Schrödinger. *What is Life? With Mind and Matter and Autobiographical Sketches.* Cambridge University Press, Cambridge, 2010.

el problema no existeix com a tal, que és una mera il·lusió que s'esvairà tan bon punt disposem d'una descripció prou acurada dels mecanismes neuronals.

No puc estar més en desacord amb el darrer punt de vista. Vaig estudiar neurociència fruit d'una admiració davant l'enigma de la consciència, però com més m'he endinsat en l'estudi del cervell, més i més l'he trobada a faltar. La consciència s'escola entre les mans quan intentes atrapar-la. Però no puc admetre que sigui una simple il·lusió; si d'alguna cosa tinc una certesa absoluta és precisament del fet que *sento*. Això és, ara per ara, tot el que puc afirmar del món.

Appendices

Appendix A

Some notes on stochastic processes

This section is not intended to provide an exhaustive set of definitions and results related to diffusion processes but rather a minimal "toolkit" that will help in understanding Appendix B, which is directly related to neuronal models. Therefore, we just provide here basic concepts and ideas, without going into their technical and mathematical details. All the results are mainly based on chapters 1, 2 and 3 of [Ricciardi, 1977].

A.1 Preliminary definitions and notation

Stochastic process

A stochastic process is a collection of random variables $\{X(t,\xi)\}_{t\in T}$ indexed by a set $T \subset \mathbb{R}$. Usually the index t represents time, and the process describes some phenomenon which evolves over time. The variable ξ can be interpreted as the outcome of a random experiment. Given an outcome ξ_0, $X(t,\xi_0)$ is a deterministic function of t which is called a *realization* or *sample path*. On the contrary, given a time t_0, $X(t_0,\xi)$ is a random variable. If t_0 and ξ_0 are fixed, $X(t,\xi)$ is a number in a state space S. We say that the process is *discrete* when S is a discrete set of points. If S contains a continuum of possible states, the process is *continuous*. The index $t \in T$ can also be either discrete or continuous. In the following we will denote the stochastic process simply by $X(t)$ and we will only consider processes continuous in space and time.

Markov process and transition p.d.f.

Given a continuous stochastic process $X(t)$ and instants $t_1 < \cdots < t_n$, we denote by

$$\rho\left(X(t_n) \,|\, X(t_{n-1}), \cdots, X(t_1)\right) \tag{A.1.1}$$

the probability density function (p.d.f.) of $X(t_n)$ conditioned on $X(t_{n-1}), \cdots, X(t_1)$. We say that the process is *Markov* if for all n and $t_1 < \cdots < t_n$,

$$\rho\left(X(t_n) \,|\, X(t_{n-1}), \cdots, X(t_1)\right) \;=\; \rho\left(X(t_n) \,|\, X(t_{n-1})\right). \tag{A.1.2}$$

In other words, a process $X(t)$ is Markov if the distribution of possible states in the future depends on the present state but not on the previous history of the process. In this case, once the p.d.f. of the initial state, $\rho(X(t_0))$, is known, all the important information about the process is contained in the so-called *transition p.d.f.* of the process:

$$\rho\left(x, t \,|\, x_0, t_0\right) = \rho\left(X(t) = x \,|\, X(t_0) = x_0\right). \tag{A.1.3}$$

We say that a Markov process is *stationary* if for all x, x_0, t, t_0,

$$\rho\left(x, t \,|\, x_0, t_0\right) = \rho\left(x, t - t_0 \,|\, x_0, 0\right). \tag{A.1.4}$$

From now on we will assume that the possible states of the process are real numbers. The transition p.d.f. of any Markov process satisfies the so-called Smolukowski equation:

$$\rho(x, t \,|\, x_0, t_0) \;=\; \int\limits_{\mathbb{R}} \rho(y, \tau \,|\, x_0, t_0)\, \rho(x, t \,|\, y, \tau)\, \mathrm{d}y \tag{A.1.5}$$

for any x_0, x and $t_0 < \tau < t$.

Laplace transform and moment generating function

The *Laplace transform* of a function $f(x)$, $x \in [0, \infty)$, is the function

$$\mathcal{L}\{f\}(\lambda) := \int\limits_0^\infty e^{-\lambda x} f(x)\, \mathrm{d}x \tag{A.1.6}$$

defined on the complex numbers λ whose real part is positive. The following properties hold:

(i) Differentiation:

$$\mathcal{L}\{f'\}(\lambda) = \lambda\mathcal{L}\{f\}(\lambda) - f(0) \tag{A.1.7}$$

(this holds as long as $\lim_{x\to\infty} e^{-\lambda x} f(x) = 0$).

(ii) Convolution:

$$L\{f * g\}(\lambda) = \mathcal{L}\{f\}(\lambda) \cdot \mathcal{L}\{g\}(\lambda), \tag{A.1.8}$$

where the convolution of two functions f and g is defined as follows:

$$(f * g)(x) = \int_{-\infty}^{\infty} f(t)\, g(x-t)\, \mathrm{d}t. \tag{A.1.9}$$

For functions f, g only defined in $[0, \infty)$, we define the convolution as

$$(f * g)(x) = \int_{0}^{x} f(t)\, g(x-t)\, \mathrm{d}t. \tag{A.1.10}$$

Given a continuous random variable X with p.d.f. f, the *moment generating function of* X evaluated at $\lambda \in \mathbb{C}$ is the expectation of $e^{\lambda X}$ (whenever it exists):

$$M(\lambda) := \mathbb{E}\left[e^{\lambda X}\right] = \int_{-\infty}^{\infty} e^{\lambda x} f(x)\, \mathrm{d}x. \tag{A.1.11}$$

The name of the function comes from the following relationship between the moments of X and the derivatives of $M(\lambda)$:

$$\mathbb{E}[X^n] = \left[\frac{\mathrm{d}^n M(\lambda)}{\mathrm{d}\lambda^n}\right]_{\lambda=0}. \tag{A.1.12}$$

If X takes values in $[0, +\infty)$, the moment generating function of X is related to the Laplace transform of f through

$$M(\lambda) = \mathcal{L}\{f\}(-\lambda). \tag{A.1.13}$$

A.2 Kinetic equations

Forward equation

Let us consider a continuous Markov process $X(t)$. Recall that the transition p.d.f. of $X(t)$ satisfies the Smolukowski equation (A.1.5). We can rewrite this equation as follows:

$$\rho(x,t+\Delta t\,|\,x_0,t_0)-\rho(x,t\,|\,x_0,t_0) \;=\; \int\limits_{\mathbb{R}} \rho(y,t\,|\,x_0,t_0)\,\rho(x,t+\Delta t\,|\,y,t)\,\mathrm{d}y-\rho(x,t\,|\,x_0,t_0). \tag{A.2.1}$$

Now we consider a function $R(x)$ with the property that all its derivatives and the function itself vanish sufficiently rapidly when $x\to\pm\infty$. Multiplying both sides of Eq. (A.2.1) by $R(x)/\Delta t$ and integrating over x we have

$$\begin{aligned}\int\limits_{\mathbb{R}} R(x)\tfrac{\rho(x,t+\Delta t\,|\,x_0,t_0)-\rho(x,t\,|\,x_0,t_0)}{\Delta t}\,\mathrm{d}x \;&=\; \tfrac{1}{\Delta t}\int\limits_{\mathbb{R}} R(x)\int\limits_{\mathbb{R}}\rho(y,t\,|\,x_0,t_0)\,\rho(x,t+\Delta t\,|\,y,t)\,\mathrm{d}y\,\mathrm{d}x\\ &\quad -\tfrac{1}{\Delta t}\int\limits_{\mathbb{R}} R(x)\rho(x,t\,|\,x_0,t_0)\,\mathrm{d}x.\end{aligned} \tag{A.2.2}$$

Let us consider the Taylor expansion of $R(x)$ around $x=y$:

$$R(x)=R(y)+\sum_{n=1}^{\infty}\frac{1}{n!}\frac{\mathrm{d}^nR(y)}{\mathrm{d}y^n}(x-y)^n. \tag{A.2.3}$$

Substituting $R(x)$ in the right-hand side of Eq. (A.2.2) by this Taylor expansion and taking the limit $\Delta t\to 0$ we have

$$\begin{aligned}&\int\limits_{\mathbb{R}} R(x)\,\tfrac{\partial}{\partial t}\rho(x,t\,|\,x_0,t_0)\,\mathrm{d}x\\ &\quad=\; \lim_{\Delta t\to 0}\tfrac{1}{\Delta t}\int\limits_{\mathbb{R}} R(y)\rho(y,t\,|\,x_0,t_0)\int\limits_{\mathbb{R}}\rho(x,t+\Delta t\,|\,y,t)\,\mathrm{d}x\,\mathrm{d}y\\ &\qquad+\sum_{n=1}^{\infty}\tfrac{1}{n!}\int\limits_{\mathbb{R}}\tfrac{\mathrm{d}^nR(y)}{\mathrm{d}y^n}\,\rho(y,t\,|\,x_0,t_0)\left(\lim_{\Delta t\to 0}\tfrac{1}{\Delta t}\int\limits_{\mathbb{R}}(x-y)^n\,\rho(x,t+\Delta t\,|\,y,t)\,\mathrm{d}x\right)\mathrm{d}y\\ &\qquad-\lim_{\Delta t\to 0}\tfrac{1}{\Delta t}\int\limits_{\mathbb{R}} R(x)\rho(x,t\,|\,x_0,t_0)\,\mathrm{d}x.\end{aligned} \tag{A.2.4}$$

Just noticing that

$$\int\limits_{\mathbb{R}}\rho(x,t+\Delta t\,|\,y,t)\,\mathrm{d}x=1, \tag{A.2.5}$$

Eq. (A.2.4) reduces to

$$\int_{\mathbb{R}} R(x)\,\frac{\partial}{\partial t}\rho(x,t\,|\,x_0,t_0)\,\mathrm{d}x \;=\; \sum_{n=1}^{\infty}\frac{1}{n!}\int_{\mathbb{R}}\frac{\mathrm{d}^n R(x)}{\mathrm{d}x^n}\,A_n(x,t)\,\rho(x,t\,|\,x_0,t_0)\,\mathrm{d}x, \tag{A.2.6}$$

where

$$A_n(x,t) := \lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{\mathbb{R}}(y-x)^n\,\rho(y,t+\Delta t\,|\,x,t)\,\mathrm{d}y. \tag{A.2.7}$$

Integrating by parts the right-hand side of Eq. (A.2.6) and using the fact that $R(x)$ and all its derivatives vanish for $x\to\pm\infty$ we can rewrite Eq. (A.2.6) as

$$\int_{\mathbb{R}} R(x)\,\frac{\partial}{\partial t}\rho(x,t\,|\,x_0,t_0)\,\mathrm{d}x \;=\; \sum_{n=1}^{\infty}\frac{(-1)^n}{n!}\int_{\mathbb{R}} R(x)\frac{\partial^n}{\partial x^n}\left[A_n(x,t)\,\rho(x,t\,|\,x_0,t_0)\right]\,\mathrm{d}x. \tag{A.2.8}$$

Since, apart from the conditions at $\pm\infty$, the function $R(x)$ was arbitrary, Eq. (A.2.8) can only be fulfilled if

$$\frac{\partial}{\partial t}\rho(x,t\,|\,x_0,t_0) \;=\; \sum_{n=1}^{\infty}\frac{(-1)^n}{n!}\,\frac{\partial^n}{\partial x^n}\left[A_n(x,t)\,\rho(x,t\,|\,x_0,t_0)\right]. \tag{A.2.9}$$

This is known as the *forward* differential form of the Smolukowski equation.

Infinitesimal moments

Now we notice something about the functions $\{A_n(x,t)\}_n$ given by (A.2.7). Let us consider Δt sufficiently small and let us define

$$\Delta X_t = X(t+\Delta t) - X(t). \tag{A.2.10}$$

We have

$$\mathbb{E}\left[(\Delta X_t)^n\,|\,X(t)=x\right] \;=\; \int_{\mathbb{R}}(y-x)^n\,\rho(y,t+\Delta t\,|\,x,t)\,\mathrm{d}y, \tag{A.2.11}$$

so

$$\lim_{\Delta t\to 0}\frac{\mathbb{E}\left[(\Delta X_t)^n\,|\,X(t)=x\right]}{\Delta t} = A_n(x,t). \tag{A.2.12}$$

This is the reason why the functions $\{A_n(x,t)\}_n$ are called the *infinitesimal moments* of the process. The first order moment $A_1(x,t)$ is called the *drift* of the process. The conditional variance of the increment ΔX_t is

$$\sigma^2\left[\Delta X_t\,|\,X(t)=x\right] \;=\; \mathbb{E}\left[(\Delta X_t)^2\,|\,X(t)=x\right] - \mathbb{E}\left[\Delta X_t\,|\,X(t)=x\right]^2 \tag{A.2.13}$$

so

$$\lim_{\Delta t \to 0} \frac{\sigma^2 \left[\Delta X_t \,|\, X(t) = x\right]}{\Delta t} \quad = \quad A_2(x,t) \tag{A.2.14}$$

as long as $A_1(x,t)$ is finite. The second-order infinitesimal moment $A_2(x,t)$ is known as the *infinitesimal variance* of the process. Notice that if the process is stationary (see condition (A.1.4)), all the infinitesimal moments are independent of t.

Backward equation

Eq. (A.2.9) is known as the *forward* Smolukowski equation because it involves derivatives of $\rho(x,t\,|\,x_0,t_0)$ with respect to the present time t and the present state x, whereas the initial condition (x_0,t_0) is treated as a parameter. It is also possible to derive another differential form of the Smolukowsi equation which involves derivatives with respect to the initial condition and where the present state (x,t) appears as a parameter.

We can write the Smolukowski equation (A.1.5) in the following form:

$$\rho(x,t\,|\,x_0,t_0-\Delta t) \quad = \quad \int_{\mathbb{R}} \rho(y,t_0\,|\,x_0,t_0-\Delta t)\,\rho(x,t\,|\,y,t_0)\,\mathrm{d}y, \tag{A.2.15}$$

where $\Delta t > 0$. Substracting $\rho(x,t\,|\,x_0,t_0)$ from both sides of Eq. (A.2.15) and using

$$\rho(x,t\,|\,x_0,t_0) \quad = \quad \int_{\mathbb{R}} \rho(y,t_0\,|\,x_0,t_0-\Delta t)\,\rho(x,t\,|\,x_0,t_0)\,\mathrm{d}y \tag{A.2.16}$$

we obtain

$$\rho(x,t\,|\,x_0,t_0-\Delta t)-\rho(x,t\,|\,x_0,t_0) \quad = \quad \int_{\mathbb{R}} \rho(y,t_0\,|\,x_0,t_0-\Delta t)\left[\rho(x,t\,|\,y,t_0)-\rho(x,t\,|\,x_0,t_0)\right]\mathrm{d}y. \tag{A.2.17}$$

Now we can express $\rho(x,t\,|\,y,t_0)$ as a Taylor series around $y = x_0$ as follows:

$$\rho(x,t\,|\,y,t_0) \quad = \quad \sum_{n=0}^{\infty} \frac{1}{n!}\,\frac{\partial^n \rho(x,t\,|\,x_0,t_0)}{\partial x_0^n}\,(y-x_0)^n. \tag{A.2.18}$$

Introducing this expression in the right-hand side of Eq. (A.2.17) we get

$$\rho(x,t\,|\,x_0,t_0-\Delta t)-\rho(x,t\,|\,x_0,t_0) \quad = \quad \sum_{n=1}^{\infty} \tfrac{1}{n!}\,\tfrac{\partial^n \rho(x,t\,|\,x_0,t_0)}{\partial x_0^n} \int_{\mathbb{R}} \rho(y,t_0\,|\,x_0,t_0-\Delta t)\,(y-x_0)^n\,\mathrm{d}y. \tag{A.2.19}$$

Finally, dividing both sides of Eq. (A.2.19) by $(-\Delta t)$ and taking the limit $\Delta t \to 0$ we obtain

$$\frac{\partial}{\partial t_0}\rho(x,t\,|\,x_0,t_0) \;=\; -\sum_{n=1}^{\infty}\frac{1}{n!}\,\frac{\partial^n \rho(x,t\,|\,x_0,t_0)}{\partial x_0^n}\,A_n(x_0,t_0), \tag{A.2.20}$$

where $\{A_n(x,t)\}_n$ are the infinitesimal moments defined earlier (A.2.7). This version of the Smolukowski equation is known as the *backward* kinetic equation of a Markov process.

The forward and backward equations (A.2.9), (A.2.20) contain an infinite number of terms. An important result known as the Pawula Theorem states that, if all the infinitesimal moments $\{A_n(x,t)\}_n$ exist and any even order moment is zero, then $A_n(x,t) = 0$ for $n \geq 3$. In this situation the equations reduce to

$$\begin{aligned}
\frac{\partial}{\partial t}\rho(x,t\,|\,x_0,t_0) \;&=\; -\frac{\partial}{\partial x}\left[A_1(x,t)\,\rho(x,t\,|\,x_0,t_0)\right] + \frac{1}{2}\,\frac{\partial^2}{\partial x^2}\left[A_2(x,t)\,\rho(x,t\,|\,x_0,t_0)\right] \\
\frac{\partial}{\partial t_0}\rho(x,t\,|\,x_0,t_0) \;&=\; -\frac{\partial \rho(x,t\,|\,x_0,t_0)}{\partial x_0}\,A_1(x_0,t_0) - \frac{1}{2}\,\frac{\partial^2 \rho(x,t\,|\,x_0,t_0)}{\partial x_0^2}\,A_2(x_0,t_0).
\end{aligned} \tag{A.2.21}$$

These are the so-called *Fokker-Planck* (forward) and *Kolmogorov* (backward) equations. Together, they are known as *diffusion equations*, and a stochastic process which obeys (A.2.21) is called a *diffusion process*.

A.3 First passage time

Let us imagine that we are studying a continuous, stationary stochastic process $X(t) \in \mathbb{R}$ which is Markov and is such that $X(0) = x_0$. In this case we use the simplified notation $\rho(x,t\,|\,x_0) := \rho(x,t\,|\,x_0,0)$. Let $S \in \mathbb{R}$ be a possible state of the system and define T as the time at which $X(t) = S$ for the first time. T is a continuous random variable taking values in $[0,\infty)$ and it is known as the *first passage time* of the process $X(t)$. We define $g(S,t\,|\,x_0)$ as the p.d.f. of T.

Let us consider an initial condition $x_0 < S$ (the result can be easily extended to the case $x_0 > S$). Now we take a time point $t > 0$ and a state $x > S$. A continuous sample path which departs from $x_0 < S$ and ends up at $x > S$ at time t necessarily has crossed S at some time point within $(0,t)$. We can therefore write

$$\rho(x,t\,|\,x_0) = \int_0^t g(S,\tau\,|\,x_0)\,\rho(x,t\,|\,S,\tau)\,d\tau = \int_0^t g(S,\tau\,|\,x_0)\,\rho(x,t-\tau\,|\,S)\,d\tau. \tag{A.3.1}$$

Expression (A.3.1) states that $\rho(x,t\,|\,x_0)$ (interpreted as a function of t) is the convolution of the functions $g(S,t\,|\,x_0)$ and $\rho(x,t\,|\,S)$.

Now we will apply the Laplace transform to both sides of Eq. (A.3.1). To make notation simpler, we define

$$\begin{aligned}\rho_\lambda(x|x_0) &:= \mathcal{L}\{\rho(x,\cdot\,|\,x_0)\}(\lambda),\\ g_\lambda(x|x_0) &:= \mathcal{L}\{g(x,\cdot\,|\,x_0)\}(\lambda).\end{aligned} \tag{A.3.2}$$

Taking into account the property (A.1.8), we obtain

$$g_\lambda(S|x_0) \;=\; \frac{\rho_\lambda(x|x_0)}{\rho_\lambda(x|S)}. \tag{A.3.3}$$

This result indicates that the Laplace transform of the first passage time p.d.f. can be computed once the Laplace transform of the transition p.d.f. is known. But even when this function is not known, it is possible to derive a relationship between the moments of T under certain circumstances, as we will show now.

Let us assume that $X(t)$ is a stationary diffusion process, that is, its transition p.d.f. satisfies the backward equation (A.2.20) with $A_n(x_0) = 0$ for $n \geq 3$:

$$\tfrac{\partial}{\partial t}\rho\,(x,t\,|\,x_0) \;=\; A_1(x_0)\tfrac{\partial}{\partial x_0}\rho\,(x,t\,|\,x_0) + \tfrac{1}{2}A_2(x_0)\tfrac{\partial^2}{\partial x_0^2}\rho\,(x,t\,|\,x_0) \tag{A.3.4}$$

(where we have used the fact that $\frac{\partial}{\partial t_0}\rho\,(x,t\,|\,x_0,t_0) = -\frac{\partial}{\partial t}\rho\,(x,t-t_0\,|\,x_0,0)$). Applying the Laplace transform (with respect to t) to both sides of this equation, and taking into account the property (A.1.7), we have

$$\lambda\,\rho_\lambda(x|x_0) - \delta(x-x_0) \;=\; A_1(x_0)\,\tfrac{\partial}{\partial x_0}\rho_\lambda(x|x_0) + \tfrac{1}{2}A_2(x_0)\,\tfrac{\partial^2}{\partial x_0^2}\rho_\lambda(x|x_0). \tag{A.3.5}$$

Using that $x > S > x_0$ and (A.3.3) we obtain

$$\lambda\,g_\lambda(S|x_0) \;=\; A_1(x_0)\,\tfrac{\partial}{\partial x_0}g_\lambda(S|x_0) + \tfrac{1}{2}A_2(x_0)\,\tfrac{\partial^2}{\partial x_0^2}g_\lambda(S|x_0). \tag{A.3.6}$$

This is a second-order ordinary differential equation for the Laplace transform of the first passage time p.d.f., and it provides enough information to derive a relationship between the moments of the first passage time T. As indicated in previous sections, the moment generating function of a non-negative random variable X with p.d.f. f is related to the Laplace transform of f through $M(\lambda) := \mathbb{E}[e^{\lambda X}] = \mathcal{L}\{f\}(-\lambda)$. If we denote by $M_\lambda(S|x_0)$ the moment generating function of the

first passage time T (which is non-negative), we can write

$$-\lambda\, M_\lambda(S|x_0) \;=\; A_1(x_0)\,\tfrac{\partial}{\partial x_0} M_\lambda(S|x_0) + \tfrac{1}{2} A_2(x_0)\,\tfrac{\partial^2}{\partial x_0^2} M_\lambda(S|x_0). \tag{A.3.7}$$

Recall that the moments of T can be computed as $\mathbb{E}[T^n] = \left[\frac{\partial^n}{\partial \lambda^n} M_\lambda(S|x_0)\right]_{\lambda=0}$. Since these moments depend on S and on the initial condition x_0, it is more convenient to use the notation $\mu_n(S|x_0) := \mathbb{E}[T^n]$. Thus, differentiating n times with respect to λ at both sides of (A.3.7) and evaluating at $\lambda = 0$ we finally have

$$A_1(x_0)\,\frac{\partial}{\partial x_0}\mu_n(S|x_0) + \frac{1}{2} A_2(x_0)\,\frac{\partial^2}{\partial x_0^2}\mu_n(S|x_0) \;=\; -n\,\mu_{n-1}(S|x_0). \tag{A.3.8}$$

Expression (A.3.8) provides a second-order ordinary differential equation for the moments of the first passage time as a function of the initial condition x_0, with conditions

$$\begin{aligned} &\mu_0(S|x_0) = 1 \qquad \text{for all } x_0, \\ &\mu_n(S|S) = 0 \qquad \text{for } n \geq 1. \end{aligned} \tag{A.3.9}$$

Solving (A.3.8) for $n \geq 1$ requires, however, an additional boundary condition. One possibility is to require that the moments of the first passage time do not change substantially with x_0 when $x_0 \to \pm\infty$ (that is, when the starting point is very far from S):

$$\lim_{x_0 \to \pm\infty} \frac{\partial}{\partial x_0}\mu_n(S|x_0) = 0. \tag{A.3.10}$$

Appendix B

Mean-field theory for networks of leaky integrate-and-fire neurons

B.1 Activity of a LIF neuron with Poisson input

This section is mainly based on chapter 15 of [Feng, 2003], on [Brunel, 2000] and on chapter 8 of [Kampen, 2011].

Preliminary considerations

Let us consider a *leaky integrate-and-fire* (LIF) neuron whose voltage obeys the equation

$$\tau \frac{\mathrm{d}V(t)}{\mathrm{d}t} = -(V(t) - V_L) + \tau I(t), \tag{B.1.1}$$

where τ is a time constant, V_L is the resting potential and $I(t)$ is an external input. Every time V reaches a threshold V_θ, the neuron generates an action potential and the voltage is immediately reset to V_r, where it remains for a resting period τ_r.

Let us assume that the external input is generated from the spikes of a set of K neurons which are connected to the neuron under study through synaptic weights $\{J_i\}_{i=1}^K$. We suppose that an action potential emitted by neuron i induces an instantaneous jump on the voltage V, of magnitude J_i:

$$I(t) = \sum_{i=1}^{K} J_i \sum_j \delta(t - t_i^j), \tag{B.1.2}$$

where $\{t_i^1, t_i^2, \ldots\}$ are the spike times of pre-synaptic neuron i and δ denotes the Dirac delta function.

From now on we consider $V_L = 0$ (this can be done without loss of generality noticing that the shifted variable $W = V - V_L$ follows the same equation with $V_L = 0$). If the spike times of the input neurons were known, the solution to Eqs. (B.1.1), (B.1.2) would be

$$V(t) = \left(V(0) + \sum_{i=1}^{K} J_i \sum_j e^{t_i^j/\tau} \right) e^{-t/\tau}. \tag{B.1.3}$$

We are interested, however, in a situation where the exact firing times of input neurons are just known statistically. The key assumption that we make is that every input neuron fires independently as a Poisson process. We denote by $\{\nu_i\}_{i=1}^K$ the rates of these processes.

Properties of the input

Let us define $N_i(t)$ as the total number of spikes emitted by neuron i before time t. Under the previous assumptions, the instantaneous external input is also a random variable and can be expressed as

$$\begin{aligned} I(t) &= \sum_{i=1}^{K} J_i \lim_{s\to 0} \frac{N_i(t+s) - N_i(t)}{s} \\ &= \sum_{i=1}^{K} J_i \lim_{s\to 0} S_i(t,s), \end{aligned} \tag{B.1.4}$$

where we have introduced the random variable $S_i(t,s) := \frac{N_i(t+s)-N_i(t)}{s}$. Since $N_i(t)$ is a Poisson process with parameter ν_i, the difference $N_i(t+s) - N_i(t)$ follows a Poisson distribution with parameter $\nu_i s$, which has mean and variance $\nu_i s$. Thus,

$$\begin{aligned} \langle S_i(t,s)\rangle &= \tfrac{1}{s}\langle N_i(t+s) - N_i(t)\rangle &&= \nu_i, \\ \mathrm{Var}(S_i(t,s)) &= \tfrac{1}{s^2}\mathrm{Var}(N_i(t+s) - N_i(t)) &&= \tfrac{\nu_i}{s}. \end{aligned} \tag{B.1.5}$$

We want to compute the covariance between $S_i(t,s)$ and $S_i(t',s)$. Let us consider that t and s are fixed and let us interpret $C_{t,s}(t-t') := \mathrm{Cov}(S_i(t,s), S_i(t',s))$ as a function of $t'-t$. We will see that, in the limit $s \to 0$, this function equals $\nu_i\, \delta(t'-t)$.

If $(t, t+s) \cap (t', t'+s) = \varnothing$, $S_i(t,s)$ and $S_i(t',s)$ are independent variables and $C_{t,s}(t-t') = 0$. Let us take t, t' such that $(t, t+s) \cap (t', t'+s) \neq \varnothing$. Without loss of generality we can assume that $t' \geqslant t$. We consider the disjoint union $[t, t'+s] = [t, t') \cup [t', t+s) \cup [t+s, t'+s]$. We can write

$N_i(t+s)-N_i(t) = N_i(t')-N_i(t)+N_i(t+s)-N_i(t')$, in such a way that the variables $N_i(t')-N_i(t)$ and $N_i(t+s)-N_i(t')$ are independent. The same can be done with $N_i(t'+s)-N_i(t')$ so that

$$\begin{aligned} S_i(t,s) &= \tfrac{t'-t}{s}\, S_i(t,t'-t) + \tfrac{t+s-t'}{s}\, S_i(t',t+s-t') \\ S_i(t',s) &= \tfrac{t+s-t'}{s}\, S_i(t',t+s-t') + \tfrac{t'-t}{s}\, S_i(t+s,t'-t), \end{aligned} \tag{B.1.6}$$

where $S_i(t,t'-t)$, $S_i(t',t+s-t')$, $S_i(t+s,t'-t)$ are pairwise independent. Thus,

$$\begin{aligned} C_{t,s}(t-t') &= (\tfrac{t+s-t'}{s})^2 \operatorname{Var}(S_i(t',t+s-t')) \\ &= \tfrac{t+s-t'}{s^2}\nu_i. \end{aligned} \tag{B.1.7}$$

In summary, we obtain

$$C_{t,s}(\tau) = \begin{cases} 0 & \text{if } \tau \in (-\infty,-s)\cup(s,+\infty) \\ \frac{s-\tau}{s^2}\nu_i & \text{if } \tau \in (0,s) \\ \frac{s+\tau}{s^2}\nu_i & \text{if } \tau \in (-s,0). \end{cases} \tag{B.1.8}$$

$C_{t,s}$ is a triangular function whose integral equals ν_i. Its width is $2s$ and its maximum, $C_{t,s}(0) = \frac{\nu_i}{s}$. Therefore, as s approaches 0, the function becomes sharper. In fact,

$$\lim_{s\to 0} C_{t,s}(\tau) = \nu_i\, \delta(\tau). \tag{B.1.9}$$

Going back to Eq. (B.1.4), we can now compute the first- and second-order moments of the input $I(t)$:

$$\begin{aligned} \langle I(t)\rangle &= \sum_{i=1}^{K} J_i\, \nu_i \\ \operatorname{Cov}(I(t),I(t')) &= \sum_{i=1}^{K} J_i^2\, \nu_i\, \delta(t'-t). \end{aligned} \tag{B.1.10}$$

The diffusion approximation

In the context presented in the previous section, we cannot expect to find a deterministic solution to Eq. (B.1.1) because the external stimulation is stochastic. Instead of this, we are interested in finding an evolution law for the transition p.d.f. of the voltage, $\rho(V,t\,|\,V_0,t_0)$. To do so, we first assume that the jumps of V are small in magnitude, that is, $J_i \approx 0$. To compensate for that, we suppose that the number of input neurons K is large. Under such assumptions, the voltage V can be assumed to be a continuous variable. Since the time-covariance of the input is a delta function,

Eq. (B.1.1) defines a Markov process, that is, the probability density at time t conditioned to the state of the system at time $t^* < t$ does not depend on which was the state of the system at times prior to t^*.

We can therefore treat the voltage as a continuous Markov process. As we have shown in Appendix A, the p.d.f. of the voltage then obeys the forward Smolukowski equation (A.2.9):

$$\begin{aligned} \frac{\partial}{\partial t}\rho(V,t\,|\,V_0,t_0) &= \sum_{n=1}^{\infty}\frac{(-1)^n}{n!}\,\frac{\partial^n}{\partial V^n}\left[A_n(V,t)\,\rho(V,t\,|\,V_0,t_0)\right] \\ A_n(V,t) &:= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{\mathbb{R}}(V'-V)^n\,\rho(V',t+\Delta t\,|\,V,t)\,\mathrm{d}V'. \end{aligned} \tag{B.1.11}$$

In order to study Eq. (B.1.11) we have to derive the exact form of the infinitesimal moments $\{A_n(V,t)\}_n$ in our system. Let us take Δt small enough so that the probability of receiving more than one stimulus within a time window of width Δt is negligible. We will assume that a subset of K_E input neurons are excitatory (E) and K_I of them are inhibitory (I), and that all the synaptic weights of the same type are equal: $J_i = J_E$ for $i \in \{1,\dots,K_E\}$ and $J_j = -J_I$ for $j \in \{1,\dots,K_I\}$, where $J_E, J_I > 0$. We assume that every E (I) neuron fires at rate ν_E (ν_I). In this case, within such a temporal window the neuron under consideration might receive

(i) a single excitatory input with probability $K_E\nu_E\Delta t$,

(ii) a single inhibitory input with probability $K_I\nu_I\Delta t$, or

(iii) no inputs with probability $1-(K_E\nu_E+K_I\nu_I)\Delta t$.

Therefore, $\rho(V',t+\Delta t\,|\,V,t)$ can be approximated by

$$\begin{aligned} \rho(V',t+\Delta t\,|\,V,t) \approx\ & [1-(K_E\nu_E+K_I\nu_I)\Delta t]\,\delta(V_0-V') \\ & +K_E\nu_E\Delta t\,\delta(V_1-V')+K_I\nu_I\Delta t\,\delta(V_2-V'), \end{aligned} \tag{B.1.12}$$

where V_0, V_1 and V_2 are the voltages at $t+\Delta t$ if the neuron does not receive any stimulus, if it receives a unique E spike or if it receives a unique I spike within the window $[t,t+\Delta t]$, respectively, provided that $V(t)=V$.

Since in the absence of stimulation the voltage decays as $V(t) = V(t_0)e^{-(t-t_0)/\tau}$, which, for a

sufficiently small time lapse can be approximated as $V(t) \simeq V(t_0)\left(1 - \frac{t-t_0}{\tau}\right)$, we obtain

$$\begin{aligned} V_0 &= V\left(1 - \tfrac{\Delta t}{\tau}\right) \\ V_1 &= \left(V\left(1 - \tfrac{\Delta t_1}{\tau}\right) + J_E\right)\left(1 - \tfrac{\Delta t_2}{\tau}\right) \\ V_2 &= \left(V\left(1 - \tfrac{\Delta t_1}{\tau}\right) - J_I\right)\left(1 - \tfrac{\Delta t_2}{\tau}\right), \end{aligned} \tag{B.1.13}$$

where $\Delta t_1 + \Delta t_2 = \Delta t$. Introducing (B.1.12) in the definition of $A_n(V,t)$ we obtain the following expression for the infinitesimal moments:

$$\begin{aligned} A_n(V,t) &= \lim_{\Delta t \to 0} \tfrac{1}{\Delta t}(V_0 - V)^n \left(1 - (K_E\nu_E + K_I\nu_I)\Delta t\right) \\ &\quad + \lim_{\Delta t \to 0} \left((V_1 - V)^n K_E\nu_E + (V_2 - V)^n K_I\nu_I\right) \\ &= \begin{cases} -\frac{V}{\tau} + K_E J_E \nu_E - K_I J_I \nu_I & \text{if } n = 1 \\ K_E J_E^n \nu_E + (-1)^n K_I J_I^n \nu_I & \text{if } n \geq 2. \end{cases} \end{aligned} \tag{B.1.14}$$

Recall that we needed $J_E, J_I \approx 0$ for the continuous approximation to be correct. Since $A_n(V,t) = \mathcal{O}(J_E^n, J_I^n)$ for $n \geq 2$, the same assumption ensures that the infinitesimal moments tend to zero as J_E^n, J_I^n. If we impose that $K_E\nu_E$ and $K_I\nu_I$ are large enough, it is reasonable to truncate the forward equation (B.1.11) at order two, which produces the so-called *diffusion approximation*:

$$\frac{\partial}{\partial t}\rho(V,t\,|\,V_0,t_0) = \frac{\partial}{\partial V}\left(\frac{V-\mu}{\tau}\,\rho(V,t\,|\,V_0,t_0)\right) + \frac{\sigma^2}{2\tau}\frac{\partial^2}{\partial V^2}\rho(V,t\,|\,V_0,t_0), \tag{B.1.15}$$

where

$$\begin{aligned} \mu &:= (K_E J_E \nu_E - K_I J_I \nu_I)\,\tau \\ \sigma^2 &:= \left(K_E J_E^2 \nu_E + K_I J_I^2 \nu_I\right)\tau. \end{aligned} \tag{B.1.16}$$

Recall from (B.1.10) that the input first- and second-order moments are

$$\begin{aligned} \langle I(t)\rangle &= K_E J_E \nu_E - K_I J_I \nu_I \\ \mathrm{Var}(I(t)) &= K_E J_E^2 \nu_E + K_I J_I^2 \nu_I. \end{aligned} \tag{B.1.17}$$

This indicates that the *drift* term of the Fokker-Planck equation (B.1.15), $A_1 = \frac{-V+\mu}{\tau}$, is related to the deterministic component of the dynamics and its *diffusion* term, $A_2 = \frac{\sigma^2}{\tau}$, depends on the fluctuations of the stochastic input.

Mean firing rate in the stationary state

From now on we will simplify the notation and write $\rho(V,t) = \rho(V,t\,|\,V_0,t_0)$. The Fokker-Planck equation (B.1.15) can be expressed as

$$\begin{aligned} \tfrac{\partial}{\partial t}\rho(V,t) &= -\tfrac{\partial}{\partial V}S(V,t), \\ S(V,t) &= -\tfrac{1}{\tau}(V-\mu)\rho(V,t) - \tfrac{\sigma^2}{2\tau}\tfrac{\partial}{\partial V}\rho(V,t), \end{aligned} \tag{B.1.18}$$

where $S(V,t)$ is the *probability flux* or *probability current* passing through V at time t. Recall that the original voltage dynamics includes a thresholding mechanism: V is reset to V_r every time V crosses a threshold V_θ and it remains at V_r for a time lapse τ_r. These conditions impose some boundary conditions on Eq. (B.1.18), as we will detail now.

First, we say that the neuron emits a spike whenever its voltage reaches V_θ. Therefore, its firing rate at time t, $\nu(t)$, equals the probability flux at $V = V_\theta$ at time t: $\nu(t) = S(V_\theta, t)$. On the other hand, since V cannot take values greater than V_θ, the probability density is zero for $V > V_\theta$. If the probability density were different from zero at $V = V_\theta$, its derivative with respect to V at this point would be infinite and so would be the flux according to Eq. (B.1.18). This would imply that the firing rate is infinite, which makes no sense. Therefore, we need to impose

$$\rho(V_\theta, t) = 0. \tag{B.1.19}$$

We also have to to introduce boundary conditions at $V = -\infty$ to ensure that the density is integrable:

$$\lim_{V\to-\infty}\rho(V,t) = 0, \qquad \lim_{V\to-\infty}V\rho(V,t) = 0. \tag{B.1.20}$$

So far we have not included the reset mechanism. In terms of the probability density, this condition means that the probability mass that crosses the threshold V_θ at time t is re-injected at V_r at time $t+\tau_r$. This introduces a discontinuity in the probability current at the reset potential V_r, in such a way that the difference between the probability current just above V_r and just below V_r at time t is precisely the firing rate at time $t-\tau_r$:

$$S(V_r^+, t) - S(V_r^-, t) = \nu(t-\tau_r). \tag{B.1.21}$$

Finally, we have to impose that the integral of the probability density at time t and the probability

that the neuron is refractory at time t add up to 1:

$$\int_{-\infty}^{V_\theta} \rho(V,t)\,\mathrm{d}V + \int_{t-\tau_r}^{t} \nu(s)\,\mathrm{d}s \;=\; 1. \tag{B.1.22}$$

We are interested in finding stationary solutions to (B.1.18), that is, solutions that are constant in time, taking into account all the previous conditions. We will denote such a solution by $\rho(V)$ and the corresponding flux by $S(V)$. According to (B.1.18), this forces the flux $S(V)$ to be constant within each of the two domains where it is continuous, $(-\infty, V_r)$ and $(V_r, V_\theta]$. In the first domain, its value coincides with the limit

$$\lim_{V\to-\infty} S(V) \;=\; \lim_{V\to-\infty} \left(-\tfrac{1}{\tau}(V-\mu)\rho(V) - \tfrac{\sigma^2}{2\tau}\tfrac{\partial}{\partial V}\rho(V)\right) \;=\; 0 \tag{B.1.23}$$

by virtue of condition (B.1.20). The value of $S(V)$ in the second domain has to coincide with its value at $V = V_\theta$: $S(V_\theta) = \nu$, where ν is the stationary firing rate. Therefore, in the stationary state we have

$$(V-\mu)\rho(V) + \frac{\sigma^2}{2}\frac{\partial}{\partial V}\rho(V) \;=\; -\nu\tau\,\Theta(V-V_r), \tag{B.1.24}$$

where Θ is the heaviside step function ($\Theta(x) = 1$ if $x \geq 0$ and 0 otherwise). Expression (B.1.24) defines an ordinary differential equation for $\rho(V)$ of the form

$$A\dot{x}(v) + B(v)x(v) = C, \tag{B.1.25}$$

where C takes different values depending on the domain. The general solution to (B.1.25) has the form $x(v) = f(v)x_h(v)$, where $x_h(v)$ is a solution to the associated homogeneous equation (i.e., the same equation with $C = 0$) and $f(v)$ is a function to be determined. On the one hand, the homogeneous solution is

$$x_h(v) = \exp\left(-\frac{1}{A}\int_{v_0}^{v} B(u)\mathrm{d}u\right). \tag{B.1.26}$$

On the other hand, the function $f(v)$ satisfies $\dot{f}(v)x_h(v) = C/A$, which implies

$$f(v) = f(v_0) + \frac{C}{A}\int_{v_0}^{v} x_h(u)^{-1}\mathrm{d}u. \tag{B.1.27}$$

Therefore, the general solution to (B.1.25) is

$$x(v) = \left(x(v_0) + \frac{C}{A} \int_{v_0}^{v} x_h(u)^{-1} \mathrm{d}u \right) x_h(v), \tag{B.1.28}$$

where x_h is defined by (B.1.26). Taking into account these observations and going back to our equation, we find

$$\rho(V) = \begin{cases} 2\frac{\nu}{\sigma}\tau \, \exp\left(-\frac{(V-\mu)^2}{\sigma^2}\right) \left(k - \int_{\frac{V_0-\mu}{\sigma}}^{\frac{V-\mu}{\sigma}} \exp(u^2)\, \mathrm{d}u \right) & \text{if } V \in (V_r, V_\theta] \\ 2\frac{\nu}{\sigma}\tau \, \exp\left(-\frac{(V-\mu)^2}{\sigma^2}\right) k' & \text{if } V \in (-\infty, V_r), \end{cases} \tag{B.1.29}$$

where k and k' are constants that have to be determined. The boundary condition (B.1.19) implies

$$k = \int_{\frac{V_0-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} \exp(u^2)\, \mathrm{d}u. \tag{B.1.30}$$

We also must impose that $\rho(V)$ is continuous at $V = V_r$, so

$$k' = \int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} \exp(u^2)\, \mathrm{d}u. \tag{B.1.31}$$

We can express the solution as

$$\rho(V) \;=\; 2\frac{\nu}{\sigma}\tau \, \exp\left(-\frac{(V-\mu)^2}{\sigma^2}\right) \int_{\frac{V_-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} \Theta\left(u - \frac{V_r-\mu}{\sigma}\right) \exp(u^2)\, \mathrm{d}u. \tag{B.1.32}$$

The boundary conditions (B.1.20) and (B.1.21) are automatically satisfied. The last condition (B.1.22) provides an equation for the stationary firing rate ν. Let us define the function

$f(V) = \frac{V-\mu}{\sigma}$. Notice that

$$\begin{aligned}
&\frac{1}{2\nu\tau} \int_{-\infty}^{V_\theta} \rho(V)\,\mathrm{d}V \\
&\quad= \frac{1}{\sigma}\int_{-\infty}^{V_r} \exp\left(-f(V)^2\right)\,\mathrm{d}V \int_{f(V_r)}^{f(V_\theta)} \exp(u^2)\,\mathrm{d}u + \frac{1}{\sigma}\int_{V_r}^{V_\theta} \exp\left(-f(V)^2\right) \int_{f(V)}^{f(V_\theta)} \exp(u^2)\,\mathrm{d}u\,\mathrm{d}V \\
&\quad= \int_{-\infty}^{f(V_r)} \exp\left(-t^2\right)\mathrm{d}t \int_{f(V_r)}^{f(V_\theta)} \exp(u^2)\,\mathrm{d}u + \int_{f(V_r)}^{f(V_\theta)} \exp\left(-t^2\right) \int_{t}^{f(V_\theta)} \exp(u^2)\,\mathrm{d}u\,\mathrm{d}t \\
&\quad= \int_{f(V_r)}^{f(V_\theta)} \exp(u^2)\,\mathrm{d}u \int_{-\infty}^{f(V_r)} \exp\left(-t^2\right)\mathrm{d}t + \int_{f(V_r)}^{f(V_\theta)} \exp(u^2) \int_{f(V_r)}^{u} \exp\left(-t^2\right)\,\mathrm{d}t\,\mathrm{d}u \\
&\quad= \int_{f(V_r)}^{f(V_\theta)} \exp(u^2) \int_{-\infty}^{u} \exp\left(-t^2\right)\mathrm{d}t\,\mathrm{d}u \\
&\quad= \frac{\sqrt{\pi}}{2}\int_{f(V_r)}^{f(V_\theta)} \exp(u^2)\,\operatorname{erfc}(-u)\,\mathrm{d}u,
\end{aligned} \tag{B.1.33}$$

where $\operatorname{erfc}(x) := \frac{2}{\sqrt{\pi}}\int_x^\infty \exp(-t^2)\,\mathrm{d}t$ is the complementary error function. Inserting (B.1.33) into condition (B.1.22) and isolating the stationary firing rate ν we obtain

$$\nu = \left(\tau_r + \tau\sqrt{\pi}\int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} \exp(u^2)\,\operatorname{erfc}(-u)\,\mathrm{d}u\right)^{-1}. \tag{B.1.34}$$

Recall that μ and σ are the mean and the standard deviation of the total input received by the neuron under study within a time window of length τ (see Eq. (B.1.16)).

Inter-spike intervals

The coefficient of variation (CV) of the inter-spike interval (ISI) of a neuron is the ratio between the standard deviation of the ISI and its mean (taking averages over time, for example). It therefore provides a measure of the variability of the spiking process. It is possible to derive an analytical formula for the CV of a neuron under the hypotheses of the previous sections, as we will show now.

The inter-spike interval of a neuron whose voltage obeys the stochastic equation (B.1.1) is precisely the first passage time T of the process $V(t)$ to hit the threshold V_θ from the initial state $V = V_r$.

Therefore, the moments $\{\mu_n\}_n$ of T obey the ordinary differential equation (A.3.8) that we derived in Appendix A:

$$A_1(V)\,\frac{\partial}{\partial V}\mu_n(V_\theta|V) + \frac{1}{2}A_2(V)\,\frac{\partial^2}{\partial V^2}\mu_n(V_\theta|V) \;=\; -n\,\mu_{n-1}(V_\theta|V). \tag{B.1.35}$$

Substituting the infinitesimal moments by their expressions (B.1.14), Eq. (B.1.35) becomes

$$\frac{\mu - V}{\tau}\,\frac{\partial}{\partial V}\,\mu_n(V_\theta|V) + \frac{\sigma^2}{2\tau}\,\frac{\partial^2}{\partial V^2}\,\mu_n(V_\theta|V) \;=\; -n\,\mu_{n-1}(V_\theta|V), \tag{B.1.36}$$

where

$$\begin{aligned} \mu &= (K_E J_E \nu_E - K_I J_I \nu_I)\,\tau \\ \sigma^2 &= (K_E J_E^2 \nu_E + K_I J_I^2 \nu_I)\,\tau \end{aligned} \tag{B.1.37}$$

as before. We also have to impose the conditions

$$\begin{aligned} &\mu_0(V_\theta|V) = 1 \qquad \text{for all } V, \\ &\mu_n(V_\theta|V_\theta) = 0 \qquad \text{for } n \geq 1. \end{aligned} \tag{B.1.38}$$

As mentioned before, solving (B.1.36) requires an additional boundary condition, and one possibility is to require that the moments do not vary substantially with V when $V \to -\infty$:

$$\lim_{V\to-\infty} \frac{\partial}{\partial V}\mu_n(V_\theta|V) = 0. \tag{B.1.39}$$

We will solve Eq. (B.1.36) with conditions (B.1.38) and (B.1.39).

Let us assume that we have solved the equation for the case $n-1$ and we want to solve it for n. Defining $f(V) = \frac{\partial}{\partial V}\mu_n(V_\theta|V)$, Eq. (B.1.36) has the form

$$f'(V) + g(V)f(V) = h(V) \tag{B.1.40}$$

with

$$\begin{aligned} g(V) &= 2\,\tfrac{\mu - V}{\sigma^2} \\ h(V) &= -2\,\tfrac{\tau n}{\sigma^2}\,\mu_{n-1}(V_\theta|V). \end{aligned} \tag{B.1.41}$$

The solution to (B.1.40) is

$$f(V) = e^{\left(\frac{\mu-V}{\sigma}\right)^2}\left(C + \int_{V_0}^{V} h(u)e^{-\left(\frac{\mu-u}{\sigma}\right)^2}\,\mathrm{d}u\right), \tag{B.1.42}$$

where V_0 is arbitrary and C is a constant which depends on V_0. Taking $V_0 = -\infty$, we have

$$\frac{\partial}{\partial V}\mu_n(V_\theta|V) = e^{\left(\frac{\mu-V}{\sigma}\right)^2}\left(C + \int_{-\infty}^{V} h(u)e^{-\left(\frac{\mu-u}{\sigma}\right)^2}\,\mathrm{d}u\right). \tag{B.1.43}$$

Condition (B.1.39) imposes $C = 0$, so

$$\frac{\partial}{\partial V}\mu_n(V_\theta|V) = e^{\left(\frac{\mu-V}{\sigma}\right)^2}\int_{-\infty}^{V} h(u)e^{-\left(\frac{\mu-u}{\sigma}\right)^2}\,\mathrm{d}u. \tag{B.1.44}$$

$\mu_n(V_\theta|V)$ is obtained simply by integrating $\frac{\partial}{\partial V}\mu_n(V_\theta|V)$:

$$\mu_n(V_\theta|V) = \mu_n(V_\theta|V_r) + \int_{V_r}^{V}\frac{\partial}{\partial v}\mu_n(V_\theta|v)\,\mathrm{d}v. \tag{B.1.45}$$

The boundary condition (B.1.38) imposes

$$\mu_n(V_\theta|V_r) = -\int_{V_r}^{V_\theta}\frac{\partial}{\partial v}\mu_n(V_\theta|v)\,\mathrm{d}v. \tag{B.1.46}$$

Therefore,

$$\begin{aligned}\mu_n(V_\theta|V) &= -\int_{V}^{V_\theta}\tfrac{\partial}{\partial v}\mu_n(V_\theta|v)\,\mathrm{d}v \\ &= -\int_{V}^{V_\theta} e^{\left(\frac{\mu-v}{\sigma}\right)^2}\int_{-\infty}^{v} h(u)e^{-\left(\frac{\mu-u}{\sigma}\right)^2}\,\mathrm{d}u\,\mathrm{d}v \\ &= 2\tau n\int_{\frac{V-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} e^{t^2}\int_{-\infty}^{t}\mu_{n-1}(V_\theta|\sigma s+\mu)\,e^{-s^2}\,\mathrm{d}s\,\mathrm{d}t.\end{aligned} \tag{B.1.47}$$

To find an analytical expression for the CV of ISIs we need only to compute $\mu_1(V_\theta|V_r)$ and $\mu_2(V_\theta|V_r)$, that is, the first and second-order moments of the first passage time when the process departs from the reset potential V_r. Taking into account that $\mu_0(V_\theta|\cdot) = 1$, the case $n = 1$ gives

$$\mu_1(V_\theta|V_r) = 2\tau\int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} e^{t^2}\int_{-\infty}^{t} e^{-s^2}\,\mathrm{d}s\,\mathrm{d}t = \tau\sqrt{\pi}\int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} e^{t^2}\mathrm{erfc}(-t)\,\mathrm{d}t. \tag{B.1.48}$$

Notice that this result is consistent with the expression that we derived for the stationary firing rate (B.1.34): the average time spent for the voltage to go from V_r to V_θ is precisely the inverse of the firing rate minus the refractory period.

The case $n = 2$ gives

$$\begin{aligned}
\mu_2(V_\theta|V_r) &= 4\tau^2\sqrt{\pi}\int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} e^{t^2}\int_{-\infty}^{t} e^{-s^2}\int_{s}^{\frac{V_\theta-\mu}{\sigma}} e^{u^2}\mathrm{erfc}(-u)\,\mathrm{d}u\,\mathrm{d}s\,\mathrm{d}t \\
&= 2\tau^2\pi\int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} e^{t^2}\int_{-\infty}^{\frac{V_\theta-\mu}{\sigma}} e^{u^2}\mathrm{erfc}(-u)\,\mathrm{erfc}(-\min(u,t))\,\mathrm{d}u\,\mathrm{d}t \\
&= 2\tau^2\pi\int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} e^{t^2}\int_{-\infty}^{t} e^{u^2}\mathrm{erfc}(-u)^2\,\mathrm{d}u\,\mathrm{d}t \\
&\quad +2\tau^2\pi\int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} e^{t^2}\,\mathrm{erfc}(-t)\int_{t}^{\frac{V_\theta-\mu}{\sigma}} e^{u^2}\mathrm{erfc}(-u)\,\mathrm{d}u\,\mathrm{d}t.
\end{aligned} \tag{B.1.49}$$

The second integral in last row of the right-hand side of (B.1.49) is

$$\begin{aligned}
I_2 &= 2\tau^2\pi\int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} e^{t^2}\,\mathrm{erfc}(-t)\int_{t}^{\frac{V_\theta-\mu}{\sigma}} e^{u^2}\mathrm{erfc}(-u)\,\mathrm{d}u\,\mathrm{d}t \\
&= 2\tau^2\pi\int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} e^{t^2}\,\mathrm{erfc}(-t)\left(\int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} e^{u^2}\mathrm{erfc}(-u)\,\mathrm{d}u - \int_{\frac{V_r-\mu}{\sigma}}^{t} e^{u^2}\mathrm{erfc}(-u)\,\mathrm{d}u\right)\mathrm{d}t \\
&= 2\tau^2\pi\left(\int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} e^{t^2}\,\mathrm{erfc}(-t)\,\mathrm{d}t\right)^2 - 2\tau^2\pi\int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} e^{u^2}\,\mathrm{erfc}(-u)\int_{u}^{\frac{V_\theta-\mu}{\sigma}} e^{t^2}\mathrm{erfc}(-t)\,\mathrm{d}t\,\mathrm{d}u \\
&= 2\tau^2\pi\left(\int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} e^{t^2}\,\mathrm{erfc}(-t)\,\mathrm{d}t\right)^2 - I_2,
\end{aligned} \tag{B.1.50}$$

which implies

$$I_2 = \tau^2\pi\left(\int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} e^{t^2}\,\mathrm{erfc}(-t)\,\mathrm{d}t\right)^2 = \mu_1(V_\theta|V_r)^2. \tag{B.1.51}$$

Combining (B.1.49) and (B.1.51) we get

$$\mu_2(V_\theta|V_r) = \mu_1(V_\theta|V_r)^2 + 2\tau^2\pi\int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} e^{t^2}\int_{-\infty}^{t} e^{u^2}\mathrm{erfc}(-u)^2\,\mathrm{d}u\,\mathrm{d}t. \tag{B.1.52}$$

This finally allows us to compute the CV of ISIs:

$$\mathrm{CV}^2 = \frac{\mu_2(V_\theta|V_r) - \mu_1(V_\theta|V_r)^2}{\left(\tau_r + \mu_1(V_\theta|V_r)\right)^2} = 2\tau^2\pi\nu^2\int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} e^{t^2}\int_{-\infty}^{t} e^{u^2}\mathrm{erfc}(-u)^2\,\mathrm{d}u\,\mathrm{d}t, \tag{B.1.53}$$

where $\nu = (\tau_r + \mu_1(V_\theta|V_r))^{-1}$ is the stationary firing rate.

B.2 A network of sparsely and homogeneously connected LIF neurons

In the previous sections we have considered the case of an isolated neuron which receives inputs from an external population. We suppose now that our neuron is embedded in a network of spiking neurons, each of which follows the dynamics defined before. What so far was the external input to the neuron now will be a recurrent input from the population itself. More precisely, we will suppose that the network has N_E excitatory (E) and N_I inhibitory (I) neurons and that each neuron receives inputs from K_E E and K_I I pre-synaptic cells from the network, which are chosen at random, and also from a pool of K_{ext} external E neurons that are supposed to be different from neuron to neuron but that fire at a constant rate ν_{ext}. We assume that in-degrees are large but connectivity is sparse ($1 \ll K_E, K_I \ll N$). This implies that the number of shared pre-synaptic neurons to any neuronal pair is negligible, so we can assume that inputs to different neurons are not correlated beyond correlations induced by time-varying firing rates. The weights from E, I and external populations are equal and they take values J_E, $-J_I$ and J_{ext}, respectively, which are small in magnitude ($J_E, J_I, J_{\text{ext}} \ll V_\theta$). Additionally, we need to assume that the spike statistics of all the neurons are approximately those of Poisson processes.

In this homogeneous scenario, the entire population can be studied using the previous formalism, where now $\rho(V,t)$ can be interpreted as a distribution of voltages over the neuronal population (more precisely, over those neurons that are not refractory) at time t. In the simplest case in which the parameters of the dynamics of E and I neurons are the same, all the neurons can be considered equivalent and the stationary firing rate ν satisfies

$$\begin{aligned} \nu &= \left(\tau_r + \tau\sqrt{\pi} \int_{\frac{V_r-\mu}{\sigma}}^{\frac{V_\theta-\mu}{\sigma}} \exp(u^2)\, \text{erfc}(-u)\, \mathrm{d}u \right)^{-1} \\ \mu &= \left[(K_E J_E - K_I J_I)\nu + K_{\text{ext}} J_{\text{ext}} \nu_{\text{ext}} \right] \tau \\ \sigma^2 &= \left[(K_E J_E^2 + K_I J_I^2)\nu + K_{\text{ext}} J_{\text{ext}}^2 \nu_{\text{ext}} \right] \tau, \end{aligned} \tag{B.2.1}$$

which constitutes a self-consistent equation for ν. In the case in which the two populations have different properties, the same arguments can be followed to obtain a system of two self-consistent equations for their stationary firing rates.

Appendix C

Spectral density of random matrices

In this section we present some results about the spectrum of random matrices that are used in Chapter 3. We do not provide their proofs here, for they depend on a large body of results. We only present the main definitions and theorems. Theorems 7-10 are presented in [Tao, 2013], and Theorem 11 is given in [Aljadeff et al., 2015a].

Definition 1. (Convergence of random variables)

A sequence $\{Z_n\}_n$ *of random variables converges* in probability *towards a random variable* Z *if for all* $\varepsilon > 0$

$$\lim_{n\to\infty} \mathrm{P}\left(|Z_n - Z| > \varepsilon\right) = 0.$$

A sequence $\{Z_n\}_n$ *of random variables converges* almost surely *towards a random variable* Z *if*

$$\mathrm{P}\left(\lim_{n\to\infty} Z_n = Z\right) = 1.$$

Almost sure convergence implies convergence in probability.

Definition 2. (i.i.d. random matrix)

An i.i.d. random matrix is an $n\times n$ *matrix* $\mathbf{X}_n$ *whose entries are independent identically distributed random variables with zero mean and unit variance. The distribution of the entries of* $\mathbf{X}_n$ *is known as the* atom distribution *of* $\mathbf{X}_n$.

Definition 3. (Spectral radius)

Let $\mathbf{A}$ *be an* $n \times n$ *matrix with eigenvalues* $\lambda_1, \cdots, \lambda_n$*. The spectral radius* $r(\mathbf{A})$ *of* $\mathbf{A}$ *is*

$$r(\mathbf{A}) = \sup_{1 \leq j \leq n} |\lambda_j|.$$

Definition 4. (Operator norm)

Let V *and* W *be two normed vector spaces and let* $\mathbf{A}$ *be a linear operator* $\mathbf{A} : V \to W$*. We define the operator norm of* $\mathbf{A}$ *as*

$$\|\mathbf{A}\|_{\text{OP}} = \sup_{\|v\|=1} \|\mathbf{A}v\|.$$

$\mathbf{A}$ *is said to be continuous if* $\|\mathbf{A}\| < \infty$*.*

Definition 5. (Empirical spectral distribution)

Given an i.i.d. random matrix $\mathbf{X}_n$*, its empirical spectral distribution* $\mu(\mathbf{X}_n)$ *is the probability measure*

$$\mu(\mathbf{X}_n) := \frac{1}{n} \sum_{j=1}^{n} \delta(\lambda_j), \tag{C.0.1}$$

where $\{\lambda_j\}_{j=1}^n$ *are the eigenvalues of* $\mathbf{X}_n$*.*

It is important to notice that since $\mathbf{X}_n$ is random, $\mu(\mathbf{X}_n)$ is also random.

Definition 6. *Given a random matrix* $\mathbf{X}_n$*, we say that* $\mu(\mathbf{X}_n)$ *converges in probability (resp. almost surely) to another probability measure* μ *on the complex plane* $\mathbb{C}$ *if for every smooth, compactly supported function* $F : \mathbb{C} \to \mathbb{C}$*, the sequence of (complex-valued) random variables* $\left\{ \int_{\mathbb{C}} F \, \mathrm{d}\mu(\mathbf{X}_n) \right\}_n$ *converges in probability (resp. almost surely) to* $\int_{\mathbb{C}} F \, \mathrm{d}\mu$*.*

Theorem 7. (Circular law for i.i.d. matrices)

Let $\mathbf{X}_n$ *be an i.i.d. random matrix. Then* $\mu\left(\frac{1}{\sqrt{n}}\mathbf{X}_n\right)$ *converges almost surely to the uniform measure* μ_c *on the unit disk.*

Theorem 8. (No outliers for i.i.d. random matrices)

Let $\mathbf{X}_n$ *be an i.i.d. random matrix whose atom distribution has finite fourth moment. Then the sequence of spectral radiuses* $\left\{ r\left(\frac{1}{\sqrt{n}}\mathbf{X}_n\right) \right\}_n$ *converges almost surely to 1 as* $n \to \infty$*.*

Notice that Theorem 8 is not a mere consequence of Theorem 7. The fact that a sequence of random measures converges to the uniform measure on the unit disk is compatible with the presence of isolated eigenvalues outside of the disk for every n. These *outliers* do not play a role in defining the limit measure. However, Theorem 8 states that when the fourth moment of the atom distribution is finite, there are no significant outliers to the circular law. As we will see next, random matrices can be perturbed in such a way that the resultant matrices do present significant outliers even when the limit measure is still given by the circular law.

Theorem 9. (Circular law for low rank perturbation of i.i.d. matrices)

Let $\mathbf{X}_n$ be an i.i.d. random matrix and, for each n, let $\mathbf{C}_n = (C_{ij}^n)_{i,j}$ be a $n \times n$ deterministic matrix with rank $o(n)$ such that

$$\|\mathbf{C}_n\|_F \;:=\; \sqrt{\sum_{i,j=1}^{n} |C_{ij}^n|^2} \;=\; \mathcal{O}(n^{1/2}).$$

Then, $\mu\left(\frac{1}{\sqrt{n}}\mathbf{X}_n + \mathbf{C}_n\right)$ converges almost surely to the circular measure μ_c.

Theorem 10. (Outliers for small low rank perturbation of i.i.d. random matrices)

Let $\mathbf{X}_n$ be an i.i.d. random matrix whose atom distribution has finite fourth moment. For each n, let $\mathbf{C}_n$ be a deterministic $n \times n$ matrix with rank $\mathcal{O}(1)$ and operator norm $\mathcal{O}(1)$. Let $\varepsilon > 0$ and suppose that for all sufficiently large n there are no eigenvalues of $\mathbf{C}_n$ in the band $\{z \in \mathbb{C} : 1+\varepsilon < |z| < 1+3\varepsilon\}$ and there are j_n eigenvalues $\lambda_1(\mathbf{C}_n), \cdots, \lambda_{j_n}(\mathbf{C}_n)$ for some $j_n = \mathcal{O}(1)$ in the region $\{z \in \mathbb{C} : |z| \geq 1+3\varepsilon\}$. Let us denote by k_n the number of eigenvalues of $\mathbf{M}_n := \frac{1}{\sqrt{n}}\mathbf{X}_n + \mathbf{C}_n$ in the region $\{z \in \mathbb{C} : |z| \geq 1+2\varepsilon\}$. Then, the random sequence $\{k_n - j_n\}_n$ converges almost surely to 0. Moreover, for n sufficiently large, these eigenvalues can be labeled so that $\lambda_i(\mathbf{M}_n) = \lambda_i(\mathbf{C}_n) + o(1)$ for each $i \in \{1, \cdots, j_n\}$.

Notice that the almost surely convergence of $\{k_n - j_n\}_n$ means that for almost every random outcome of $\{\mathbf{X}_n\}_n$, there exists n_0 (which might depend on the outcome) such that for $n > n_0$, $k_n = j_n$. Thus, it makes sense to compare the outliers of $\mathbf{M}_n$ and $\mathbf{C}_n$ for n sufficiently large as stated in Theorem 10.

A natural question to ask is what happens with the spectrum of matrices whose entries are independent but not necessarily i.i.d. In this regard, Aljadeff et al. [2015a] provided a result

which concerns matrices whose entries are arranged in blocks in such a way that entries in the same block have the same variance.

Theorem 11. (Spectrum of block structured matrices)

Let $\mathbf{X}_n = (X_{ij}^n)_{i,j}$ be an i.i.d. random matrix whose atom distribution has finite fourth moment. Let $D \in \mathbb{N}$ and let $g = (g_{cd})_{c,d}$ be a $D \times D$ deterministic matrix with real, positive entries. Let $\alpha_1, \cdots, \alpha_n \in \mathbb{R}^+$ be such that $\sum_{d=1}^{D} \alpha_d = 1$. We define

$$c_i = \left\{ c \in \{1, \cdots, D\} : \frac{i}{n} \in \left(\sum_{d=1}^{c-1} \alpha_d, \sum_{d=1}^{c} \alpha_d \right] \right\} \quad \text{for } i \in \{1, \cdots, n\}$$

and we denote by $\mathbf{G} = (G_{cd})_{c,d}$ the $D \times D$ deterministic matrix given by

$$G_{cd} = \alpha_d \, g_{cd}^2.$$

Let $\mathbf{Y}_n = (Y_{ij}^n)_{i,j}$ be an $n \times n$ random matrix whose entries are

$$Y_{ij}^n = g_{c_i c_j} X_{ij}^n.$$

Then, $\mu\left(\frac{1}{\sqrt{n}}\mathbf{Y}_n\right)$ converges almost surely to a deterministic measure μ which is radially symmetric and whose support has radius $\sqrt{r(\mathbf{G})}$. Furthermore, the spectral radius of $\frac{1}{\sqrt{n}}\mathbf{Y}_n$ converges almost surely to $\sqrt{r(\mathbf{G})}$ as $n \to \infty$.

This theorem extends the results concerning random matrices with unit variance to matrices whose entries have different variances as long as the entries are organized into a fixed and finite number of blocks with the same variance. The main result now is that the spectrum is still confined in a disk but the radius of this disk depends on the variances. In particular, the radius is given by the square root of the spectral radius of the matrix $\mathbf{G}$, which is defined as the weighted matrix of the variances of the different blocks. It is important to notice that the eigenvalues of $\mathbf{G}$ are also eigenvalues of $\frac{1}{n}\mathbf{V}_n$, where $V_n = (V_{ij}^n)_{i,j}$ is the $n \times n$ matrix of all the variances: $V_{ij}^n = g_{c_i c_j}^2$. In fact, the eigenvalues of $\frac{1}{n}\mathbf{V}_n$ are those of $\mathbf{G}$ plus zeroes. This means that the spectral radius of $\mathbf{G}$ and $\frac{1}{n}\mathbf{V}_n$ coincide. Therefore, Theorem 11 states that the spectral radius of $\frac{1}{\sqrt{n}}\mathbf{Y}_n$ and $\frac{1}{\sqrt{n}}\sqrt{r(\mathbf{V}_n)}$ coincide almost surely when $n \to \infty$.

An important difference with respect to the i.i.d. case, although not specified here, is that the

limit measure of Theorem 11 is no longer uniform in the disk of radius $\sqrt{r(\mathbf{G})}$. The work in [Aljadeff et al., 2015a] explicitly provides an analytical formula for this measure, which we have not included here because we have not used it in the context of the present thesis.

Bibliography

J. Aljadeff, D. Renfrew, and M. Stern. Eigenvalues of block structured asymmetric random matrices. *Journal of Mathematical Physics*, 56(10):103502, Oct. 2015a.

J. Aljadeff, M. Stern, and T. Sharpee. Transition to chaos in random networks with cell-type-specific connectivity. *Physical Review Letters*, 114(8):088101, Feb. 2015b.

J. Aljadeff, D. Renfrew, M. Vegué, and T. O. Sharpee. Low-dimensional dynamics of structured random networks. *Physical Review E*, 93(2):022302, Feb. 2016.

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

D. J. Amit and N. Brunel. Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, 7(3):237–252, May 1997b.

M. Bastian, S. Heymann, and M. Jacomy. Gephi: An Open Source Software for Exploring and Manipulating Networks. In *Third International AAAI Conference on Weblogs and Social Media*, Mar. 2009.

R. Beltramo, G. D'Urso, M. D. Maschio, P. Farisello, S. Bovetti, Y. Clovis, G. Lassi, V. Tucci, D. D. P. Tonelli, and T. Fellin. Layer-specific excitatory circuits differentially control recurrent network dynamics in the neocortex. *Nature Neuroscience*, 16(2):227, Feb. 2013.

D. D. Bock, W.-C. A. Lee, A. M. Kerlin, M. L. Andermann, G. Hood, A. W. Wetzel, S. Yurgenson, E. R. Soucy, H. S. Kim, and R. C. Reid. Network anatomy and in vivo physiology of visual cortical neurons. *Nature*, 471(7337):177–182, Mar. 2011.

N. Brunel. Dynamics of Sparsely Connected Networks of Excitatory and Inhibitory Spiking Neurons. *Journal of Computational Neuroscience*, 8(3):183–208, May 2000.

N. Brunel. Is cortical connectivity optimized for storing information? *Nature Neuroscience*, 19(5):749–755, May 2016.

N. Brunel and V. Hakim. Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. *Neural Computation*, 11(7):1621–1671, Oct. 1999.

E. Bullmore and O. Sporns. The economy of brain network organization. *Nature Reviews Neuroscience*, 13(5):336, May 2012.

S. R. y. Cajal. *Comparative Study of the Sensory Areas of the Human Cortex*. Theclassics Us, Sept. 2013. Google-Books-ID: BQi7ngEACAAJ.

D. J. Chalmers. *The Character of Consciousness*. Oxford Scholarship Online, 2010.

R. Chaudhuri and I. Fiete. Computational principles of memory. *Nature Neuroscience*, 19(3):394, Mar. 2016.

M. M. Churchland, B. M. Yu, J. P. Cunningham, L. P. Sugrue, M. R. Cohen, G. S. Corrado, W. T. Newsome, A. M. Clark, P. Hosseini, B. B. Scott, D. C. Bradley, M. A. Smith, A. Kohn, J. A. Movshon, K. M. Armstrong, T. Moore, S. W. Chang, L. H. Snyder, S. G. Lisberger, N. J. Priebe, I. M. Finn, D. Ferster, S. I. Ryu, G. Santhanam, M. Sahani, and K. V. Shenoy. Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nature Neuroscience*, 13(3):369, Mar. 2010.

M. C. Crair, D. C. Gillespie, and M. P. Stryker. The Role of Visual Experience in the Development of Columns in Cat Visual Cortex. *Science*, 279(5350):566–570, Jan. 1998.

G. Deco and E. Hugues. Neural Network Mechanisms Underlying Stimulus Driven Variability Reduction. *PLOS Computational Biology*, 8(3):e1002395, Mar. 2012.

W. Denk and H. Horstmann. Serial Block-Face Scanning Electron Microscopy to Reconstruct Three-Dimensional Tissue Nanostructure. *PLOS Biol*, 2(11):e329, Oct. 2004.

J. C. Eccles, M. Ito, and J. Szentágothai. *The Cerebellum as a Neuronal Machine*. Springer-Verlag, New York, 1967.

A. Einstein. *Mis ideas y opiniones*. Editorial Bosch, Barcelona, 1985.

M. Ercsey-Ravasz, N. T. Markov, C. Lamy, D. C. Van Essen, K. Knoblauch, Z. Toroczkai, and H. Kennedy. A Predictive Network Model of Cerebral Cortical Connectivity Based on a Distance Rule. *Neuron*, 80 (1):184–197, Oct. 2013.

P. Erdős and A. Rényi. On Random Graphs I. *Publicationes Mathematicae (Debrecen)*, 6:290–297, 1959.

J. Feng, editor. *Computational Neuroscience: A Comprehensive Approach*. Chapman and Hall/CRC, Boca Raton, 1st edition, Oct. 2003.

E. Fino and R. Yuste. Dense inhibitory connectivity in neocortex. *Neuron*, 69(6):1188–1203, Mar. 2011.

E. Fino, A. M. Packer, and R. Yuste. The Logic of Inhibitory Connectivity in the Neocortex. *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry*, 19(3):228–237, June 2013.

E. Gal, M. London, A. Globerson, S. Ramaswamy, M. W. Reimann, E. Muller, H. Markram, and I. Segev.

Rich cell-type-specific network topology in neocortical microcircuitry. *Nature Neuroscience*, 20(7):1004, July 2017.

J. Ginibre. Statistical Ensembles of Complex, Quaternion, and Real Matrices. *Journal of Mathematical Physics*, 6(3):440–449, Mar. 1965.

V. L. Girko. Circular Law. *Theory of Probability & Its Applications*, 29(4):694–706, Jan. 1985.

K. M. Hagihara, T. Murakami, T. Yoshida, Y. Tagawa, and K. Ohki. Neuronal activity is not required for the initial formation and maturation of visual selectivity. *Nature Neuroscience*, 18(12):1780, Dec. 2015.

K. D. Harris and G. M. G. Shepherd. The neocortical circuit: themes and variations. *Nature Neuroscience*, 18(2):170, Feb. 2015.

S. L. Hill, Y. Wang, I. Riachi, F. Schürmann, and H. Markram. Statistical connectivity provides a sufficient foundation for specific functional connectivity in neocortical neural microcircuits. *Proceedings of the National Academy of Sciences of the United States of America*, 109(42):E2885–E2894, Oct. 2012.

A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 117(4):500–544, Aug. 1952.

S. B. Hofer, T. D. Mrsic-Flogel, T. Bonhoeffer, and M. Hubener. Experience leaves a lasting structural trace in cortical circuits. *Nature*, 457(7227):313–317, Jan. 2009.

S. B. Hofer, H. Ko, B. Pichler, J. Vogelstein, H. Ros, H. Zeng, E. Lein, N. A. Lesica, and T. D. Mrsic-Flogel. Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. *Nature Neuroscience*, 14(8):1045–1052, Aug. 2011.

C. Holmgren, T. Harkany, B. Svennenfors, and Y. Zilberter. Pyramidal cell communication within local networks in layer 2/3 of rat neocortex. *The Journal of Physiology*, 551(Pt 1):139–153, Aug. 2003.

T. Hromádka, M. R. DeWeese, and A. M. Zador. Sparse Representation of Sounds in the Unanesthetized Auditory Cortex. *PLOS Biology*, 6(1):e16, Jan. 2008.

Y. Hu, J. Trousdale, K. Josić, and E. Shea-Brown. Local paths to global coherence: Cutting networks down to size. *Physical Review E*, 89(3):032802, Mar. 2014.

D. H. Hubel and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1):106–154.2, Jan. 1962.

J. S. Isaacson and M. Scanziani. How Inhibition Shapes Cortical Activity. *Neuron*, 72(2):231–243, Oct. 2011.

X. Jiang, S. Shen, C. R. Cadwell, P. Berens, F. Sinz, A. S. Ecker, S. Patel, and A. S. Tolias. Principles of connectivity among morphologically defined cell types in adult neocortex. *Science*, 350(6264):aac9462, Nov. 2015.

N. G. V. Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier, Aug. 2011.

E. R. Kandel. The Molecular Biology of Memory Storage: A Dialogue Between Genes and Synapses. *Science*, 294(5544):1030–1038, Nov. 2001.

N. Kasthuri, K. J. Hayworth, D. R. Berger, R. L. Schalek, J. A. Conchello, S. Knowles-Barley, D. Lee, A. Vázquez-Reina, V. Kaynig, T. R. Jones, M. Roberts, J. L. Morgan, J. C. Tapia, H. S. Seung, W. G. Roncal, J. T. Vogelstein, R. Burns, D. L. Sussman, C. E. Priebe, H. Pfister, and J. W. Lichtman. Saturated Reconstruction of a Volume of Neocortex. *Cell*, 162(3):648–661, July 2015.

L. C. Katz and J. C. Crowley. Development of cortical circuits: Lessons from ocular dominance columns. *Nature Reviews Neuroscience*, 3(1):34, Jan. 2002.

D. Kleinfeld, A. Bharioke, P. Blinder, D. D. Bock, K. L. Briggman, D. B. Chklovskii, W. Denk, M. Helmstaedter, J. P. Kaufhold, W.-C. A. Lee, H. S. Meyer, K. D. Micheva, M. Oberlaender, S. Prohaska, R. C. Reid, S. J. Smith, S. Takemura, P. S. Tsai, and B. Sakmann. Large-Scale Automated Histology in the Pursuit of Connectomes. *The Journal of Neuroscience*, 31(45):16125–16138, Sept. 2011.

H. Ko, S. B. Hofer, B. Pichler, K. A. Buchanan, P. J. Sjöström, and T. D. Mrsic-Flogel. Functional specificity of local synaptic connections in neocortical networks. *Nature*, 473(7345):87–91, May 2011.

A. A. Koulakov, T. Hromádka, and A. M. Zador. Correlated Connectivity and the Distribution of Firing Rates in the Neocortex. *Journal of Neuroscience*, 29(12):3685–3694, Mar. 2009.

I. D. Landau, R. Egger, V. J. Dercksen, M. Oberlaender, and H. Sompolinsky. The Impact of Structural Heterogeneity on Excitation-Inhibition Balance in Cortical Networks. *Neuron*, 92(5):1106–1121, Dec. 2016.

J.-V. Le Bé and H. Markram. Spontaneous and evoked synaptic rewiring in the neonatal neocortex. *Proceedings of the National Academy of Sciences of the United States of America*, 103(35):13214–13219, Aug. 2006.

G. W. Leibniz. *Monadología*. Biblioteca Nueva, Madrid, 2001.

A. Litwin-Kumar and B. Doiron. Formation and maintenance of neuronal assemblies through synaptic plasticity. *Nature Communications*, 5:5319, Nov. 2014.

A. Litwin-Kumar, K. D. Harris, R. Axel, H. Sompolinsky, and L. F. Abbott. Optimal Degrees of Synaptic Connectivity. *Neuron*, 93(5):1153–1164.e7, Mar. 2017.

L. Margulis and D. Sagan. *Microcosmos. Cuatro mil millones de años de evolución desde nuestros ancestros microbianos*. Tusquets Editores, Barcelona, 1995.

N. T. Markov, M. Ercsey-Ravasz, D. C. V. Essen, K. Knoblauch, Z. Toroczkai, and H. Kennedy. Cortical High-Density Counterstream Architectures. *Science*, 342(6158):1238406, Nov. 2013.

N. T. Markov, M. M. Ercsey-Ravasz, A. R. Ribeiro Gomes, C. Lamy, L. Magrou, J. Vezoli, P. Misery, A. Falchier, R. Quilodran, M. A. Gariel, J. Sallet, R. Gamanut, C. Huissoud, S. Clavagnier, P. Giroud, D. Sappey-Marinier, P. Barone, C. Dehay, Z. Toroczkai, K. Knoblauch, D. C. Van Essen, and H. Kennedy. A Weighted and Directed Interareal Connectivity Matrix for Macaque Cerebral Cortex. *Cerebral Cortex (New York, NY)*, 24(1):17–36, Jan. 2014.

H. Markram, J. Lübke, M. Frotscher, A. Roth, and B. Sakmann. Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. *The Journal of Physiology*, 500(Pt 2):409–440, Apr. 1997.

H. Markram, M. Toledo-Rodriguez, Y. Wang, A. Gupta, G. Silberberg, and C. Wu. Interneurons of the neocortical inhibitory system. *Nature Reviews Neuroscience*, 5(10):793, Oct. 2004.

H. Markram, E. Muller, S. Ramaswamy, M. W. Reimann, M. Abdellah, C. A. Sanchez, A. Ailamaki, L. Alonso-Nanclares, N. Antille, S. Arsever, G. A. A. Kahou, T. K. Berger, A. Bilgili, N. Buncic, A. Chalimourda, G. Chindemi, J.-D. Courcol, F. Delalondre, V. Delattre, S. Druckmann, R. Dumusc, J. Dynes, S. Eilemann, E. Gal, M. E. Gevaert, J.-P. Ghobril, A. Gidon, J. W. Graham, A. Gupta, V. Haenel, E. Hay, T. Heinis, J. B. Hernando, M. Hines, L. Kanari, D. Keller, J. Kenyon, G. Khazen, Y. Kim, J. G. King, Z. Kisvarday, P. Kumbhar, S. Lasserre, J.-V. Le Bé, B. R. C. Magalhães, A. Merchán-Pérez, J. Meystre, B. R. Morrice, J. Muller, A. Muñoz-Céspedes, S. Muralidhar, K. Muthurasa, D. Nachbaur, T. H. Newton, M. Nolte, A. Ovcharenko, J. Palacios, L. Pastor, R. Perin, R. Ranjan, I. Riachi, J.-R. Rodríguez, J. L. Riquelme, C. Rössert, K. Sfyrakis, Y. Shi, J. C. Shillcock, G. Silberberg, R. Silva, F. Tauheed, M. Telefont, M. Toledo-Rodriguez, T. Tränkler, W. Van Geit, J. V. Díaz, R. Walker, Y. Wang, S. M. Zaninetta, J. DeFelipe, S. L. Hill, I. Segev, and F. Schürmann. Reconstruction and Simulation of Neocortical Microcircuitry. *Cell*, 163(2):456–492, Oct. 2015.

A. Mason, A. Nicoll, and K. Stratford. Synaptic transmission between individual pyramidal neurons of the rat visual cortex in vitro. *The Journal of Neuroscience*, 11(1):72–84, Jan. 1991.

L. Mazzucato, A. Fontanini, and G. L. Camera. Dynamics of Multistable States during Ongoing and Evoked Cortical Activity. *Journal of Neuroscience*, 35(21):8214–8231, May 2015.

V. B. Mountcastle. Modality and Topographic Properties of Single Neurons of Cat's Somatic Sensory Cortex. *Journal of Neurophysiology*, 20(4):408–434, July 1957.

V. B. Mountcastle. The columnar organization of the neocortex. *Brain: A Journal of Neurology*, 120 (Pt 4):701–722, Apr. 1997.

M. Newman. The Structure and Function of Complex Networks. *SIAM Review*, 45(2):167–256, Jan. 2003.

D. Q. Nykamp. A mathematical framework for inferring connectivity in probabilistic neuronal networks. *Mathematical Biosciences*, 205(2):204–251, Feb. 2007.

D. Q. Nykamp, D. Friedman, S. Shaker, M. Shinn, M. Vella, A. Compte, and A. Roxin. Mean-field

equations for neuronal networks with arbitrary degree distributions. *Physical Review. E*, 95(4-1):042323, Apr. 2017.

D. H. O'Connor, S. P. Peron, D. Huber, and K. Svoboda. Neural Activity in Barrel Cortex Underlying Vibrissa-Based Object Localization in Mice. *Neuron*, 67(6):1048–1061, Sept. 2010.

M. Okun and I. Lampl. Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nature Neuroscience*, 11(5):535–537, May 2008.

M. Okun, N. A. Steinmetz, L. Cossell, M. F. Iacaruso, H. Ko, P. Barthó, T. Moore, S. B. Hofer, T. D. Mrsic-Flogel, M. Carandini, and K. D. Harris. Diverse coupling of neurons to populations in sensory cortex. *Nature*, 521(7553):511, Apr. 2015.

S. Ostojic. Two types of asynchronous activity in networks of excitatory and inhibitory spiking neurons. *Nature Neuroscience*, 17(4):594, Apr. 2014.

S. Pajevic and D. Plenz. Efficient Network Reconstruction from Dynamical Cascades Identifies Small-World Topology of Neuronal Avalanches. *PLOS Comput Biol*, 5(1):e1000271, Jan. 2009.

R. Perin, T. K. Berger, and H. Markram. A synaptic organizing principle for cortical neuronal groups. *Proceedings of the National Academy of Sciences*, 108(13):5419–5424, Mar. 2011.

V. Pernice, B. Staude, S. Cardanobile, and S. Rotter. How Structure Determines Correlations in Neuronal Networks. *PLOS Computational Biology*, 7(5):e1002059, May 2011.

V. Pernice, M. Deger, S. Cardanobile, and S. Rotter. The relevance of network micro-structure for neural dynamics. *Frontiers in Computational Neuroscience*, 7, 2013.

K. Rajan and L. F. Abbott. Eigenvalue spectra of random matrices for neural networks. *Physical Review Letters*, 97(18):188104, Nov. 2006.

S. Ramaswamy, J.-D. Courcol, M. Abdellah, S. R. Adaszewski, N. Antille, S. Arsever, G. Atenekeng, A. Bilgili, Y. Brukau, A. Chalimourda, G. Chindemi, F. Delalondre, R. Dumusc, S. Eilemann, M. E. Gevaert, P. Gleeson, J. W. Graham, J. B. Hernando, L. Kanari, Y. Katkov, D. Keller, J. G. King, R. Ranjan, M. W. Reimann, C. Rössert, Y. Shi, J. C. Shillcock, M. Telefont, W. Van Geit, J. Villafranca Diaz, R. Walker, Y. Wang, S. M. Zaninetta, J. DeFelipe, S. L. Hill, J. Muller, I. Segev, F. Schürmann, E. B. Muller, and H. Markram. The neocortical microcircuit collaboration portal: a resource for rat somatosensory cortex. *Frontiers in Neural Circuits*, page 44, 2015.

M. W. Reimann, J. G. King, E. B. Muller, S. Ramaswamy, and H. Markram. An algorithm to predict the connectome of neural microcircuits. *Frontiers in Computational Neuroscience*, page 120, 2015.

A. Renart, J. de la Rocha, P. Bartho, L. Hollender, N. Parga, A. Reyes, and K. D. Harris. The asynchronous state in cortical circuits. *Science (New York, N.Y.)*, 327(5965):587–590, Jan. 2010.

L. M. Ricciardi. *Lecture Notes in Biomathematics | Diffusion Processes and Related Topics in Biology*, volume 14. Springer-Verlag, Heidelberg, 1st edition, 1977.

M. Rigotti, O. Barak, M. R. Warden, X.-J. Wang, N. D. Daw, E. K. Miller, and S. Fusi. The importance of mixed selectivity in complex cognitive tasks. *Nature*, 497(7451):585, May 2013.

A. Roxin. The Role of Degree Distribution in Shaping the Dynamics in Networks of Sparsely Connected Spiking Neurons. *Frontiers in Computational Neuroscience*, 5, Mar. 2011.

A. Roxin, N. Brunel, D. Hansel, G. Mongillo, and C. v. Vreeswijk. On the Distribution of Firing Rates in Networks of Cortical Neurons. *Journal of Neuroscience*, 31(45):16217–16226, Nov. 2011.

A. J. Sadovsky and J. N. MacLean. Scaling of Topologically Similar Functional Modules Defines Mouse Primary Auditory and Somatosensory Microcircuitry. *The Journal of Neuroscience*, 33(35):14048–14060, Aug. 2013.

C. Schmeltzer, A. H. Kihara, I. M. Sokolov, and S. Rüdiger. Degree Correlations Optimize Neuronal Network Sensitivity to Sub-Threshold Stimuli. *PLOS ONE*, 10(6):e0121794, June 2015.

E. Schrödinger. *What is Life? With Mind and Matter and Autobiographical Sketches*. Cambridge University Press, Cambridge, 2010.

H. Setareh, M. Deger, C. C. H. Petersen, and W. Gerstner. Cortical Dynamics in Presence of Assemblies of Densely Connected Weight-Hub Neurons. *Frontiers in Computational Neuroscience*, 11, 2017.

M. N. Shadlen and W. T. Newsome. Noise, neural codes and cortical organization. *Current Opinion in Neurobiology*, 4(4):569–579, Aug. 1994.

G. Silberberg, A. Gupta, and H. Markram. Stereotypy in neocortical microcircuits. *Trends in Neurosciences*, 25(5):227–230, May 2002.

W. R. Softky and C. Koch. The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *Journal of Neuroscience*, 13(1):334–350, Jan. 1993.

H. Sompolinsky, A. Crisanti, and H. J. Sommers. Chaos in Random Neural Networks. *Physical Review Letters*, 61(3):259–262, July 1988.

S. Song, P. J. Sjöström, M. Reigl, S. Nelson, and D. B. Chklovskii. Highly Nonrandom Features of Synaptic Connectivity in Local Cortical Circuits. *PLOS Biol*, 3(3):e68, Mar. 2005.

O. Sporns, C. J. Honey, and R. Kötter. Identification and Classification of Hubs in Brain Networks. *PLOS ONE*, 2(10):e1049, Oct. 2007.

O. Stetter, D. Battaglia, J. Soriano, and T. Geisel. Model-Free Reconstruction of Excitatory Neuronal Connectivity from Calcium Imaging Signals. *PLOS Comput Biol*, 8(8):e1002653, Aug. 2012.

J. Szentágothai. The Ferrier Lecture, 1977: The Neuron Network of the Cerebral Cortex: A Functional

Interpretation. *Proceedings of the Royal Society of London B: Biological Sciences*, 201(1144):219–248, May 1978.

T. Tao. Outliers in the spectrum of iid matrices with bounded rank perturbations. *Probability Theory and Related Fields*, 155(1-2):231–263, Feb. 2013.

T. Tao and V. Vu. Random matrices: universality of ESDs and the circular law. *The Annals of Probability*, 38(5):2023–2065, 2010.

N. M. Timme, S. Ito, M. Myroshnychenko, S. Nigam, M. Shimono, F.-C. Yeh, P. Hottowy, A. M. Litke, and J. M. Beggs. High-Degree Neurons Feed Cortical Computations. *PLOS Computational Biology*, 12 (5):e1004858, May 2016.

C. Tomm, M. Avermann, C. Petersen, W. Gerstner, and T. P. Vogels. Connection-type-specific biases make uniform random network models consistent with cortical recordings. *Journal of Neurophysiology*, 112(8):1801–1814, Oct. 2014.

J. T. Trachtenberg, B. E. Chen, G. W. Knott, G. Feng, J. R. Sanes, E. Welker, and K. Svoboda. Long-term in vivo imaging of experience-dependent synaptic plasticity in adult cortex. *Nature*, 420(6917):788–794, Dec. 2002.

C. van Vreeswijk and H. Sompolinsky. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science (New York, N.Y.)*, 274(5293):1724–1726, Dec. 1996.

J. C. Vasquez, A. R. Houweling, and P. Tiesinga. Simultaneous stability and sensitivity in model cortical networks is achieved through anti-correlations between the in- and out-degree of connectivity. *Frontiers in Computational Neuroscience*, 7, Nov. 2013.

M. Vegué, R. Perin, and A. Roxin. On the Structure of Cortical Microcircuits Inferred from Small Sample Sizes. *Journal of Neuroscience*, 37(35):8498–8510, Aug. 2017.

X.-J. Wang and H. Kennedy. Brain structure and dynamics across scales: in search of rules. *Current Opinion in Neurobiology*, 37:92–98, Apr. 2016.

Y. Wang, H. Markram, P. H. Goodman, T. K. Berger, J. Ma, and P. S. Goldman-Rakic. Heterogeneity in the pyramidal network of the medial prefrontal cortex. *Nature Neuroscience*, 9(4):534–542, Apr. 2006.

M. Xue, B. V. Atallah, and M. Scanziani. Equalizing excitation-inhibition ratios across visual cortical neurons. *Nature*, 511(7511):596–600, July 2014.

L. Zhao, B. Beverlin, T. Netoff, and D. Q. Nykamp. Synchronization from Second Order Network Connectivity Statistics. *Frontiers in Computational Neuroscience*, 5, July 2011.

Y. Zuo, G. Yang, E. Kwon, and W.-B. Gan. Long-term sensory deprivation prevents dendritic spine loss in primary somatosensory cortex. *Nature*, 436(7048):261–265, July 2005.