

Information, Noise, Coding, Modulation: What about the Brain?

Claude Berrou, Olivier Dufor, Vincent Gripon, and Xiaoran Jiang

Télécom Bretagne

CNRS Lab-STICC

Brest, France

firstname.surname@telecom-bretagne.eu

Abstract—At the microscopic level, the brain is fundamentally a matter of physics and chemistry, as all the components of the universe are. At the macroscopic scale, behavior, psychology and affects are the main dimensions of its life. To convert atoms and molecules into intelligence, some kind of information has to be fixed in the grey matter of the cerebral cortex. The way this "mental information" is materialized and processed is still an enigma, probably the most puzzling problem addressed to science nowadays. At this mesoscopic level of the brain functioning, the concepts to consider are likely the same as those considered in communication and information theory, mainly information, noise, coding and modulation. This paper proposes some ideas that could help understand some features of the brain in an information-processing perspective.

Keywords—Neuroscience, information, communication, distributed coding, modulation, noise.

I. INTRODUCTION

When trying to build bridges from information science towards biology, in particular neurobiology, many researchers may be quickly discouraged by the tremendous complexity of the physicochemical mechanisms brought into play in living organisms. The few essential principles relevant to information representation and processing are most often hidden under the biological canopy. This is particularly true when considering the brain, both a tremendous biological device with multiple concomitant processes and an outstanding "learning and reasoning machine" with a unique thought at any one time. Taking the problem from the psychological side is not easier, as so many and often conflicting theories on human intelligence and aptitudes have been defended from time immemorial. The intermediate so-called mesoscopic level is perhaps the least difficult to deal with.

A simple way to start with confidence the reverse engineering of the brain is to share the standpoint of Jeff Hawkins: "Most scientists say that because the brain is so complicated, it will take a very long time for us to understand it. I disagree. Complexity is a symptom of confusion, not a cause. Instead, I argue we have a few intuitive but incorrect assumptions that mislead us. The biggest mistake is the belief that intelligence is defined by intelligent behavior" [1]. In this spirit, why not consider that some methods or techniques that human intelligence has imagined to improve processing technologies could be in return profitable to computational

neuroscience? Some techniques that have recently emerged in the telecommunication field may belong to this catalogue of biologically acceptable solutions, especially distributed coding, cooperative communications and spatial modulation.

A striking example is provided by the analogy that can be made between a variable-processor of a Low Density Parity Check (LDPC) decoder and a neuron when this is reduced to the McCulloch-Pitts model [2]. Both are able to aggregate signals, positive/excitatory or negative/inhibitory, and to produce an output depending on the summation result. Of course, the analogy does not hold for the parity check-processors since no algebraic computation seems natural in the brain. However, viewing neurons as variable-processor nodes in the cortical network may be an inspiring starting point in the specification of neuro-inspired message passing machines. In the next sections, some ideas that could help understand the brain at the informational level are propounded. The underlying rationale is that of Ockham's razor principle: always proceed with no more hypotheses and/or variables than needed for the specific problem you have to solve.

II. THE CEREBRAL MATERIAL

Not taking into account the metabolic components (glial cells, blood vessels, etc.), the brain is "simply" made of neurons which communicate with each other besides receiving and sending various physiological (sensory-motor) signals. The inputs of the neuron (synapses) are numerous (10,000 is a typical order of magnitude) and there is only one output (axon). There are about some tens of billions neurons in the brain, each one being controlled by various non trivial biological and genetic processes. One of them, called *synaptic plasticity*, is fundamental in computational/informational neuroscience and says that a synapse connecting neurons A and B may be reinforced if the activity of A triggers systematically that of B. This is known as the Hebb's rule [3]. Anti-Hebbian principles have also been proposed to explain synapse weakening or disappearance [4].

Signals that neurons convey are of the impulse type. The frequency and instants of these pulses, which are called *action potentials* or *spikes*, is very variable. On this point, the informational modeling of the nervous system divides up into two quite distinct approaches: the *spiking neuron* and the *binary neuron*.

This work was supported by the European Research Council under Grant ERC-AdG2011 290901 NEUCOD.

A. Spiking neuron

Using an appropriate model, for instance that of Hodgkin-Huxley [5] or simpler ones [6], spikes can be realistically generated. It is possible to associate binary value 1 with the presence of a spike, at a given time, and value 0 otherwise. Information being encoded both spatially and temporally and delays being taken into account in the circuitry, neural dynamic systems may be assessed with a high level of biological plausibility, but also with limited systemic complexity (up to some thousands neurons seems feasible with a standard computer). This model is particularly useful for the modeling of sensory-motor mechanisms [7] or for studying the influence of the various physicochemical cell parameters.

B. Binary and multi-valued neuron

When reducing the functionality of the neuron to that of the McCulloch-Pitts model, systems with millions of units can be simulated with acceptable response time. It is then possible to devise and test vast networks in which advanced concepts such as sparse coding [8] and clustering [9] may be implemented. Neuron with binary output and binary inputs (*i.e.* with synaptic weights equal to 0 or 1) is the simplest model that can allow studying the brain at a really informational and systemic (*e.g.* hierarchic) large scale. If needed, multi-valued neuronal output may be considered to take variable spiking frequency into account, that is, variable energy. But this multi-valued signal is rather to be regarded as a measure of reliability (like the logarithm of likelihood ratio in message passing decoders) than as information encoding and mapping.

As for the connections between neurons, whatever the mathematical model, they are mainly of three types: short or long range excitatory or short range inhibitory [10]. Long range signals, transmitted through white matter and the *corpus callosum*, enable communication and cooperation between the various functional areas of the brain within and between both hemispheres. Short range signals serve for local dedicated processing. This, added to the fact that inhibitory neurons are a minority in the cerebral cortex (about 20% [11]), may suggest that information storing, recovery and conveying is essentially provided by excitatory neurons. Section V is more explicit about this point.

Despite the analogy we drew in the introduction between LDPC variable-processors and neurons, it is difficult to admit that the neuron could be the repeating fundamental operator in the brain, due to its great diversity in shape, location within the six layers of the grey matter, axonal projection, neurotransmitter sensitivity and behavior. If one wants to find such an omnipotent unit to emulate the "universal" mental operator, the concept of microcolumn (also called minicolumn) may be very useful. This very heterogeneous group of about 100 neurons repeats itself quasi-uniformly in a large part of the grey matter and is able to send and receive excitatory signals to and from all regions of the brain, as well as lateral short range inhibitory signals [12]-[14]. These lateral inhibitory signals may be, for instance, necessary to the implementation of the Winner-Take-All mechanism which helps a strongly active unit extinguish other competing ones in its close vicinity, for reasons of limited available energy (blood and oxygen).

III. THE SHANNON'S MODEL OF COGNITION

How can a system with a limited number of operators construct and transmit or store an accurate and robust representation of a richly detailed and fleeting source of information? Such a question is familiar in the field of sensing, telecommunication and mass storage and has found many satisfying responses, in video and audio applications for instance, based on the Shannon's model: firstly, information is cleared of useless components and then "smart redundancy" is added to allow error correction at the receiver or reading head side. This well-known source coding - channel coding scheme seems relevant to schematize the way the brain captures and stores what it considers as essential information (Fig. 1).

The first part of the acquisition work (source coding) has attracted a lot of interest from the computational neuroscience community under the name of "machine learning". From the seminal perceptron [15] to deep neural networks [16], from decision trees [17] to dictionary learning for sparse coding [18], many architectures and methods have been proposed so far to mimic the brain as an intelligent learner and classifier. In contrast, much less effort has been made on the second part (channel coding), which is related to the problem of robust and durable memorization and even further to reasoning. The use of two distinct expressions: *computational neuroscience* or *informational neuroscience* could be helpful to mark the difference between the two classes of problems.

Like many artificial systems (*e.g.* a video camera), the brain, as a capturing and storing machine, has to transform signals coming from an analog environment (the surrounding world and the physiological body) into stable, robust and durable information in order to recover it easily in case of necessity. The best way that engineers have found to achieve this in artificial systems is *digitizing*. Indeed, smart redundancy is much more resource-efficient to implement in digital systems than in analog ones. In the latter, exploiting the concept of redundancy amounts to multiplying (duplicating, triplicating, etc.) the material and/or the functions, that is, so-called functional redundancy which is not a thrifty solution.

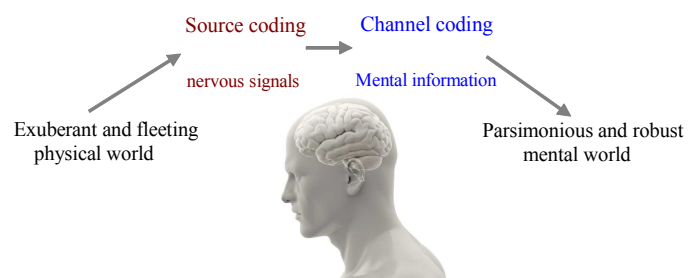


Fig. 1. The Shannon's model of cognition: richly detailed information provided by the external world is strongly compressed and then "smart redundancy" is added to enable robust and durable memorization.

If evolution, in its continuous race towards efficiency, has made good choices, it is not inconceivable that long term cerebral memory could be built on digital principles. Indeed, more than one billion years ago, evolution already made the

digital choice for genetic memory, that is, DNA with its quaternary alphabet (A, T, C, G). It would not be so surprising that the *mental information* that the brain has to preserve firmly during the time of a life is also of the digital type. In spite of all attacks (chemical, radiative, degenerative) that our brains sustain continuously, they are able to maintain and associate a myriad of pieces of information for tens of years and without any error. "9 times 8 equals 72" is learnt during the childhood and may still be valid seventy years later!

In information sciences, in particular telecommunications, the distinction between the two notions of signal and information is well understood. It is much less clear when it comes to the nervous system. Because all nervous cells use the same kind of signalization - action potentials with same shape and magnitude -, one could think that all these cells have the same functional role. But, when considering functionality and not cellular biology, what have in common for instance a neuron of the premotor cortex which contributes to a movement of the hand and another, in the prefrontal cortex, which helps remind the result of an elementary multiplication? Obviously, these neurons use the same signalization but do not handle the same kind of information. One has still to understand how these particular cells that were initially adopted by evolution to map out the sensorial system have also become suited to the implementation of the long term memory (LTM).

IV. NOISE AND VARIABILITY

The brain is a noisy and changeable material. It could be compared, for instance, with an integrated circuit having billions of transistors, each one producing impulses fired at random according to a Poisson's distribution. Moreover, these transistors would be clustered in processing units whose size and organization could be modified at any time. This circuit would probably not work properly. The brain does!

A. Noise

Two types of noise have to be considered. Firstly most neurons have a tonic activity due to fluctuations of their membrane potential. A continuous flow of ions (K^+ , Na^+ , Cl^-) through this membrane maintains its voltage at a constant resting mean value of about -65 mV. The flow of ions being a random walk, the voltage varies around this value and may reach a sufficient level to depolarize the inner side of the membrane and make the neuron fire [19]. Fluctuations in neurotransmitter vesicle release [20] may also contribute a lot in the generation of spurious spikes as well as incoming random impulses coming from other neurons. The latter point indicates that spurious spikes may be strongly correlated among neighboring cells.

Secondly, we have to consider another kind of noise whose effect is not the generation of spurious spikes but the suppression of good ones. For chemical reasons (lack of neurotransmitter at the arrival of a spike), the voltage-current conversion at the synaptic junction can fail. For instance, authors have shown that in natural conditions, the average release probability of glutamate (the main excitatory neurotransmitter of the nervous system) in cortical neurons is

0.5 to 0.64 at low stimulation frequencies and may be as low as 0.1 for frequencies higher than 5 Hz [21], [22].

These two kinds of noise, positive (insertion) and negative (deletion), suggest modeling communication paths between neurons or columns as insertion/deletion channels with high error rate. The way mental information propagates between the different areas of the brain must have adapted to these very severe conditions (see section VI).

Lastly, it is well known that noise can play a positive role in the dynamics of non linear systems, for instance to escape local minima or to synchronize oscillators. Though several authors have also considered neural noise as beneficial in mental information processing (see [23] for instance), in order to be coherent with the analogy between information/communication systems and the brain that we want to deepen, noise has to be regarded in a classical way, that is, as the opponent of information.

B. Variability

There are no two human brains identical. With tools that continue to improve, especially functional magnetic resonance imaging, neural activity may be mapped with enough accuracy to assess inter- and intra-subject variability on both structural [24] and functional [25] bases. This variability has been proved both in space and in time [26].

One very interesting point, highlighted in [25] for instance, is that variability does not express itself uniformly across brain regions. Heteromodal (associative) regions are more plastic than unimodal (specific) ones. Translated in terms of information processing, this would mean that acquisition and compression principles have been somewhat stabilized during evolution and are reproducible whereas long term memory and associative mechanisms are more tentative and adaptive in their implementation. In particular, one can see variability in frontal and parietal cortices (the main associative areas) as a consequence of redundancy rate adaptation. In order to avoid overloading and interference due to excessive density in a dedicated memorization network, one may imagine that this would extend the circuitry in its vicinity to benefit from more additional discriminating (redundant) information. In other words, the redundancy rate in the "channel coding operation of the brain" would depend on the amount of knowledge.

V. CODING

The expression "neural code" may have various meanings depending on the topic in discussion. We are interested here in the way the many elements of knowledge that our reason exploits are robustly stored in the cortex in spite of its poorly reliable components. Once again, two visions in principle conflict: that of the "grand-mother cell" and that of neuronal assemblies. The former [27] says that a fundamental and recurrent memory such that a familiar face or name is fixed by a unique neuron in a specialized cortical region. This neuron is at the top of a hierarchical processing that reduces the amount of information step by step. The latter [3], [28] upholds the idea that small groups of neurons (or microcolumns) activate synchronously to express elements of knowledge. In order not to limit memory capacity, these groups have to overlap.

When a single node in a network (or a single bit in a binary dataset) is devoted to the expression of a particular piece of information, the only one way to protect it against erasure or error is repetition. This means that the number of grand-mother cells would be several times larger than the number of elements to store. The counting law of elements of knowledge as a function of the network population would then be significantly less than linear. In contrast, when several nodes are used together to materialize elements, the notion of redundant coding may lead to more subtle protection strategies as it is for bits in a codeword. This leads to the concept of assembly or coactivation coding.

On another essential point, considering the high level of noise that affects communication between neurons as well as the natural short term fluctuations of synaptic conductance [29], the argument of analog memorization does hardly hold. Analog memorization hypothesis, which means that knowledge is embedded in synaptic weights through hebbian and anti-hebbian rules, is flimsy when it comes to robustness and durability.

All the arguments above advocate for the vision of a binary assembly-based LTM. Assembly-based models have already been proposed and evaluated with a high level of biological accuracy. In [30], [31] for instance, the authors demonstrate the possibility for neurons (or groups of neurons that we can assimilate to microcolumns) to activate synchronously when they are modeled as spiking pyramidal cells with the Hodgkin-Huxley equations. Inhibitory interneurons are used to make the discrimination between all possible patterns easier. However, because these assemblies cannot overlap, the diversity (the total number of patterns that can be stored with later good enough recovery rate) follows a linear law.

We have recently demonstrated that Willshaw networks [32], revisited in order to allow them to handle non binary alphabets, are able to store patterns according to a quadratic law and with quasi-optimal efficiency [33]-[35]. This good performance results from the redundancy naturally offered by the more than sufficient connections binding an interdependent and coactivating assembly of nodes. To speak in terms of graph theory, a fully interconnected sub-graph in a graph is called a clique (Fig.2 (a)). The number c of its vertices is called the clique order. Let us consider the case of degenerated cliques, such as that of Fig. 2 (b) and suppose the node degree α ($2 \leq \alpha \leq c - 1$) is the same for every node. Likening this pattern to a graphical codeword whose symbols are the edges, we can observe that the minimum Hamming distance is given by the comparison of two patterns having $c - 1$ common nodes, thus differing in 2α edges:

$$d_{\min} = 2\alpha \quad (1)$$

On the other hand, when the size of the graph is large in comparison with that of patterns, the coding rate is given by the ratio between the minimal number of edges necessary to specify a clique and the number of actual edges [33]. This coding rate is thus given by:

$$R = \frac{\left\lfloor \frac{c+1}{2} \right\rfloor}{\frac{\alpha c}{2}} \quad (2)$$

For c even, this reduces to $1/\alpha$, leading to a merit factor

$$F = Rd_{\min} = 2 \quad (3)$$

For c odd, the value is a little bit more favorable. This very simple demonstration shows that a clique-based code, even if edges are missing (uniformly), offers error correction properties thanks to a merit factor larger than 1 (the merit factor of the well-known (8,4,4) Hamming code is also equal to 2). This is really a good mathematical argument to support the hypothesis of cell-assembly cerebral coding. Even if connections disappear due to some damaging, the quasi-clique is still able to respond synchronously or to resonate in response to some relevant stimuli. The missing connections can also be repaired thanks to the Hebbian rule. Note also that this model involves only excitatory connections. Among other recent studies worth of mentioning about this kind of approach, we can cite [36] for improved diversity and [37] for resilience properties.

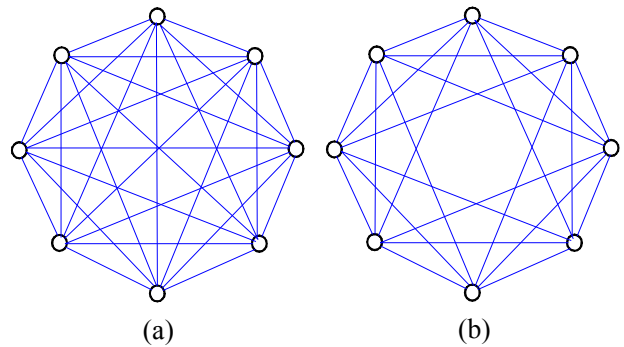


Fig. 2. (a) A perfect clique of order $c = 8$. (b) The same clique with regular degeneration (diametral connections were removed) resulting in a node degree $\alpha = 6$. Both patterns, as codewords of a binary graphical code, benefit from the same error correction properties because the merit factor F is independent of the node degree (if edges are uniformly removed) and always equal to 2 (or slightly larger for c odd).

Cerebral processing is fundamentally dynamic. So, in addition to the existence of atemporal elements of knowledge (cliques or quasi-cliques as described above), sequences have also to be considered. To implement sequences in the network of microcolumns, edges have to be replaced with arrows, which are more natural, biologically speaking (communication between neurons is unidirectional, from axons to synapses). Fig. 3 depicts a short fraction of a sequence composed of three successive patterns. Each symbol of a pattern is connected to all symbols of the subsequent one, which offers full redundancy in the immediate connectivity. Moreover, to reflect the anticipation effects brought to the fore in many experiments, hearing for instance [38], the activation of a pattern at time t may prepare those to come after $t + 1$. In Fig. 3, the activation of circles, at time t , prepares that of squares at time $t + 2$.

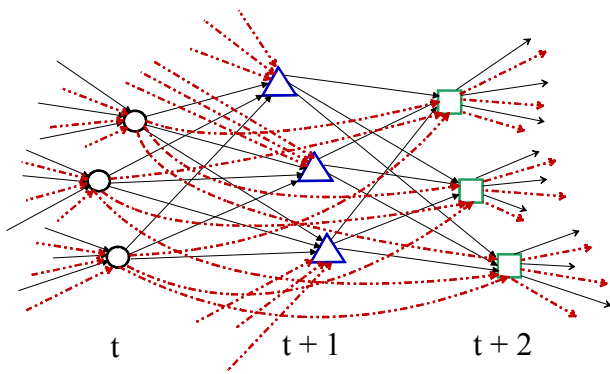


Fig. 3. A fraction of a sequence with three patterns (circles at time t , triangles at $t+1$ and squares at $t+2$). All circles are connected to all triangles and squares. The activation of circles triggers that of triangles and anticipates that of squares. Each pattern may also be interconnected as a clique (not represented here).

Both spatial assembly redundancy and sequential redundancy, with or without anticipation, may be combined to make the activation of a node more reliable. If c is the number of symbols in a pattern, fully interconnected as a clique and if this pattern has the control over the r subsequent ones, the number of stimulating signals that each node receives is $c - 1 + rc$. Of course, implementing so many nexuses has an immediate effect on the network density and may lead to overloading. This problem has been studied in [39].

VI. NETWORK AND MODULATION

The most astonishing feature of neural communication is probably the fact that a neuron is not aware of where the spikes it fires will arrive. Thus, the brain cannot be compared with a classical network, such as Internet, with addresses and routers. However, the neuron does not project its axon at random as its genetic program imposes a type of connection, short or long range and in the latter case its preferred interregional relations. So the brain appears to be a peculiar kind of network presenting both broadcasting and routing characteristics.

To be as simple as possible, we can see the cortical network as composed of two distinct families of functional modules: unimodal (specialized) and multimodal (hetero-associative). Unimodal functions deal with dedicated problems such as vision, face recognition, hearing, etc. Multimodal elements, also called hubs [40], receive various pieces of information from other parts of the brain so as to elaborate some kind of syntheses that could be useful for other cerebral processes. All this is reminiscent of what is called cooperative communication in the field of telecommunication [41]. By analogy with the decode-forward principle of multi-relay communication, a speculative vision of the role of a multimodal region would be that of an associate-decode-forward (ADF) function. Receiving simultaneous stimuli from some local networks (A, B, C, ...), for instance signals sent by a resonating assembly inside A, by another one inside B, etc., the hub (X) will try to make sense of this event. To do this, X will attempt to find in its own experience, the trace of a coactivating assembly (a clique or a quasi-clique) being the most representative as possible of the on going event. If the result is positive, X will forward it to some parts of the brain that could be concerned. Because X is not a mere re-transmitter of signals stemming from A, B, C, ...,

but performs a local task of recovery, with possible correction and/or completion, the analogy with the decode-forward principle seems quite relevant.

Around this idea of cerebral cooperative communication, many schemes linking specialized and hetero-associative functional units may be contemplated. Plausible and powerful computational architectures will certainly emerge in the years to come, benefiting from the continuous progress obtained in cerebral connectivity thanks to structural and functional imagery.

Communication between local cortical networks is achieved through the grey matter (for short range nexuses) and the white matter (for both short and long range nexuses). Signals are transmitted on multi-wire channels, called bundles or tracts and the presence of several bundles between two regions provides full duplex communication. The estimated number of wires linking two regions (about one million [42]) is much more in accordance with the population of microcolumns than with that of neurons, which reinforces the hypothesis of the microcolumn as the fundamental processing unit.

Currently, there is no technical means to observe the traffic within bundles of axons. Only indirect methods may give some estimation. For instance, based on energetic considerations, [22] suggests a 1% - 16% margin as the fraction of simultaneous active axons at a given time. This could be reminiscent of what is called spatial modulation in MIMO systems with either one transmit antenna [43] or multiple ones [44], and would suggest that mental information is conveyed by sparse combinations of active axons. One of our studies to come in the next future will deal with such a model of cerebral communication, taking into account noisy insertions and deletions as introduced in section IV.

VII. CONCLUSION

This paper has been written for the benefit of colleagues who could be interested in joining the small but expanding community of *informational neuroscience*. This field of research is to be distinguished from computational neuroscience (and machine learning as one of its important applications) and deals with the way mental information is encoded, retrieved and propagates. Informational neuroscience is to channel coding what computational neuroscience is to source coding. However, as in communication systems, compression and protection are also two very related and even overlapping matters in the brain functioning.

The principles that the brain has adopted to elaborate and fix mental information have still to be discovered and/or validated. It would not be so surprising that some ideas that have emerged from the imagination of researchers in information and communications technology, especially distributed coding and cooperative communications, could also be those chosen by life and evolution to bring them some intelligence! In this paper, we have proposed some leads to cope with this cross-fertilization of information/communication science and neuroscience, with the underlying conviction that the fundamental principles of cognition are not so numerous. Among these, the hypothesis of cerebral memory and communication based on digital principles deserves much attention. As a very speculative but inspiring example of

analogy, would it not be possible to see the inter-hemisphere communication as a kind of turbo processing, allowing both parts of the brain to work at the same time on common data (mainly sensorial) and specific ones (their own experience and architecture) and finally to agree on a shared decision, that is, a fixed point solution of their distributed process?

REFERENCES

- [1] J. Hawkins (with S. Blakeslee), *On intelligence*, pp. 5-6, Owl Books, 2005.
- [2] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bulletin of Mathematical Biophysics*, vol. 5, pp. 115-133, 1943.
- [3] D. Hebb, *The organization of behavior*, Wiley & Sons, 1949.
- [4] J. Lisman, "A mechanism for the Hebb and the anti-Hebb processes underlying learning and memory," *Proc. Natl. Acad. Sci.*, vol. 86, pp. 9574-9578, Dec. 1989.
- [5] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *J. Physiol.*, vol. 117, pp. 500-544, 1952.
- [6] E.M. Izhikevich, "Which model to use for cortical spiking neurons?," *IEEE Trans. Neural Netw.*, vol. 15, n° 5, pp. 1063-1070, Sept. 2004
- [7] J. Keat, P. Reinagel, R. C. Reid, and M. Meister, "Predicting every spike: a model for the responses of visual neurons," *Neuron*, vol. 30, issue 3, pp. 803-817, May 2001.
- [8] P. Földiák "Forming sparse representations by local anti-Hebbian learning," *Biol. Cybern.*, vol. 64, pp. 165-170, 1990.
- [9] M. Kerszberg, S. Dehaene, and J. P. Changeux, "Stabilization of complex input-output functions in neural clusters formed by synapse selection," *Neural Netw.*, vol. 5, pp. 403-413, 1992.
- [10] P. Somogyi, G. Tamás, R. Lujan, and E. H. Buhl, "Salient features of synaptic organisation in the cerebral cortex," *Brain Res. Rev.*, vol. 26, pp. 113-135, 1998.
- [11] J. S. Isaacson and M. Scanziani, "How inhibition shapes cortical activity," *Neuron*, vol. 72, pp. 231-243, Oct. 2011.
- [12] V. B. Mountcastle, "The columnar organization of the neocortex," *Brain*, vol. 120, n° 4, pp. 701-722, 1997.
- [13] E. G. Jones, "Microcolumns in the cerebral cortex," *Proc. Natl. Acad. Sci.*, vol. 97, n° 10, pp. 5019-5021, May 2000.
- [14] D. P. Buxhoeveden and M. F. Casanova, "The minicolumn hypothesis in neuroscience," *Brain*, vol. 125, n° 5, pp. 935-951, 2002
- [15] F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain," *Psychol. Rev.*, vol. 65, n° 6, pp. 386-408, 1958.
- [16] Y. Bengio, "Learning deep architectures for AI," *Found. and Trends in Mach. Learn.*, vol. 2, n° 1, pp. 1-127, Jan. 2009.
- [17] I. K. Sethi, "Entropy nets: from decision trees to neural networks," *Proc. of the IEEE*, vol. 78, n° 10, pp. 1605-1613, Oct. 1990.
- [18] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online dictionary learning for sparse coding," *Proc. ICML '09*, pp. 689-696, Montreal, June 2009.
- [19] J. A. White, J. T. Rubinstein, and A. R. Kay, "Channel noise in neurons," *Trends Neurosci.*, vol. 23, n° 3, pp. 131-137, Mar. 2000.
- [20] C. Modchang, S. Nadkarni, T. M. Bartol, W. Triampo, T. J. Sejnowski, H. Levine, and W.-J. Rappel, "A comparison of deterministic and stochastic simulations of neuronal vesicle release models," *Phys Biol*, vol. 7, n° 2, 7pp, May 2010.
- [21] 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," *J. Physiol.*, vol. 500.2, pp. 409-440, Apr. 1997.
- [22] S. B. Laughlin and T. J. Sejnowski, "Communication in neuronal networks," *Science*, vol. 301, n° 5641, pp. 1870-1874, Sept. 2003.
- [23] M. D. McDonnell and L. M. Ward, "The benefits of noise in neural systems: bridging theory and experiment," *Nature Rev. Neurosci.*, vol. 12, pp. 415-426, July 2011.
- [24] P. Fillard, V. Arsigny, X. Pennec, K. M. Hayashi, P. M. Thompson, and N. Ayache, "Measuring brain variability by extrapolating sparse tensor fields measured on sulcal lines," *Neuroimage*, vol. 34, pp. 639-650, 2007.
- [25] S. Mueller, D. Wang, M. D. Fox, B. T. Thomas Yeo, J. Sepulcre, M. R. Sabuncu, R. Shafee, J. Lu, and H. Liu, "Individual variability in functional connectivity architecture of the human brain," *Neuron*, vol. 77, issue 3, pp. 586-595, Feb. 2013.
- [26] G. K. Aguirre, E. Zarahn, and M. D'Esposito, "The variability of human, BOLD hemodynamic responses," *Neuroimage*, vol. 8, pp. 360-369, 1998.
- [27] J. S. Bowers, "On the biological plausibility of grandmother cells: implications for neural network theories in psychology and neuroscience," *Psychol. Rev.*, vol. 116, n° 1, pp. 220-251, 2009.
- [28] W. Singer, "Synchronization of cortical activity and its putative role in information processing and learning," *Annu. Rev. Physiol.*, vol. 55, pp. 349-374, 1993.
- [29] Z. Piwkowska, M. Pospisichil, R. Brette, J. Sliwa, M. Rudolph-Lilith, T. Bal, and A. Destexhe, "Characterizing synaptic conductance fluctuations in cortical neurons and their influence on spike generation," *J Neurosci Methods*, vol. 169, pp. 302-322, 2007.
- [30] E. Fransén, A. Lansner, and H. Liljenström, "A model of cortical associative memory based on Hebbian cell assemblies," *Comput. Neural Syst.*, pp. 431-435, Kluwer, 1993.
- [31] E. Fransén and A. Lansner, "A model of cortical associative memory based on a horizontal network of connected columns," *Network: Comput. Neural Syst.*, vol. 9, n° 2, pp. 235-264, 1998.
- [32] D. J. Willshaw, O. P. Bunetman, and H. C. Longuet-Higgins, "Non-holographic associative memory," *Nature*, vol. 222, pp. 960-962, 1969.
- [33] V. Gripon and C. Berrou, "Sparse neural networks with large learning diversity," *IEEE Trans. Neural Netw.*, vol. 22, n° 7, pp. 1087-1096, July 2011.
- [34] V. Gripon, V. Skachek, W. J. Gross, and M. Rabbat, "Random clique codes," *Proc. 7th Int'l Symp. on Turbo Codes and Iterative Inform. Processing*, pp. 121-125, Gothenburg, Sweden, Aug. 2012.
- [35] B. Kamary Aliabadi, C. Berrou, V. Gripon, and X. Jiang, "Storing sparse messages in networks of neural cliques," *IEEE Trans. Neural Netw.*, vol. 25, n° 5, pp. 1087-1096, May 2014.
- [36] A. H. Salavati and A. Karbasi, "Multi-level error-resilient neural networks," *Proc. IEEE ISIT*, pp. 1064-1068, Boston, 2012.
- [37] F. Leduc-Primeau, V. Gripon, M. G. Rabbat, and W. J. Gross, "Cluster-based associative memories built from unreliable storage," *Proc. ICASSP*, Florence, Italy, May 2014.
- [38] A. M. Leaver, J. Van Lare, B. Zielinski, A. R. Halpern, and J. P. Rauschecker, "Brain activation during anticipation of sound sequences," *J. Neurosci.*, vol. 29, n° 8, pp. 2477-2485, Feb. 2009.
- [39] X. Jiang, "Storing sequences in binary neural networks with high efficiency," Ph.D. thesis, Télécom Bretagne, n° 2013telb0303, France, Jan. 2014.
- [40] M. P. van den Heuvel and O. Sporns, "Network hubs in the human brain," *Trends Cogn. Sci.*, vol. 17, n° 12, pp. 683-696, Dec. 2013.
- [41] A. Nosratinia and A. Hedayat, "Cooperative communications in wireless networks," *IEEE Comm. Mag.*, pp. 74-80, Oct. 2004.
- [42] A. Schüz and V. Braitenberg, "The human cortical white matter: quantitative aspects of cortico-cortical long-range connectivity," in *Cortical areas: unity and diversity*, pp. 377-386, Taylor & Francis, London, 2002.
- [43] R. Y. Mesleh, H. Hass, S. Sinanović, C. Wook Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, n° 4, pp. 228-2241, July 2008.
- [44] J. Wang, S. Jia, and J. Song, "Generalised spatial modulation system with multiple active transmit antennas and low complexity detection scheme," *IEEE Trans. Wireless Comm.*, vol. 11, n° 4, pp. 1605-1615, Apr. 2012