Modelling the brain: Elementary components to explain ensemble functions

Egidio D’Angelo(1)(2) and Claudia Gandini Wheeler-Kingshott(1)(2)(3)

(1) Department of Brain and Behavioral Sciences, University of Pavia
Via Forlanini 6, I-27100 Pavia, Italy

(2) Brain Connectivity Centre, C. Mondino National Neurological Institute
Via Mondino 2, I-27100 Pavia, Italy

(3) NMR Research Unit, Department of Neuroinflammation, Queen Square MS Centre
UCL Institute of Neurology, London WC1N 3BG, UK

received 30 December 2016

Summary. — The brain is organized on multiple levels. The lowest meaningful one pertains to the molecular realm, followed by subcellular structures like the synapses, by cells like the neurons, and by microcircuits, mesocircuits and large-scale circuit assemblies. This stratified structure has so far hampered the interpretation of brain functions in terms of elementary electrochemical events occurring in the membranes of neurons and synapses. Each organization level is governed by emerging rules that do not simply account for the summation of events at the lower levels but require the understanding of highly non-linear interactions occurring in complex feed-forward and feed-back loops. Moreover, various forms of plasticity can persistently modify the neural circuits and their connections depending on the interactions of the organism with the environment. The brain appears thus to operate as a complex adaptive dynamical system and interpreting its function requires understanding the time-dependent evolution of multiple local activities and their rewiring during behaviour. While experimental evidence is instrumental to any further consideration on how the brain might operate, interpreting its multiscale organization in mechanistic terms requires the development of appropriate models. In this work we will illustrate how low-level representations of neuronal activity, intermediate level large-field networks and high-level connectomics can be used to explain how ensemble brain functions might emerge from elementary neuronal components.
1. – Modelling the brain: the nature of the issue

Understanding the brain is a core issue for Neuroscience and this concept has been recently casted into a theoretical framework [1]. Moreover, a surge of interest has addressed the possibility of modelling brain functions [2-5]. But what would be the gain of having a model for understanding the brain? At the very least, a model would instantiate Richard Feynmann’s reflection that “What I cannot create, I do not understand”. In fact, there are specific and compelling reasons indicating that constructing a model is essential toward the understanding of how the brain works. And, as a consequence, a brain model would foster the reproduction of functions in artificial machines and would provide new cues for curing brain diseases. But this is not all what a model of the brain would mean in this context, there is much more.

1’1. Brain organization and function: the complexity issue. – The brain is the most fascinating and probably the most complex structure of the universe. With its $10^{12}$ neurons and $10^{15}$ synapses, the human brain generates an internal representation of the world and self, controls behaviours, perceives sensations, commands movements, feels emotions, generates thoughts, stores and retrieves memories and makes all of this conscious. The number of publications on brain structure and function has shown a tremendous increase in the last years(1) but still we do not understand how the brain works. Or, to be more precise, the fundamental question on how the highest brain functions arise from molecular properties of neurons remains unanswered. Why? There are several reason to consider, but first of all we have to face the issue of brain complexity.

(1) One can count the articles in Pubmed (www.pubmed.gov) which contain the word “neuron” in their titles or abstracts. Impressively, the number increases from about 1000 in the 70s to about 25000 in 2015! However, clearly this search does not capture all the neuroscience articles, in particular those related to brain imaging or psychophysics. When the search is expanding to include “neuron OR neural OR neuronal OR brain”, this number almost quadruples (e.g. from about 25000 to 100000 articles in 2015). Interestingly, about one every five of these papers also contains the word “model” (e.g. about articles 18000 in 2015). The current pace of scientific publication in neuroscience is so high that it is becoming almost impossible to keep up.
Complexity depends on the number of interactions and possible states assumed by a system and not by the number of elements only, and this applies to any physical system and to brain too [6]. The brain (2% of body mass) is certainly much more difficult to understand than the muscles (40% of body mass), for which we can provide a direct explanation of force generation based on their molecular properties and mechanical arrangement! What is somehow misleading is that the brain is made of principal cells (the neurons), supporting cells (the glial cells) and blood vessel cells, so that in this respect it does not differ from other body structures, with whom it shares fundamental biological and pathological mechanisms. After all, neurons are cells and the molecular networks controlling the membrane, cytoplasmic and nuclear functions of neurons are no more complex than those of other cells in principle. Thus, the reason of our failure to understand the brain does not seem to reside in molecular and cellular aspects (though these play a critical role, as explained below) but rather in the complexity of neuronal interactions and on their multi-layered architecture. These issues will now be considered in turn.

1'2. The multiscale organization of the brain. – Unanimated matter is best conceived as being made of elementary components, e.g. a large collection of atoms or molecules, whose properties have an immediate reflection into those of the structure they constitute. For example, in a star, fundamental laws of physics predict how atomic properties generate mass, volume, temperature, light, gravity, radiation and so forth. In other words, astrophysicists can jump directly from elementary to ensemble properties and can therefore understand how a star is generated and evolves. This clearly does not apply to the brain, in which self-organization of biomolecules and biostructures generates a multi-layered system (fig. 1). We recognize at least 8 distinct anatomo-functional levels, which can also be referred as to microscale (1-6), mesoscale (7), and macroscale (8).

1) Bio-molecules (DNA, enzymes, etc.)
2) Simple subcellular structures (bio-membranes, calcium stores, etc.)
3) Complex subcellular structures (synapses, dendritic spines, axon hillock, etc.)
4) Aggregates of specialized subcellular structures (multi-synaptic microcircuits, synaptic glomeruli, etc.)
5) Cells (neurons, glial cells, blood-vessel cells)
6) Local multicellular aggregates (local neuronal microcircuits, including glial components and blood vessels)
7) Interconnected microcircuits (e.g. thalamo-cortical circuit, other major brain structures)
8) Large-scale networks (the brain)

Each one of these levels has its own complexity and can be investigated through specific techniques. Normally, the properties of one level can be used to predict those of the higher hierarchical level or can be demonstrated to descend from those of the lower hierarchical level, but longer jumps (e.g. from molecules to brain) are unpractical and unconstrained. This is one of the reasons why a multiscale model of the brain is so important.
Fig. 1. – The multiscale organization of the brain. The figure illustrates the multiple levels characterizing brain organization, from molecules to cells, circuits and behaviour. Reprinted from [7].

Historically, mostly for practical and methodological reasons, disciplines have evolved to deal with these different organization levels and this has caused a fragmentation of actions rather than an advantage towards the final goal of understanding the brain. Nor clarity was added by the diatribe on the brain-mind problem, which has been dividing philosophers around the concepts of dualism and monism. The problem originates from the observation that, while brain and mind are related to one other, the brain is material while the mind is immaterial, leading to various conceptual solutions dating back to Aristotle, Plato, Kant and Descartes, just to mention a few main ones(2). This issue has been reinterpreted by neuroscientists in seminal papers and, since the 50’s [8, 9], more and more importance has been given to the fact that traceable brain activity is causative for mental function and dysfunction. Modern neurophilosophy is telling us that dualism is not likely to provide the solution but rather it supports the concept that mental functions derive from the brain, in a way that reflects the ensemble activity of the underlying structures (a huge impulse in this direction has recently been given by MRI and connectomics, as explained below [10-12]). Clearly one may speculate whether an appropriate model that reflects the multi-layered structure of the brain could eventually generate high-order functions — like behaviour and thought — and eventually consciousness [13].

1.3. The properties of molecules. – Knowing that the brain is made of molecules does not help much by itself to explain its functioning, unless the relationship of molecules with higher level phenomena is known. The problem is that this relationship appears to be elusive when considering the huge number of molecules and the complexity of their interactions and functions. The importance of molecules could emerge only if they were embedded into detailed molecular-cellular level computational models [14-17]. For example, one may reconstruct a model of the molecular interactions deriving from the activation of a membrane receptor, with activation of intracellular transduction cascades and production of second messengers, that would eventually modulate effectors like enzymes, structural proteins, ionic channels, membrane receptors and even the genome. These mechanisms, in turn, would generate mechanistic predictions about phenomena like neuromodulation, synaptic plasticity, homeostasis, neurodegeneration, neural growth, etc. The level of representation of molecular properties can go down to the atomic level, for example using molecular dynamics models. Unfortunately, beside their attractiveness, models based on explicit reconstructions of molecular structure-function-dynamic relationships are probably too complex and laborious at present to be used in the context of large-scale brain simulations.

An important aspect of molecular properties, that could bring about relevant consequences once molecular properties are accounted for, is the emergence of stochasticity that would lead beyond a deterministic interpretation of brain function and behaviour [18]. However, at present, kinetic descriptions of chemical transformations based on deterministic differential equations are commonly used to describe the underlying processes. An example of this is the classical Hodgkin-Huxley model used for modelling the molecular properties of ionic channels [19,20].

1.4. The activity of neurons and microcircuits. – Brain activity is based on the continuous exchange of information between neurons (e.g., see [7], and appendix A for biophysical foundations). Neurons are specialized cells generating electrical signals across their membrane and chemical signals at the synapses. In essence, the membrane of neurons is polarized due to the establishment of electrochemical potentials. This causes a negative resting membrane potential between −60 and −70 mV depending on the neuron type, although some neurons have an oscillating membrane potential and are never strictly at rest. Whether resting or oscillating, the neuron initial state can be perturbed giving rise to an action potential. This is a rapid (∼1 ms) membrane potential transition from negative to positive potentials and back, which activates in an all-or-none fashion when a threshold around −40 mV is crossed. Sophisticated mechanisms can regulate the process of action potential generation forming patterns that represent the neuronal signals. The action potentials travel at high speed along the axons to reach the synapses. Here, complex molecular mechanisms allow releasing chemical neurotransmitters that reach the nearby neurons generating a postsynaptic potential. When this potential crosses the threshold, new action potentials are generated in the postsynaptic neuron and information flows through the neuronal chain.

The processes of action potential generation are highly non-linear with respect to time and voltage, as are those of synaptic transmission and signal transduction [19]. Moreover, in addition to synapses that excite the postsynaptic neuron, there are those that inhibit it. Finally, neurotransmission can activate biochemical transduction systems also independently from ionic current control across the neuronal membrane.
Neurons assemble into local aggregates, called local microcircuits. These are formed by $10^4$–$10^5$ neurons that generate intricate connection patterns. While neurons are the elementary cellular components, it is at the level of microcircuits that the fundamental brain computations take place. Neurons, by receiving about $10^3$ synapses each (between $10^1$–$10^6$, depending on the neuron type), have the capacity of integrating a huge amount of information and to generate action potential patterns that reflect the non-linear transformation that neuronal processing operates. Local microcircuits in turn exploit neuronal processing to perform parallel and distributed computations on the inputs, that are themselves coming from other microcircuits. At present, the function and dynamics of signal processing in local neuronal networks can be precisely resolved using “realistic” bottom-up modelling strategies (see below).
Fig. 3. – The multiscale experimental approach. Different experimental approaches are used at different brain scales. Patch-clamp recordings are used for single neurons (e.g. a Golgi cell; [26]) and the cells are reconstructed using immunofluorescence. Multi-photon confocal microscopy is used for recording multiple neurons simultaneously in neuronal microcircuits (e.g. granule cells; [27]). Long-range connections are reconstructed using MRI tractography (e.g. a cerebello-prefrontal tract; [28]). Brain function and pathology are analyzed using resting-state fMRI (e.g. AD and MCI; [29]).

Local microcircuits perform specific operations on the inputs. The activity of multiple microcircuits can be coordinated and propagated to neighboring connected structures forming integrated systems that are functional units on the mesoscale. An example is the cortico-thalamic circuit, in which cortical microcircuits interact with thalamic microcircuits to form cortico-thalamic loops. Another example is the loop formed by cerebellar cortical microcircuits (or microzones) with deep cerebellar nuclei and inferior olive to generate the cerebellar microcomplex [22] (fig. 2). Beyond this, large-scale circuits involving multiple mesoscale or microscale circuits can interconnect distant brain areas, for example the cerebro-thalamic loops with cerebellum microcomplexes. At the large-scale level, the main problem is to resolve the geometrical organization of these regions (the so-called “connectome”, see below) and to determine how the connectome is related to system function and dynamics [23-25]. How these different organization and functional levels correspond to different experimental approaches is shown in fig. 3.
Fig. 4. – The brain as a complex adaptive system. The brain generates an internal virtual reality (gray shadowing) that is continuously compared to the external reality through the implementation of schemes and sensory feed-back. The changes to the schemes are maintained through plasticity mechanisms determining learning and memory.

1.5. Principles of brain functioning. – At the macroscale, there are several guiding principles that can help understanding how the brain ultimately operates.

The brain operates as an autonomous system modulated by senses (fig. 4). This means that the brain generates an internal virtual representation of reality that is continuously confronted with the signals conveyed by senses. The external inputs are remapped in the internal space, where they undergo complex spatio-temporal transformations, and interfere with the ongoing activity of the brain. Indeed, the brain is never resting and shows internal rhythms attesting the coherence and frequency of underlying neuronal oscillations. Therefore, the brain has to be treated as a dynamical system, and the sole structure-function relationship is insufficient to understand how the brain works and evolves in life [30].

The brain requires continuous tuning to operate. Since there is no way to pre-tune all synapses genetically, information coming from the senses is used to this purpose(3). The way synapses transmit signals and neurons generate action potentials is not fixed and specific mechanisms of synaptic and non-synaptic plasticity are thought to support this function.

(3) With its $10^{12}$ neurons and $10^3$ connections/neuron, and assuming that each connection requires 1 bit, the brain would require $10^{15}$ bits to be tuned at least once. Beside the fact that tuning is continuously reshaped, it is easy to demonstrate that biological systems do not have a way to transmit such information to the progeny. The human genome is made of $3 \times 10^9$ base-pairs, each one occurring in 4 possible configurations corresponding to occurrence of one of the 4 nucleotides (ACGT). Thus, each base-pair contains $\log_2 4 = 2$ bits of information and the whole genome contains $6 \times 10^9$ bits (corresponding to about 5.5 GB), i.e. nearly 3 orders of magnitude below the number of synapses. Thus the genome cannot program all brain connections, neither in case it would be used only once and just for this purpose. The information required for brain wiring must come from the environment.
The brain operates as a predictive machine. This is not an intuitive issue at all. The brain exploits its internal representation to predict future system states and anticipate their occurrence through actions\(^4\). This also allows consciousness to be instantaneous and continuous and movement to be controlled in real-time. The continuous internal activity of the brain provides the reference frame with respect to which all other signals are remapped.

The brain proceeds through the implementation of adaptable schemes (fig. 4). Schemes are based on previous memory and cognitive processing, but then are tuned through sensory feed-back deriving from experience. It should be noted that the brain evolved to allow animals to move and that the motor system implements basic coordination schemes that need, then, to be tuned for the specific environmental conditions and cases of use. The schemes are thus tuned on the basis of active interaction with the environment. Cognition and higher functions can be thought as deriving from this initial design. In order to support this process, appropriate circuits have evolved, for example those involving the cerebellum and the cerebral cortex (see below).

Therefore, at the macroscale, the brain can be conceived as a dynamic adaptive system operating through predictive tunable schemes on the basis of an internal virtual representation of the world and self. Clearly here we have taken a top-down attitude, we have considered what the system does and hinted at how it might operate. Eventually, an appropriate bottom-up model should be able to uncover these emerging system properties.

1.6. Problems descending from complexity and the need for a brain model. – There is a series of drawbacks descending from the framework explained above.

- Difficulty in analyzing microcircuit activity. While single neurons can be rather well investigated and understood in their biophysical and biochemical mechanisms, understanding a connected microcircuit (typically \(10^3\)–\(10^4\) neurons) remains challenging. This is a critical technical issue, requiring the development of new imaging and electrophysiological tools [31,32].

- Structure-function-dynamics relationship not always clear. A common approach used to investigate the brain is that of defining its structure, performing stimulus-response experiments to investigate its functions, and then reconnect these aspects to explain the dynamic behaviour of the system in space and time. However, the relationship among structure, function and dynamics remains unclear in many cases [30].

- Elusive link between brain function and consciousness. Ultimately, the brain generates a virtual reality of which we are aware, a property called consciousness. But the link between brain functioning and consciousness lacks critical experimental measurements [6].

\(^4\) Suppose that a car moves at 100 km/h, i.e. 27.7 m/s. Since the cerebral cortex employs in the best case about 100 ms to elaborate a percept, this means that the time elapsed from car position and when the driver recognizes a turn is about 2.7 m. This is enough to drive the car out of the road. Clearly, the brain needs to anticipate the occurrence of events when it is engaged in sensory-motor loops!
Elusive signal coding strategies and multidimensional mapping. The way the brain remaps and process space and time inside its circuits still creates conceptual problems. These fundamental physical dimensions are encrypted in the neuronal network space and their processing is hard to decipher [33].

Incomplete understanding of stochasticity. The brain has several stochastic processes running inside its circuits, nevertheless we use to treat it as a deterministic machine. The implications of stochasticity in brain processing, such as emerging from molecular level studies, are far from clear [34].

All these elements underline the fact that, in essence, there is no single accepted theory on how the brain works, that could be tested and falsified. The absence of a unified theory for the brain is to be searched in a long-standing ontological issue, the Turing-Goedel theorem, stating that a machine cannot understand another with a similar or higher complexity [35]. But then, is understating the brain possible for humans at all? We believe the answer is yes, we can understand the brain [1], provided that we have a theoretical framework and a model. A model constructed in a way that it is grounded on neuronal biophysics, that reflects accurately the structure, functions and dynamics of brain circuits and that allows generating predictions that can be tested and falsified. Ultimately, the biophysics of neuronal signal processing and the architecture of the brain have to emerge into higher-order functions. Here we will illustrate how such a model is not utopia, but is actually already in fieri[5].

1.7. Realistic modelling: a bottom-up approach to the brain. – A model taking into account biological details through a construction-validation process is called realistic[6]. The approach for reconstructing a model from its components proceeds bottom-up, in contrast to top-down models that anticipate an intuition on how the system works and then elaborate a plausible explanation. The realistic bottom-up approach implements

[5] In Physics, there are several cases in which theory and models have produced strong advancements in the understanding of natural phenomena. Maybe the most dramatic example is provided by the theory of general relativity, which has generated models of the Universe. There are also concrete examples in the field of climate and material science that closely resemble the case of the brain [30]. Numerical weather prediction uses computational models based on physical principles of the atmosphere and oceans to predict the weather based on initial conditions. Manipulating the vast datasets provided by satellites and on-earth observations and performing the complex calculations necessary to modern numerical weather prediction requires some of the most powerful supercomputers in the world. Weather models use systems of differential equations based on the laws of physics, fluid motion, and chemistry, and use a coordinate system, which divides the planet into a 3D grid. Winds, heat transfer, solar radiation, relative humidity, and surface hydrology are calculated within each grid cell, and the interactions with neighboring cells are used to calculate atmospheric properties in the future. The accuracy of numerical predictions is affected by the density and quality of observations used as input and by deficiencies in the numerical models. A critical issue lies in the chaotic nature of the partial differential equations that govern the atmosphere, which are impossible to solve exactly causing error propagation and limiting the extent of predictions to a few days rather than long term predictions. As will become clear in the main text of the manuscript, very similar issues are faced when tackling the brain and similar computational and modelling strategies can be used for brain modelling, although the underlying physics and parameters are obviously different.

[6] Calling these models realistic does not mean that other models are unrealistic! This rather indicates that these models are based on realistic biophysical mechanisms and are therefore biophysically detailed.
BRAIN MODELLING

307

process of reverse engineering, in which construction is based on local rules of elementary interactions, while general rules about the system are extracted a posteriori from the ensemble behaviour of the construction.

For the brain, top-down may be used but the success is limited by the exceeding complexity, the multiscale organization and the overwhelming number of details that make the brain a still enigmatic machine [3, 4, 21, 37]. For example, intuition cannot easily jump from molecules to consciousness, nor it can account for all the elements that could be critical to determine function and dynamics. Therefore, bottom-up approaches appear to be a winning card and are probably essential in the attempt to modelling the brain. In addition, the bottom-up approach has further specific advantages. First, it can incorporate all relevant details of brain function, down to molecular dynamics and up to large-scale connectivity. In this way it naturally implements multiscale architectures. Secondly, it can account for brains typical of different animal species. Once “scaffold” models for neurons and microcircuits are designed, their microscopic parameters and modular connectivity can be modified leading to species-specific variants. In this way, bottom-up modelling also helps addressing evolutionary principles and explaining how different functions emerge from specific neuronal properties and microcircuit organization. Thirdly, it can be improved and updated as soon as new relevant data are provided. Therefore, a bottom-up model co-evolves with experimental research, of which it becomes an inherent component. Finally, since general biophysical and biological rules are used for construction, then the bottom-up strategy can compensate for missing knowledge accelerating the process of system reconstruction. This appears as an essential element of the strategy that can prevent a never-ended collection of pieces of the puzzle. Thus, the fundamental importance and independence of biological discoveries notwithstanding, bottom-up realistic models can easily incorporate novelties and predict missing knowledge, promoting research in critical directions and accelerating the reconstruction of the global picture.

It is in this sense that the model constructed through a realistic bottom-up approach can promote the development of a theory of the brain. A theory that can be updated, tested and falsified. This iterative process promotes new experiments that eventually will allow researchers to improve the model and so forth. The details of this iterative procedure will be explained below.

1.8. Realistic modelling strategies: construction, validation, propagation. – To summarize, we are facing the most complex structure of the Universe but we do not know its “project”, that is actually what we would like to discover. Critical data are missing and system complexity is so high that we will never be able to obtain all the data in a reasonable time. But we have the constructing rules and a dataset sufficient to generate an initial (or scaffold) model of the brain through a bottom-up realistic approach. The general plan and the general organization of the brain can be reconstructed through a process of reverse engineering proceeding through a series of well defined steps:

– In the reconstruction phase, the neurons and brain circuits are reconstructed through a model compensating for missing data.

– The model will then be subjected to validation against experimental data that had not been used for reconstruction.

– Finally the model will be investigated through simulations in order to obtain predictions about the system functional states.
The construction rules encompass the laws of cellular biophysics and of connectivity in neuronal assemblies that have recently been defined in the exemplar reconstruction of the cortical microcolumn [38]. This strategy is waiting for generalization though the reconstruction of other brain microcircuits, specifically those of hippocampus, cerebellum, and basal ganglia.

In order to implement the bottom-up modelling strategy for large neuronal assemblies, supercomputing resources are needed. De facto, the realistic bottom-up strategy is now becoming feasible since supercomputers are reaching the exa-flop scale, providing the computational power needed for large-scale network simulations [5](7). Likewise, neuromorphic computing architecture may in the future transform the way brain simulations are carried out bringing them into hardware and providing at the same time new electronic computing architectures.

Informatics is causing a revolution in the way brain science is developing(8). The large data-sets required for bottom-up brain modelling are becoming available through specialized databases, in which data are collected, curated and organized. Among these, the Allen Institute for Brain Science has developed the Allen Brain Atlas over the last decade, which covers multiscale data from genomics to proteomics, cell types and connectomics(9). Advanced brain atlases bringing whole-brain reconstruction to the subcellular level are being produced combining MRI technologies with histology, electron microscopy, and advanced techniques like knife-edge scanning microscopy [39](10),(11).

Databasing initiatives are also promoting neuron modelling(12). These databases contain the required information and drivers that allow to reconstruct neuronal and microcircuit models through specific modelling platforms(13). Therefore, understanding the brain requires modelling the brain and this in turn requires informatics, databasing and high-performance computing (see also(4)).

This visionary strategy embracing realistic bottom-up brain modelling, supercomputing and neuromorphic hardware, and the implementation/exploitation of large databases, has been elaborated into large-scale projects pioneered by the European Flagship, Human Brain Project (HBP; [38]). Clearly, brain investigation requires big-science and advanced infrastructures fueling at the same time the advancement of science and technology in a virtuous cycle [2, 3, 40–42].

19. The process of network simplification: from micro to macro-circuit models. – In order to scale-up from local microcircuits to models of interconnected brain regions (from microscale to mesoscale and macroscale), some further steps need to be taken. While in principle, given enough computational power and appropriate neuroinformatic strategies, a full simulation of multiple interconnected realistic microcircuits could be achieved, this

---

(7) High-performance computing (HPC) is becoming available through an open-access scheme based on international initiatives like the Partnership for Advanced Computing in Europe (PRACE) and the Neuroscience Gateway (NSG) in USA as well through Cloud Computing.

(8) Several links can be found at https://www.openconnectomeproject.org/links.


(13) NEURON (https://www.neuron.yale.edu/neuron/).
would not help much our intuition towards what the system is doing. Therefore, a first step to make the interpretation of implicit model computations affordable, is to simplify it.

The simplification process is not trivial and should occur under supervised guidance. This means retaining, even after simplification, the fundamental computations and dynamics that are thought to characterize the real system and the realistic model. Simplification involves a top-down process, in which it is important to identify constraints derived from experiments and to decide whether they have to be retained in the simplified model. Thus, a good simplified model should be one that does not introduce arbitrary choices (as it will become clearer later in the article) and is at the same time computationally efficient. A way to achieve these goals is 1) to identify the biological target of any simulation, 2) to identify the properties of neurons that are relevant and need to be retained, and 3) to reproduce them with minimal computational efforts. Examples in this direction have been provided in recent works using the generalized leaky integrate-and-fire models (GLIF: [43, 44]), which allow to represent several aspects of neuronal electroresponsiveness accurately. The switch from bottom-up models to simplified ones is needed in a set of often coexisting cases:

- When the scale of simulations requires representing a huge number of neurons and connections. For example, this approach has been used to model a whole corticothalamic system [45].

- When circuit models have to be embedded into control loops. For example, simplified models should eventually be able to generate realistic microcircuit interactions, allowing simulations into closed-loop systems using simulated neuro-robots [46].

- When circuit models have to be accelerated to real-time performance in order to drive a real robot [47].

- When circuit models have to be transformed into hardware to generate neuromorphic computers [48, 49].

- When a theoretical analysis is needed [50]. Actually, simplified models represent the link between the pure bottom-up strategy enforced through realistic microcircuit reconstruction and the top-down inference from the observation of high-level brain functions. A crucial achievement that exploits the confluence of theories at different scales has been provided by generating neuronal masses and connecting them to investigate the interplay of local microcircuit dynamics and ensemble dynamics. This has allowed to reconnect simplified models to statistical physics and thermodynamics. The computation of entropy and information out of model simulations has allowed developing the concepts of metastable states operating at the edge of chaos to explain the inner physics of brain functioning [23-25, 50, 51].

Therefore, the role of mesoscale/macroscale simplified models is just that of giving a substrate in terms of circuits and mechanisms to the conceptual scheme reported in fig. 5.

10. Wrap-up: an integrated view of the whole brain modelling process. – The general scheme that derives from the considerations above is illustrated in fig. 5. Molecular and cellular neurophysiology provide, through biological recordings, the fundamental observations needed to generate realistic models of neurons and microcircuits that need than
Fig. 5. – The different steps of brain modelling. Note that all required information is originated by biological investigations. However, while realistic modelling proceeds bottom-up, relevant top-down intervention is needed to generate control systems and simplified models. Description to be simplified and embedded into control systems designed to enable behaviour. At the same time, biological recordings provide critical information about the nature and localization of plasticity in microcircuit synapses and neurons. Once the system has been reconstructed, it can be connected to a simulated or a real robot that will allow the circuit interaction with the environment. In this way, by tracking the activity of each underlying component within the microcircuit itself during the ongoing interaction with the environment, it becomes possible to discover the cellular basis for the emergence of behaviour.

Importantly, data generated by these modelled control systems could then be compared to those derived from high-level measurements in vivo (e.g., LFP, fMRI, hd-EEG). As a last step, the properties of the system can be analyzed with tools deriving from the field and neuronal masses approach [23, 51-53] to obtain theoretical insight. It should be noted that, notwithstanding the absolute relevance of biological data, the model is our only way to access information otherwise inaccessible, as the model is able to provide a full set of information at different mechanistic levels far from experimental reach. Examples of all these procedures are provided in sect. 3.

2. – Brain connectivity

2.1. Brain organization and function: the problem of complexity. – In sect. 1 we have highlighted the problem of complexity from a bottom-up point of view. Similarly to the cellular level it is possible to approach the investigation of brain properties at macroscopic level in vivo implying that a top-down approach is also needed. Fundamental information can be obtained using tomographic techniques such as magnetic resonance imaging (MRI) as well as electrophysiological measurements using for example electroencephalography (EEG).
In this section we will focus on MRI, as its contribution to understanding the brain from a top-down level has been amazing and we will refer to other techniques as it becomes necessary. MRI is an in vivo non-invasive way to investigate brain properties from structural, functional and metabolic point of view. The versatility of MRI allows the user to sensitize the measured signal to specific properties of the underlying tissue. What is measured are ensemble properties of the brain, averaged at millimetre scale, which reflect though the cellular properties that determine tissue structure and function (fig. 6).

The arguments of layered complexity addressed at cellular and microcircuits levels are supported even at millimetre scale. In fact, with MRI it is possible to obtain quantitative metrics of local tissue characteristics that reflect microstructural properties such as axonal density, fibre coherence and orientation complexity, myelin and macromolecular water fractions, iron content, and even properties such as tortuosity and mean axon diameter. Recently, development of acquisition strategies report even measurements of g-ratio, i.e. the ration between the inner and outer diameter of an axon, hence giving local tissue information that could be related to signal conduction, showing alterations in diseases such as multiple sclerosis or dementia. Thanks to magnetic resonance spectroscopy (MRS) we can also study average metabolic properties linked to axonal integrity, energy consumption, gliosis to say a few. By employing dedicated hardware, metabolism
and physiological aspects of the brain become more approachable, for example through phosphorous spectroscopy that gives quantifications of ATP and PH, or through sodium imaging where intra and extracellular sodium ions quantification becomes possible at ultra-high field.

But MRI is not only incredibly powerful for the assessment of microstructural and metabolic properties of tissue. It is also able to access information about blood perfusion, quantifying blood volume, blood flow and arrival transit time, opening the view over an expanding range of aspects. Thanks to MRI sensitivity to blood oxygenation level, functional MRI was introduced [54] as a mean of studying brain function. It is now well known that when performing a task, there is a blood oxygenation level dependent (BOLD) signal change of a few % that can be statistically significant and therefore can be used to map which brain areas are responsible for a specific task. By collecting a wealth of results obtained with BOLD functional MRI (fMRI), with EEG or with other imaging methods such as positron emission tomography (PET), magnetoencephalography (MEG) or near infrared spectroscopy (NIRS), we now can clearly map where the sources of many brain functions are in vivo in humans.

All in all, therefore, we have a very sophisticated set of tools for assessing tissue microstructure, metabolism and local function. These tools also provide advanced technical approaches to understand how the brain works and to answer questions like how motion is generated or where does cognition come from. In order to do so, we have to step up a level and start looking at how these brain regions that share consistent micro, meso and macroscopic properties across the human race (and sometimes even across races) interact with each other and are structurally and functionally connected.

Before moving into a more specific discussion of what can be learnt from quantitative MRI of the brain structure and function, it is important to stress that, when analyzing MRI data, one has to take advantage of a cascade of models and computational strategies that affect the outcome of the research. Recent attention has highlighted how structural and functional imaging studies can heavily depend on model assumptions as well as on pre-processing steps implemented in own algorithms or in available software packages [55]. For functional imaging, in particular, statistical modelling is also a source of debatable results, with false positive rates inflated by the wrong assumptions [56,57].

In MRI we have the signal, which is an “integrated” truth of the functional and structural properties of the brain. This “truth” is, indeed, influenced by the biophysical properties of the tissue, at molecular and cellular level, which evolve dynamically over timeframes that span from instantaneous cellular processes to changes lasting milliseconds or more, which are happening over a scale comparable to that of the MRI experiment itself. We can say that the physiological basis of the MRI signals are dynamic, and introduce variability in the results of repeated scans and contribute to group analysis outcomes. Caffeine intake, time of the day for scanning, day of the week even, hydration or dehydration, are all variable that is impossible to control and that influence our brain dynamics and connectivity. Bottom-up models of integrated brain functions are indeed one possible way to cope with this large physiological and methodological variability and dependency on parameterization and assumption in our analysis.

2.2. The multiscale organization of the brain reflected at macroscopic level. – The multi-scale organization of the brain is also reflected at macroscopic level. While looking at microscopic properties, it became apparent that it was possible to start investigating network properties both from a structural and functional point of view. How is this possible? Given that it is not feasible to track the location of each single neuron in vivo
nor it is possible to measure the action potential of a single channel or synapsis, how can we access network properties and investigate how the brain works in vivo? We must start from top-down assumptions. We can measure properties and hypothesize that these properties reflect a known biophysical property.

We can assume for example models of axonal organization to derive metrics that reflect axonal density and verify that what we measure is consistent with biophysical properties. This is a huge problem that the MRI research community is constantly tackling and that may find a solution when bottom-up models will find their way to meet top-down ones. Even harder is to verify functional imaging results, where the best way is to compare multi-modal recording in humans in vivo. An alternative to prove the validity of fMRI is by multi-electrode-array recordings in animals where single or multi unit spikes can be assessed directly and compared with fMRI results.

A similar problem comes when we want to reconstruct fibre pathways and networks based on the sensitivity of MRI to water diffusion in tissue. By introducing signal dephasing associated to positional change (diffusion) of water molecules in the magnetic field, it is possible to probe tissue microstructure as water molecules will be hindered and restricted in their movement by cells membranes. Mapping the probability density function of the water displacement in tissue in vivo, for example, with constrained spherical deconvolution of the signal \[58\], we can reconstruct possible fibre pathways (tractography) connecting regions of the brain. While tracers experiments have confirmed the ability of tracking real fibre pathways in non-human primates \[59\], there is also substantial evidence that these tractograms are affected especially by false positives \[60,61\]. This is due to the intrinsic limitation of diffusion MRI to detect synapsis, distinguish crossing from kissing fibres, and differentiate afferent from efferent fibres from specific brain regions. It is essential therefore that macroscopic networks of specific systems, obtained from MRI diffusion tractography, are supported by a priori hypothesis based on anatomical knowledge.

In parallel, fMRI data not only reveals local changes in BOLD signal associated to specific tasks, but can be analyzed to determine the functional connectivity between regions during the task. This process is a statistical analysis of time series of signal at voxel level under the assumption that regions that are functional connected will oscillate with the same patterns and respond to the task in the same way. The fact that two regions are functionally connected it does not mean that these are also structurally connected as their ability to support the same function could depend on a third party.

Scaling up the question of the brain functional and structural connectivity, one reaches the problem of reconstructing the human connectome (fig. 7), a challenge that was at first proposed by the human connectome project (HCP) (http://www.humanconnectome.org/data/). The overwhelming amount of data (1200 healthy subjects collected between 2012-2015) collected is available to the research community for developing ever-sophisticated tools to read data and contribute to the understanding of brain function.

\[23\]. The activity of neurons and the human connectome. – Mathematical models can come into rescue and help understanding structural and functional connectivity at network level. These models are again top-down as they start from assumptions of how the network may work and may be connected because of the impossibility of direct measures of neuronal connectivity and function in humans in vivo. We should also not forget that fMRI is an indirect measure of function (fig. 8). Between the action potential generation at cellular level and a BOLD response there are several processes at neuronal and synaptic level whose interaction gives raise to the measured changes.
Fig. 7. – Building the human structural connectome. Grey matter can be segmented into sub-regions based on anatomical or functional parcellations that can be considered as nodes of a graph. In the figure, “AAL parcellations” is a color-coded representation of different grey matter regions obtained with the anatomical automatic labelling (AAL), one of the most reliable digital brain atlas. “Tractography streamlines” is a representation of the streamlines obtained using constrained spherical deconvolution (CSD) tractography. Counting streamlines that reach pairs of AAL regions, it is possible to obtain measures of connectivity, shown as connecting edges in the “Connectivity” box of the figure. Graphs are then shown as matrices with AAL regions as columns and rows labels, and connectivity measures as values. Graphs are weighted by the “number of streamlines” between pairs, but also by structural properties of tracts connecting pairs, such as “Fractional anisotropy” or “Mean diffusivity” (courtesy of Thalis Charalambous, UCL, UK and Fulvia Palesi, UniPV, Italy).

A specific structural network can be constructed from nodes and edges, represented into a graph that can then be analyzed mathematically. From graph theory analysis it has been possible to study the brain as a large scale network and to reveal emerging properties such as that of small worldness, i.e. the brain is organized into small worlds clustered around hubs, characterized by short path length and high clustering coefficients [62]. Nodes of the graph are typically associated to grey matter regions, where signal processing takes place, synapsis occur and dynamics are evolving. Edges are identified with white matter tracts that are connecting the nodes. Typical networks are identified in terms of regions that emerge as contributing to specific functions and the white matter tracts that connect these regions. There are several ways to build the connectome though as reported in appendix B, reflecting different properties of the brain.

Interestingly, MRI is sensitive to the synchronized oscillations that pervade the brain even when at rest [63]. As pointed out in sect. 1, the brain is never shut down, but rather in resting or active condition. Based on low-frequency analysis of signal time series acquired in absence of tasks, it is possible to identify a set of networks incorporating areas of the brain statistically similar in their oscillatory patterns. Such networks have been identified as being the core regions for specific functions, implying that the brain has a pre-defined scheme ready to be excited in order to function [64]. Correlations
Fig. 8. – Resting state networks and functional connectivity. Resting state networks (RSNs) obtained by independent component analysis (ICA) of signal fluctuations in time between brain areas. The circular representation of the RSNs shows the AAL anatomical regions that are involved in each RSNs (colour coded) and their functional connectivity as represented by the semi-circular lines. Abbreviations are: auditory network (AN), right (R) and left (L) ventral attention networks (VAN), language network (LN), sensory motor network (SMN), executive control network (ECN), default mode network (DMN), medial visual network (MVN), lateral visual network (LVN), cerebellar network (CBLN), salience network (SN), task positive network (TPN), frontal cortex network (FCN), anterior insular network (AIN), Basal ganglia network (BGN) (courtesy of Gloria Castellazzi, UniPV, Italy).

between regions can define a functional connectivity matrix that represent the functional connectome of the brain, where pairs of nodes are “functionally connected” depending on their functional connectivity. As this relies on group analysis of signal fluctuation in time, the functional connectome is defined at group level (fig. 9).

It is worth considering that resting state functional networks are not completely independent either, as structural and functional connectivity exists between regions belonging to different RSNs. It is therefore possible to investigate the connectivity between functional networks by calculating full and partial correlations between the mean time course signal of each RSN, which can be used as a mean for investigating the top level organization of the brain.

2.4. Toward a theoretical interpretation of brain functions. – The connectomic reconstruction obtained using high-definition MRI, EEG and MEG technologies is fundamental to reconnect local to global brain activities [23, 24]. This allows to reconcile under a single theoretical framework apparently antithetic hypothesis, localizationism and globalism. The first claims that specific functions have specific localization in the brain, and actually fMRI studies support this view by showing that certain areas are activated in relation to specific behaviours, as expected from previous neurophysiological and anatomical studies. The second claims that higher brain functions are distributed and involve multiple brain areas. There is evidence that local and global activities are reconnected through brain dynamics [63, 65], in such a way that global dynamics over distributed brain areas emerge from the local dynamics of each brain area and, at the same time, constraint local dynamics. The system therefore shows circular causality and becomes self-organizing [66].
Fig. 9. – Full and partial correlations of resting state networks. Resting state networks (RSNs) are identified by independent component analysis (ICA) on signal fluctuation of individual voxels. When considering the overall RSNs’ signal behaviour it is possible to determine full and partial correlations between RSNs themselves. The top panel shows one RSN per each column of the graph metric. The graph metric shows the colour coded full and partial functional correlations between pairs of RSNs. The star diagrams in the right panels are showing the partial and full RSNs correlations. Notice how the cerebellum displays both partial and full correlations with the medial visual networks, the basal ganglia and the default mode network, i.e. both sensory and cognitive networks.

Clearly here top-down and bottom-up modelling approaches find a merging point through the concepts of functional connectivity and effective connectivity. The former is the statistic dependence between remote neurophysiological events, and is normally assessed through simple correlations or coherence analysis in fMRI and EEG. The latter is the influence that a system exerts on another, explicitly depending on the underlying model of neuronal dynamics [67]. Thus, functional connectivity can be extracted from connectomics and connectivity matrices, while effective connectivity can be generated using neuronal microcircuit models like those obtained using fields and neuronal masses [23,51,52]. Eventually, the local dynamics generated in local circuits communicate on the large-scale through long-range connections.

In order to understand the interplay among several neuron aggregates, the attractor theory can be used [68]. The aggregates operate as multi-stable attractors that tend to settle around stability points depending on their internal effective connectivity. Interactions between aggregates can set-up complex global dynamics. For example, structured firing fluctuations around a low-frequency equilibrium state lead the aggregates to generate activities resembling those of the resting state fMRI networks. The transitions between multistable attractors are driven by noise and the average uncertainty that a given attractor is associated with, starting from a random condition, provides an estimate of system entropy. Interestingly, when the inter-area connections are too low or too high, there is only one attractor state possible and the entropy is zero. The number of attractors increases for intermediate connectivity strengths along with entropy [23].
This theoretical analysis suggests therefore that, like in real networks, interesting computations occur when the information (entropy) processed by the system is non-zero, as it occurs at an intermediate connectivity state near a phase transition. At this point, the correlation with empirical functional connectivity could be obtained by deconvolving the fMRI signal with the haemodynamic response function to obtain neuronal activity, e.g. using the Ballon-Windkessel model [69, 70]. This comparison actually shows that entropy for real networks in resting-state fMRI is very similar to that obtained from the attractor theory.

In summary, these results suggest that theoretical models can strongly support our understanding of how brain works. It can be envisaged that the availability of more and more precise microcircuit models could substitute neural masses and lead to understanding in great details the relationship between low-level (molecular and cellular) properties and the global dynamics in which they are engaged [63].

3. – Examples from the cerebellar circuit

3’1. The cerebellar circuit and its models: foundations of the issue. – The history of neuroscience is profoundly bound to the cerebellum. With more than 50% of all brain neurons, the cerebellum forms the second major cortical structure of the brain (fig. 2). From the anatomical view point, the cerebellum has not just fostered the generation of the concept of neuron [71] but also of one of the first integrated brain theories, the Motor Learning Theory of David Marr [72]. This theory is remarkable in several respects and makes predictions about the role of the cerebellum in behaviour, hinting towards the neuronal nature of functions [73,74]. The Nobel laureate J.C. Eccles, in his foreword to a seminal book written by Masao Ito [75], wrote

“For me the most significant property of the cerebellar circuitry would be its plastic ability, whereby it can participate in motor learning, that is the acquisition of skills. This immense neuronal machine with the double innervation of Purkinje cells begins to make sense if it plays a key role in motor learning […] it could be optimistically predicted that the manner of operation of the cerebellum in movement and posture would soon be known in principle”.

However, Marr did not consider either the existence of forms of plasticity in addition to long-term depression (LTD) at Purkinje cell synapses, nor the relevance of intrinsic neuronal dynamics (he assumed implicit rate coding), nor the impact of network geometry (he used only statistics of connections). While Marr’s theory is a brilliant example of synthesis of concepts and still guides our thinking on how the cerebellum might work, its foundations have been weakened by a series of recent discoveries showing that the olivocerebellar circuit expresses more than 15 recognized forms of plasticity, shows remarkable non-linear dynamics in its neurons, and demonstrates connectivity patterns that were not recognized previously [21]. This is a case in which the fate of a top-down models, as venerable as it might be, is undermined by disrupting discoveries that weaken its foundations.

3’2. Cerebellar realistic models: from experiments to simulations and back. – In order to take into account the relevant molecular and cellular details and the geometrical structure of network connectivity in an easily updatable framework, a realistic multiscale modelling approach is needed. At present, all cerebellar neurons have been carefully reconstructed in the simulation platform PYNN (http://neuralensemble.org/PyNN/) [76]
and advanced dataset are available for single neurons and for microcircuit construction and validation. In the cerebellum, the alternate progress of experiments and models has been pioneered and developed since 15 years already and provides an almost complete case of application of the principles described above [21,77]. We will proceed here through a series of exemplar cases.

The models of cerebellar granule cells: Granule cells are small neurons in which most ionic conductances have been resolved experimentally and a precise hypothesis on the mechanism of action potential generation was proposed [78]. However, when the 7 known ionic currents of the neuron were placed in a realistic model, this was unable to predict the whole set of granule cell functional states. In particular, the newly discovered oscillatory and resonant behaviours in a low-frequency band (the theta band on EEG) could not be resolved. This prompted the search for the hypothesized missing current, that was actually discovered and characterized. Once this last current was introduced in the model, this was able to reproduce reliably all the granule cell electrophysiological behaviours [79].

A second case concerns the generation of action potentials in granule cells. While the general mechanisms were accounted for by a single-compartment model, the size of the action potential was larger than that measured experimentally suggesting that some elements were missing. Sometime later, immunolocalization experiments revealed that Na channels are located primarily in the axonal initial segment [82] and this result was soon confirmed by single channel and whole-cell recordings [81]. The construction of a new multicompartmental model, which was keeping these details into consideration, explained what was happening [80]. The spikes were actually generated in the axon initial segment and then back-propagated passively into the soma loosing amplitude and explaining the anomaly of the first model.

A third case concerns the transmission of action potentials in the granule cell axon, the parallel fibre. While experimental measurements show that the action potential can travel long distances without shape or velocity alteration along the parallel fibres, the model showed a progressive reduction of the action potential size and anomalies in the subsequent spike after hyperpolarization. The reason for this effect came to light in consequence of experiments using advanced imaging techniques with voltage sensitive dyes. These recordings suggested that the membrane resistance of the axon tended to infinity, at odds with the Hodgkin-Huxley model [84] that assumes a finite membrane leakage. Once this high resistance was placed in the axon of the granule cell model, only single action potentials were generated by the axon, but not the repetitive firing, resembling the alteration caused by mutations of a growth factor called FHF (fibroblast growth factor homologous factor). FHF is a modulatory factor that shifts the inactivation curve of the Na channel. When this curve was shifted, the model became able to generate repetitive firing in the axon at appropriate frequency and transmission speed. The absence of FHF in the axon was subsequently demonstrated [83] (fig. 10).

The models of cerebellar Golgi cells: Golgi cells are the main interneurons of the cerebellar granular layer and their active membrane properties were revealed experimentally [26]. Golgi cells are neurons showing short response bursts, pace-making and phase-rest in vivo. While certain of these properties, it was still unclear whether these properties originated by intrinsic factors or by network dynamics (i.e. through the intervention of other neurons). The realistic model predicted that the Golgi neurons can themselves generate all these properties based on the ionic channel complement [85,86]. Further developments of these models allowed to account for electrical communication between Golgi cells through gap junctions [87-89].
Fig. 10. – The case of cerebellar granule cell modelling. Accurate biological determinations have started repeated cycles of modelling which, in turn, revealed weaknesses in previous hypotheses and promoted the experimental investigation further. The case started with accurate determination of the main ionic currents generating granule cell electroresponsiveness leading to the first model and the identification of an additional yet undiscovered current [78]. The inability of this first model to account for the small GrC spike amplitude led to a second model accurately describing action potential generation [80]. This model solved the issue of action potential generation and accounted for specific properties of the Na current [81,82]. The precise description of action potential conduction finally led to the last model including a new model of the axon [83].

The models of the cerebellar granular layer: The whole granular layer circuit was reconstructed using the granule cell and Golgi cell models reported above using dynamical synapses [90, 91] (fig. 11). In this way, the properties of single neuron models were propagated into the circuit [92]. Interestingly, the model predicted a set of emerging properties about the spatio-temporal organization of signal processing. These include the conversion of noisy inputs into coherent low-frequency oscillations, the organization of responses to single active fibre bundles into centre-surround, and the generation of logical operations inside the circuit. At the mechanistic level, the properties of neurons match the circuit time constant and frequency dependencies, so that intrinsic oscillations and resonance are instrumental to make the whole circuit oscillating at the same frequency.

The models of granule cells were used to predict the generation of LFPS in vivo. The models were used to generate extracellular current that were then let circulate in the extracellular space reconstructing the electric field [93,94]. This allowed to explain the origin of LFPS in vivo, showing that are generated by dense neuronal clusters in apparent contrast with the Marr theory that predicted sparse granule cell activation. Again, realistic models challenge foundational theories constructed top-down.

Other cerebellar neuron models: Further combined experimental and modelling investigation are now being carried out for the Purkinje cells and stellate cells [95]. These examples show that realistic modelling is able to optimally interact with experimental determinations providing the basis for a mechanistic understanding of microcircuit computation.
Fig. 11. – The model of the cerebellar granular layer microcircuit. The assembly of GrC and GoC models through precise connectivity rules and the development of synaptic transmission models has allowed full modelling of the granular layer microcircuit. This specific example is taken from a 10000 neuron network [92], that is able to reproduce all the known spatiotemporal dynamics of the microcircuit. These include oscillations, resonance, bursting and centre-surround response patterns.

A quest for modelling across scales starting from the molecular-level: Cerebellum modelling fully embraces the concepts of multiscale modelling [21, 77, 96]. It would be therefore critical to start from the molecular level in order to incorporate fundamental knowledge about the underlying low-level phenomena that generate in a natural way the complex set of mechanisms controlling plasticity, homeostasis and neuromodulation. For example, while a wealth of information is available on multiple forms of plasticity in the cerebellar glomerulus, they have for the moment been described only in theoretical models [97]. Preliminary models of the biochemical cascades leading to cerebellar LTP and LTD are available and a unified mechanisms explain LTP and LTD dependence on frequency, duration, phase (STDP), and membrane voltage has been proposed [98](14).

(14) A further elaboration of the integrated control of glomerular plasticity has been presented in the Bachelor Degree thesis of Leonardo Daniel Herbas Burgos, University of Pavia, 10 September 2016, “Simulazione di un modello realistico per la plasticità sinaptica tra le fibre muscoidi e la cellula dei granuli” (Simulation of a realistic model for the synaptic plasticity between the mossy fiber and the granule cell).
Fig. 12. – Robotic models and simulations of cerebellar microcircuit functions. This figure shows a robotic simulation of an associative learning task using a cerebellar spiking neural network (SNN). The cerebellum circuit was simplified and embedded into a robotic control system, in which it provided the substrate to integrate spatio-temporal information in different associative learning tasks. (eye blink classical conditioning (EBCC)-like, vestibulo-ocular reflex (VOR) and upper limb reaching perturbed by force fields). Note that this model uses multiple bidirectional plasticities and allows to analyze the firing pattern of single neurons and its evolution during learning (modified from [77, 104, 106]).

These preliminary results suggest that complex multiparametric plasticity rules could be reconstructed based on the molecular interactions. For example, in the cerebellar glomerulus, the biochemical mechanisms activated by glutamate and GABA receptors influence one another and control several processes through generation of second messengers and calcium waves in the cytoplasm, regulating membrane channels and receptors as well as presynaptic neurotransmitter release [97,99]. A full understanding of the system would therefore require a precise reconstruction of cerebellar glomeruli in terms of molecular mechanisms. Interestingly, the same molecular mechanisms can even explain vascular motility [100], generating a close bridge between microcircuit functions and the origin of MRI signals.

3.3. Cerebellar spiking neuro-robots: closed-loop simulations of behaviour. – In order to analyze the circuit at work in closed-loop, a hybrid system with a spiking neural network (SNN) of the cerebellum embedded into a classical controller was constructed [101-105] (fig. 12). The microcircuit had to be simplified and the choice was, to begin with,
to use leaky integrate-and-fire (LIF) neurons. These are very simple neuron models lacking internal non-linear dynamics and are rate-modulated, but have the advantage to be easily implemented and to be computationally inexpensive. These LIF models were tuned toward the fundamental properties of specific neurons in terms of membrane time constant and synaptic inputs, so that the general reactivity of the system was maintained. Thus, while a direct insight on factors, like connectivity and plasticity, could be gained, the question remains on how more realistic neuronal dynamics would modify robot behaviour. Eventually, these models have provided remarkable insight on how the system operates in closed loop. The controller and SNN architecture was adapted to achieve very fast computations, up to real-time, so that massive simulations lasting for the time required (minutes to hours) could be run and the impact of long-term synaptic plasticity on the network could be investigated along with the motor control strategies adopted by the system. Importantly, we have been able to introduce different multiple forms of plasticity tuned toward the time constants of the real circuit and to investigate their impact.

The cerebellum acting in closed loop demonstrated the ability to learn and control motor tasks for which it was not programmed explicitly, revealing therefore that it implemented a general algorithm that could be used under many different circumstances. These included eye-blink classical conditioning, vestibulo-ocular reflex, force-field adaptation, obstacle avoidance tasks, and continuous motor control toward a complex target trajectory. The cerebellar network was therefore spontaneously operating as a generalized adaptable controller.

The impact of multiple plasticities: The first advantage of having a SSN of the cerebellum embedded into a robotic controller was to be able to investigate the role of multiple forms of plasticity [77]. The cerebellum plays a critical role in adaptive motor control and its complex plasticity mechanisms implement fundamental operations of prediction, timing and learning. The spiking cerebellar robot proved able to reproduce a cerebellar-driven associative paradigm, the EyeBlink Classical Conditioning (EBCC). Bidirectional plasticity at 3 different sites (the parallel fibre - Purkinje cell synapse, the Purkinje cell - DCN synapse, the mossy fibre - DCN synapse) was required to reproduce the whole set of properties of human EBCC, comprising timing and response rate, fast acquisition, stabilization, extinction, and re-acquisition. Importantly, learning proceeded through two steps determining a faster and slower learning phase, as indeed revealed experimentally. Thus, through this closed-loop modelling, the unanswered role of the multiple plasticity forms of cerebellum starts to come to light.

This approach is relevant for its ability to match the top-down intuition of Marr’s motor-learning theory [106]. The Marr’s theory envisaged that a circuit algorithm could be resolved on the basis of microcircuit computation and implementation. Actually, the SNN of the robot generated implicit spiking computations able to produce associative sensory-motor behaviours. That is, we have reversed the original procedure: rather than anticipating an algorithm and looking for possible computations and implementations capable of generating it (inverse problem), we have followed a bottom-up approach yielding a behavioural response (an adaptive sensori-motor association) built on network constructive principles and plasticity rules. In addition, rather than investigating the cerebellar circuit in isolation, we have engaged it into the feed-back and feed-forward loops of an entire sensori-motor system operating in closed-loop. Therefore, a main conceptual pillar derived from Marr’s theory has been satisfied, although with extensions to the original concepts.
The cerebellum as a generalized controller: The second remarkable fall out has been to address one of the core questions about the cerebellar network: how can the cerebellum perform its operation of forward/feed-back controller [104]? The cerebellum receives command from the motor cortex and sensory inputs form the spinal cord, with whom it is integrated both in feed-forward loops (delivering corrective terms to the spinal cord) and in feed-back loops (delivering correcting terms to motor cortex). In the feed-forward scheme, the cerebellum receives sensory inputs and produces motor corrective terms, implementing therefore an “inverse model” of the kinematics and dynamics of movement. In the feed-back scheme, the cerebellum receives motor inputs from the cerebral cortex and produces sensory corrective terms implementing therefore a “forward model” [107,108].

Eventually, closed-loop robotic simulations allowed to identify the role of the multiple controller loops and plasticity forms to determine unique properties of biological learning and motor control, including generalization, acceleration and dynamic memory transfer.

Closed-loop robotic simulations: how far from human behaviour? In summary, neuro-robotic simulations provide a unique tool to understand how the elementary properties of neurons and the architecture of circuit organization impact on behaviour. Of special interest from our cerebellar work is that models reconstructed from mice data enabled the robot to reproduce behaviours of humans. Nonetheless, for impressive it might be, this result does not mean that the goal of simulating humans behaviour has been achieved. This simply means that there are elementary components of behaviour that are maintained across species, and the EBCC is one of these. Other more complex behaviours are species-specific. This issue is relevant once considering that there are special properties of neurons and microcircuits [109] and that there are even bigger differences in the architecture of large-scale connections and modules, that differentiate species one from the other [110,111]. For example, humans have many more cortical microcolumns and than mice have, and their connectivity is much more complex too. Therefore, scaling up from mice to humans is not just a question of size but requires specific knowledge of cellular and microcircuit properties, as well as of the connectome, that need to be incorporated in robotic simulations. This scale-up modelling exercise between species needs to rely on techniques such as MRI that take integrated signals and interpret them currently with top-down models, indirectly reconstructing functional and structural connectivity properties for example between cortical areas. Here is where bottom-up realistic models could help validating results in humans in vivo, translating cellular and microcircuits properties to large-scale systems.

Another important aspect pertains the predictive power of these robotic simulations. Alterations in the cerebellar microcircuit model have been shown to precisely predict EBCC alterations in human pathology (the correspondence to humans reflects again the considerations given above). And, as far as the elementary aspects of circuit functioning are concerned, these results may be very useful for their potential biomedical applications. The challenge is now to substitute the current simplified models of neurons and microcircuits with more realistic ones so that, from their activity during a specific behavioural task, scientists should be able to infer the underlying coding strategies at the microscopic level. Moreover, by generating a complex cerebellar connectivity would allow to move toward macroscale brain modelling, enabling the robotic system to face “human-like” behaviours more complex than EBCC.
3.4. The cerebellar connectome: toward macroscale brain modelling. – As explained in sect. 2, MRI and connectomics are fundamental to linking low-level to high-levels brain phenomena and provide therefore critical data for understanding brain structure and function and for brain modelling. At the very least, MRI tractography can be used to reconstruct the connectome in large-scale brain models, and fMRI can provide fundamental validation cues for model simulations. Surprisingly enough, MRI studies have rarely addressed the cerebellum and only recently an MRI perspective on cerebellar connectivity and functioning started to appear (fig. 13; see also figs. 6–9).

An important discovery has been that the cerebellum is not only connected to motor areas but also to associative areas, in particular to the prefrontal cortex and the temporal cortex, two regions controlling higher cognitive and memory functions [28]. This connectivity is reciprocal and accounts for up to 80% of all cerebellar connections with the neocortex (although MRI tractography is not strictly quantitative and can generate false positives, this percentage is impressive). This observation was an indication that the cerebellum was involved not just in motor but also in cognitive processing, in line with an extensive analysis of available literature [22].

On the functional side, fMRI investigations have revealed that the cerebellum is engaged in multiple aspects of sensori-motor and cognitive processing. During voluntary movement, for example a squeeze-ball task [112], several cerebellar areas are involved along with sensori-motor and associative areas in the cerebral cortex, including primary and secondary motor areas, visual areas, temporal areas and prefrontal areas. Connectomic maps have also been obtained from resting state fMRI signals [29]. These
connectomic reconstructions actually indicate the cerebellar and cerebrocortical areas that are functionally correlated suggesting possible architectures for cerebrocortical interactions. Interestingly, the combination of tractographic and functional data has allowed an advanced reconstruction of the cerebellar connectome, in which both edges and nodes are defined and weighted. In the future, the application of Dynamic Causal Modelling and Psycho-physiological Interactions techniques [25] may be used to investigating how distributed signal processing occurs in the network.

It is therefore envisaged that, like in the basal ganglia [113], the cerebellar modules are connected to different and multiple cerebro-cortical areas forming closed loops controlling different aspects of behaviour [37]. This hypothesis is currently under investigation using a combination of psychophysiological tests, fMRI and TMS techniques.

3.5. High-level model validation from integrative physiology. – TMS (transcranial magnetic stimulation) is a technique that allows non-invasive localized brain stimulation elaborating causal relationships between brain regions and functionalities. The application of TMS pulses to the cerebellum proved able to impair both sensori-motor and cognitive processing. In both cases, the cerebellum is engaged in processing precise and fast response timing, which is disrupted once the system is perturbed. Moreover, in both cases the cerebellum is involved in predictive actions (see fig. 12).

In sensori-motor paradigms, repetitive TMS on the vermian and paravermian cerebellum impairs memory consolidation in eye-blink classical conditioning (EBCC) [114]. Interestingly, the availability of precise learning curves has allowed to implement robotic simulations that have fitted the experimental data [105,106]. The robotic simulations, by allowing to explore internal model parameters, have provided a hypothesis for learning in the cerebellar circuit, in which patterns are first rapidly acquired in the cerebellar cortical circuit and are subsequently transferred to the deep-cerebellar nuclei for consolidation of memory. These results support previous observation carried out using different learning paradigms [115].

In cognitive processing, single-pulse TMS on the lateral hemispheres impairs motion detection and visual pattern discrimination [116,117]. In this case, rather than interfering with cerebellar learning, the protocol interferes with cerebellar processing on a fast time scale. Presentation of visual patterns for less than 50 ms activates the cerebellum but not the cerebral cortex and this is enough to generate a prediction on pattern motion. Interfering with cerebellar processing using TMS significantly altered task performance.

An interesting development has been to apply TMS to a psychopathological condition, the Bordeline Personality Disorder (BPD) [118]. BPD is a complex behavioural disorder characterized by a loss of impulsivity control presumably involving alterations of the cerebello-prefrontal axis. Indeed, BPD subjects confronted with an affective go-no-go task performed worse than healthy subjects, but inhibitory cerebellar TMS was able to improve BPD performance toward control values.

4. – Conclusions and outstanding questions

It can be envisaged that multiscale models will help answering fundamental biomedical questions. Can modelling help to reconstruct the fMRI signal starting from neuronal activities? Can modelling help to predict pathological states starting from neuronal activities? Is it possible to develop accurate neuromorphic hardware accounting for biological neuronal activities? Will ultimately be possible to generate a theory of the brain through a reverse engineering process?
Cerebellar modelling can help addressing the issues, since it is available at different levels, from realistic modelling of neurons and microcircuits to robotics and *in vivo* in humans connectomics, and is providing one of the most compelling examples of the integrated application of experiments and modelling to neuroscience. This modelling is intrinsically multiscale and contains both bottom-up and top-down elements. This modelling is progressing rapidly thanks to i) the availability of detailed data on cellular and microcircuit neurophysiology and on *in vivo* connectomics, ii) the predictions of foundational theories, and iii) the possibility of implementing sensori-motor loops in robotic simulations. Thus, cerebellar modelling clearly illustrates how low-level representations of neuronal activity, intermediate level large-field networks and high-level connectomics can be used to explain how ensemble brain functions might emerge from elementary neuronal components.

* * *

This work was born from the fruitful discussion developed in the years in the research groups of the University of Pavia and UCL and has extended to several other groups through various Institutional (Centro Fermi) and European projects (especially Spike-Force, SENSOPAC, REALNET, CEREBNET) and, more recently through the visionary work of the Human Brain Project. A fundamental step on the way of this paper was also the recent course held in Erice on “Modelling the brain: from Neurons to Networks and integrated circuits” by the Ettore Majorana Foundation and Centre for Science and Culture.

**Appendix A.**

The biophysical principles of realistic neuronal modelling

“Realistic” neuronal modelling indicates models that are biophysically detailed and that can generate membrane electroresponsiveness based on know principle of neuronal membrane cellular biophysics. The membrane model is based on the “parallel electrical equivalent circuit” [19, 20]. The inside and the outside of the plasma membrane are connected through parallel electrical resistances representing the ionic conductances. Moreover, a capacitive branch represents the hydrophobic non-conductive lipidic bilayer. Across the membrane a potential difference, \( V_m \), is established. The conductances \( g_{Na}, g_K, g_{Cl}, g_{Ca} \) correspond to the main permeant ions, \( Na^+ \), \( K^+ \), \( Cl^- \) and \( Ca^{2+} \), and \( E_{Na}, E_K, E_{Cl} \) and \( E_{Ca} \) are the equilibrium potentials for these ions. In addition, there is an aspecific leakage conductance \( g_{leak} \) with the associated \( E_{leak} \). The resistive branches are effective current generators with tunable resistance. Thus, when a current \( I_m \) flows through the membrane, it divides between the capacitance \( C_m \) and the conductances \( g_K, g_{Na}, g_{Cl}, g_{Ca} \) and \( g_{leak} \). According to this equivalent electric circuit, the membrane equation is

\[
I_m = I_c + (I_{Na} + I_{Ca} + I_{Cl} + I_{leak}) + I_{inj},
\]

\[
I_m = C_m \frac{dV}{dt} + \sum_i \left[ g_i (V_m - E_i) \right] + I_{inj},
\]

where \((V_m - E_i)\) is actually the driving force for each \(i\)-th ionic current: \((V_m - E_K)\), \((V_m - E_{Na})\), \((V_m - E_{Cl})\), \((V_m - E_{Ca})\) \( (V_m - E_{leak}) \). The membrane equation, due to the capacitive term, is a first-order differential equation with exponential solution.
Importantly, the conductances $g_K$, $g_{Na}$, $g_{Cl}$ and $g_{Ca}$ are themselves a function of $V_m$ and $t$ (while $g_{leak}$ is voltage and time independent). The standard description of voltage- and time-dependent conductances has been pioneered by Hodgkin and Huxley [84,119,120], who showed that each ionic conductance depends on the probability that some gating particles are in the permissive ($y$) or non-permissive state ($1 - y$). Moreover, there are both activation ($y_{act}$) and inactivation ($y_{inact}$) particles that can be replicated in numerous copies ($m, n$) inside each channel. There can be multiple such particles in each ion channel and each one can oscillate between $y$ and $1 - y$. Eventually, each ionic conductance depends on the probability that the activation or inactivation particles are in the permissive state scaled by a maximum value $g_{max}^i$:

$$g_i = g_{max}^i \cdot y_{act}^m \cdot y_{inact}^n.$$  

The $y$ to $1 - y$ conversion occurs at a rate determined by gating constants, $\alpha$ and $\beta$, following first-order reaction kinetics and bringing the reaction from the initial value $y_0$ to the final value $y_{\infty}$. The voltage dependence of the gating particles can be approximated by Boltzmann and Arrhenius equations. By considering each $i$-th activation or inactivation particle, the mathematical description of the membrane can be represented by an ordinary differential equation (ODE) system:

$$\begin{aligned}
\frac{dV}{dt} &= \frac{1}{t_m} \left( V_m - \sum_i \left[ g_i \left( V_m - E_i \right) \right] + I_{inj} \right), \\
\text{where} \quad t_m &= C_m g_{tot} \\
g_i &= g_{max}^i \cdot y_{act}^m \cdot y_{inact}^n, \\
\frac{dy_i}{dt} &= \alpha_i - \left( \alpha_i + \beta_i \right) \cdot y_i.
\end{aligned}$$

In a neuron, there are several gating particles describing the many ion channel types that populate the membrane. This yields a large ODE system, which is usually solved through numerical methods [121,122]. Once appropriately parameterized, the solution of this ODE system gives the membrane potential time course [79,92]. A variant of this approach can be applied to describe the synaptic vesicle cycle causing neurotransmitter release [123].

**Appendix B.**

**Principles of connectomics**

In order to sample the human connectome it is important first of all to distinguish how to represent it and what properties are going to be exploited. Graph theory was identified as able to represent a set of nodes (e.g. discrete grey matter regions, anatomically or functionally defined) and define a connectivity metric between each pair of nodes (fig. 14). Such metric can be calculated in several ways, depending on the property of the brain to be studied, either as structural, functional or effective connectivity [124]. Moreover, a modular structure can be identified as subtending the brain connectome, with key nodes identified as hubs of a typical small world network [62].
Nodes can be defined based on common atlases, e.g. automated anatomical labeling (AAL) system [125], including the neo-cortex only or deep grey matter regions and the cerebellar cortex. Such parcellation of grey matter regions to be used as the connectome nodes must then be registered to each individual subject’s space and used for the edge definition.

The structural connectome can be created from diffusion weighted imaging data, applying methods like tractography to define connectivity between grey matter nodes [126]. This method relies on the ability of tractography to depict connections and estimate number of streamlines between nodes in a consistent manner, therefore it is imperative to take advantage of the latest software packages implementing the most robust and model-free (if possible) methods for tractography. The edge metric for this connectome becomes therefore the number of streamlines between pairs of nodes. Other tract properties can be used as node metrics, including emerging ones such as the $g$-ratio or the axonal density along the tract.

It is important to note that the structural connectome can be built on the individual subject basis as nodes and edges can be determined specifically for each brain.

Another metric used to calculate a structural connectome is the average thickness of the grey matter parcellations. From an evolution point of view, cortical regions belonging to the same structural network are growing with similar properties, including average thickness. At group level it is indeed possible to determine patterns of correlations between grey matter regions, whose value can be used as edge in the graph metric of the connectome.

Similarly, it is possible to build the functional connectome of the brain by analyzing the functional connectivity of brain regions, i.e. establishing correlations between the synchronous fluctuations of the MRI signal along the time series of the acquired data.
Once a graph has been built with nodes and connectivity strengths, i.e. edges, whether these are properties of tracts or correlation coefficients reflecting similarities of properties of an underlying biomarker (e.g. cortical thickness, functional connectivity), then it is possible to determine macroscopic characteristics of the graph that collapse network properties in a handful of measures such as global efficiency, nodal degree, edge density, segregation (for the mathematical representation of the graph and its properties see [124]).

REFERENCES


Stix G., in Scientific American Blog - February 25, 2013, Big Neuroscience: Billions and Billions (Maybe) to Unravel Mysteries of the Brain.


