Error-correcting codes and neural networks

Yuri I. Manin[1]

To Sasha Beilinson, most cordially

Abstract Encoding, transmission and decoding of information are ubiquitous in biology and human history: from DNA transcription to spoken/written languages and languages of sciences. During the last decades, the study of neural networks in brain performing their multiple tasks was providing more and more detailed pictures of (fragments of) this activity. Mathematical models of this multifaceted process led to some fascinating problems about "good codes" in mathematics, engineering, and now biology as well. The notion of "good" or "optimal" codes depends on the technological progress and criteria defining optimality of codes of various types: error-correcting ones, cryptographic ones, noise-resistant ones etc. In this note, I discuss recent suggestions that activity of some neural networks in brain, in particular those responsible for space navigation, can be well approximated by the assumption that these networks produce and use good error-correcting codes. I give mathematical arguments supporting the conjecture that search for optimal codes is built into neural activity and is observable.

> *Io mirava, e chiedea:*
> *Musa, la lima ov'è? Disse la Dea:*
> *lima è consumata; or facciam senza.*
> *Ed io, ma di rifarla*
> *Non vi cal, soggiungea, quand'ella è stanca?*
> *Rispose: hassi a rifar, ma il tempo manca.*
>
> *Giacomo Leopardi, Canti. XXXVI, Scherzo*

Mathematics Subject Classification 92C20 · 97M60

✉ Yuri I. Manin
manin@mpim-bonn.mpg.de

[1] Max–Planck–Institut für Mathematik, Bonn, Germany

Birkhäuser

Introduction and summary

Recently it became technically possible to record simultaneously spiking activity of large neural populations (cf. Refs. 1, 2 in [21]). The data supplied by these studies show *"signatures of criticality"*. This means that the respective neural populations are functioning near the point of a phase transition if one chooses appropriate statistic models of their behaviour.

In some papers the relevant data for retina were interpreted as arguments for *optimality* of encoding visual information (cf. [22]), whereas in other works such as [21] it was argued that criticality might be a general property of many models of collective neural behaviour unrelated to the encoding optimality.

In this note I test the philosophy relating criticality with optimality using the results of recent works suggesting models of encoding of stimulus space that utilise the basic notions of the theory of error-correcting codes. The recent collaboration of neurobiologists and mathematicians, in particular, led to the consideration of binary codes used by brain for encoding and storing a stimuli domain such as a rodent's territory through the combinatorics of its covering by local neighbourhoods: see [1,2, 24].

These binary codes as they are described in [1,2,24] (cf. also a brief survey for mathematicians [14]) are *not* good error-correcting codes themselves. However, these codes *as combinatorial objects* are produced by other neural networks during the stage where a rodent say, studies a new environment. During this formation period, other neural networks play an essential role, and I suggest that the *auxiliary codes* used in this process are error-correcting ones, changing in the process of seeking *optimality*.

In laboratory experiments, dynamics of these auxiliary codes is described in terms of *criticality,* i. e. working near phase transition boundaries.

I approach the problem of relating criticality to optimality of such neural activities using a new statistical model of error-correcting codes explained in [15]. In this model the thermodynamic energy is measured by the complexity of information to be encoded; this complexity in turn being interpreted as the length of *a maximally compressed* form of this information, Kolmogorov style. As far as I know, the respective Boltzmann partition function based upon L. Levin's probability distributions ([9,10]) has never appeared before [15] in the theory of codes.

In this setting, *the close relationship between criticality and optimality* becomes an established (although highly non-obvious) mathematical fact, and I argue that the results of [1,2,24] allow one to transpose it in the domain of neurobiology.

In the main body of the paper below, Sect. 1 is dedicated principally to the mathematics of error-correcting codes, in particular, of good ones, whereas Sect. 2 focusses on mathematical models of neurological data and related problems from our perspective.

1 Codes and good codes

1.1 Codes

In this article a *code* $\mathcal{C}$ means a set of words of finite length in a finite *alphabet A*, i. e. a subset $\mathcal{C} \subset \coprod_{n \geq 1} A^n$. Here we will be considering only finite codes or even codes consisting of words of a fixed length.

Informally, codes are used for the representation ("encoding") of certain information as a sequence of code words. If code is imagined as a *dictionary*, then information is encoded by *texts,* and all admissible in some sense (for example, "grammatically correct") texts constitute *a language.* The main motivation for encoding information is the communicative function of a language: text can be transmitted from a *source* to a *receiver/target* via a *channel.*

An informal notion of "good" or even "optimal" code strongly depends on the circumstances in which the encoding/transmission/decoding takes place. For example, *cryptography* provides codes helping avoid unauthorised access and/or falsification of information, whereas construction of *error-correcting codes* allows receiver to reconstruct with high probability the initial text sent by the source through a noisy channel.

Of course, it is unsurprising that, after the groundbreaking work by Shannon–Weaver, quality of error-correcting codes is estimated in probabilistic terms. One starts with postulating certain statistical properties of transmitted information/noisy channel and produces a class of codes maximising the probability to get at the receiver end the signal close to the one that was sent. In the best case, at the receiver end one should be able to correct all errors using an appropriate decoding program.

In this paper, I will be interested in a large class of *error-correcting codes* whose quality from the start is appreciated in purely combinatorial terms, *without appealing to probabilistic notions at all.* The fact that good codes lie near/on the so called *asymptotic bound*, whose existence was also established by combinatorial means (see [23] and references therein), was only recently interpreted in probabilistic terms as *criticality* of such codes: cf. [12,15,16].

The class of probability distributions that popped up in this research was introduced in [9]. It involves a version of Kolmogorov complexity, has a series of quite non-obvious properties (e.g. a fractal symmetry), and was subsequently used in [13] in order to explain Zipf's law and its universality.

In the remaining part of this section, I will briefly explain these results.

1.2 Combinatorics of error-correcting codes

From now on, we consider finite codes consisting of words of a fixed length $\mathcal{C} \subset A^n, n \geq 1.$. The cardinality of the alphabet A is denoted $q \geq 2$. If $A, \mathcal{C}$ are not endowed with any additional structure (in the sense of Bourbaki), we refer to $\mathcal{C}$ as an *unstructured code.* If such an additional structure is chosen, e.g. A is a linear space over a finite field $\mathbf{F}_q$, and $\mathcal{C}$ is a linear subspace in it, we may refer to such a code as *a linear one,* or generally using the name of this structure.

Consider, for example, neural codes involved in place field recognition according to [2,24]. Here $q = 2$, and in [24] these codes are identified with subsets of $\mathbf{F}_2^n$, but they are not assumed to be linear subspaces.

Besides alphabet cardinality $q = q(\mathcal{C})$ and word length $n = n(\mathcal{C})$, two most important combinatorial characteristics of a code $\mathcal{C}$ are the (logarithmic) *cardinality*

$$k(\mathcal{C}) := \log_q \text{card}(\mathcal{C})$$

and the *minimal Hamming distance* between different code words:

$$d(\mathcal{C}) := \min \{d(a, b) \mid a, b \in \mathcal{C}, a \neq b\},$$
$$d((a_i), (b_i)) := \text{card}\{i \in (1, \ldots, n) \mid a_i \neq b_i\}.$$

In the degenerate case card $\mathcal{C} = 1$ we put $d(\mathcal{C}) = 0$. We will call the numbers q, $k = k(\mathcal{C}), n = n(\mathcal{C}), d = d(\mathcal{C})$, *code parameters* and refer to $\mathcal{C}$ as an $[n, k, d]_q$-code.

These parameters play the following role in the crude estimates of code quality.

The arity $q(\mathcal{C})$ and the number of pairwise different code words $q^{k(\mathcal{C})}$ essentially determine, how many "elementary units" of source information can be encoded by one-word messages. Codes with very large q (as Chinese characters in comparison with alphabetic systems) can encode thousands of semantic units by one-letter words, whereas alphabetic codes with $q \approx 30$ in principle allow one to encode 30^4 units by words of length 4.

The minimal distance between two code words gives an upper estimate of the number of letters in a code word misrepresented by noisy channel that can still be safely corrected at the receiving end. Thus, if the channel is not very noisy, a small $d(\mathcal{C})$ might suffice, otherwise larger values of $d(\mathcal{C})$ are needed. When $d(\mathcal{C})$ is large, this slows down the information transmission, because the source is not allowed to use all q^n words for encoding, but still must transmit a sequence n-letter words.

Finally, when solving engineering problems, it might be preferable to use structured codes, e.g. linear ones, in order to minimise encoding and decoding efforts at the source/target ends. But in this paper, I will not deal with it.

In the remaining part of the paper, in particular, in discussions of optimality, the alphabet cardinality q is fixed.

1.3 From combinatorics to statistics: I. The set of code points and asymptotic bound

Consider an $[n, k, d]_q$-code $\mathcal{C}$ as above and introduce two rational numbers: the (relative) *transmission rate*

$$R(\mathcal{C}) = R_q(\mathcal{C}) := \frac{[k(\mathcal{C})]}{n(\mathcal{C})}, \tag{1.1}$$

and the *relative minimal distance*

$$\delta(\mathcal{C}) = \delta_q(\mathcal{C}) := \frac{d(\mathcal{C})}{n(\mathcal{C})}. \tag{1.2}$$

In (1.1) we used the integer part $[k(\mathcal{C})]$ rather than $k(\mathcal{C})$ itself in order to obtain a rational number, as was suggested in [12]. The larger n is, the closer is (1.1) to the version of $R(\mathcal{C})$ used in [15,16] and earlier works.

According to our discussion above, an error-correcting code is *good* if in a sense it maximises simultaneously the transmission rate and the minimal distance.

A considerable bulk of research in this domain is dedicated either to the construction/engineering of (families of) "good" error-correcting codes or to the proofs that "too good" codes do not exist. Since a choice of the transmission rate in a given situation is dictated by the statistics of noise in a noisy channel, we may imagine this task as maximisation of $\delta(\mathcal{C})$ for each fixed $R(\mathcal{C})$.

In order to treat this problem as mathematicians (rather than engineers) do, we introduce the notion of a *code point*

$$cp(\mathcal{C}) := (R(\mathcal{C}), \delta(\mathcal{C})) \in [0, 1]^2 \cap \mathbf{Q}$$

and denote by V_q the set of code points of all codes of given arity q. Let the latter set be P_q.

Let U_q be the closed set of *limit points* of V_q. We will call *limit code points* elements of $V_q \cap U_q$. The remaining subset of *isolated code points* is defined as $V_q \setminus V_q \cap U_q$.

More than thirty years ago it was proved that U_q consists of all points in $[0, 1]^2$ lying below the graph of a certain continuous decreasing function α_q:

$$U_q = \{(R, \delta) \mid R \leq \alpha_q(\delta)\}. \tag{1.3}$$

Moreover, $\alpha_q(0) = 1$, $\alpha_q(\delta) = 0$ for $1 - q^{-1} \leq \delta \leq 1$, and the graph of α_q is tangent to the R-axis at $(1, 0)$ and to the δ-axis at $(0, 1 - q^{-1})$.

This curve is called *the asymptotic bound*. For a modern treatment and a vast amount of known estimates for asymptotic bounds, cf. [23].

Thus, an error-correcting code can be considered a good one, if its point either lies in U_q and is close to the asymptotic bound, or is *isolated*, that is, lies above the asymptotic bound.

The main result of [12] was the following description of limit and isolated code points *in terms of the computable map cp : $P_q \to V_q$ rather than topology of the unit square.*

We will say that a code point $x \in V_q$ has *infinite* (resp. *finite*) *multiplicity*, if $cp^{-1}(x) \subset P_q$ is infinite (resp. finite).

1.3.1 Theorem [12]

(a) Code points of infinite multiplicity are limit points. Therefore isolated code points have finite multiplicity.

(b) Conversely, any point (R_0, δ_0) with rational coordinates satisfying the inequality $0 < R_0 < \alpha_q(\delta_0)$ (resp. $0 < R_0 < \alpha_q^{lin}(\delta_0)$) is a code point (resp. linear code point) of infinite multiplicity.

The existence of isolated codes is established, but we are pretty far from understanding the whole set of them. According to our criteria, an isolated code having the transmission rate matching the channel noise level would be really the best of all good codes. But it is totally unclear how such codes can be engineered. To the contrary, trials and corrections might lead us close to a crossing of an appropriate asymptotic bound, and we will later stick to this subclass of "optimal codes".

1.4 From combinatorics to statistics: II. Crossing asymptotic bound as a phase transition

The main result of [15] (see also [16]) consisted in the suggestion to use on the set P_q the probability distribution in which energy level of a code is a version of its *Kolmogorov complexity*. This distribution was introduced and studied by L. Levin. Furthermore, we proved that the respective Boltzmann partition function using this distribution produces the phase transition curve exactly matching the asymptotic bound α_q.

In order to make more precise the analogy with phase transition in physical systems, it is convenient to introduce the function inverse to α_q which we will denote β_q. This means that the equation of asymptotic bound is now written in the form $\delta = \beta_q(R)$, and the domain below α_q is defined by the inequality $\delta < \beta_q(R)$.

It was argued in [16], sec. 3, that with our partition function, the transmission rate is a version of density, whereas the curve $\delta = \beta_q(R)$ becomes an analog of the (temperature, density)-phase transition curve.

1.5 Criticality and optimality for error-correcting codes

Imagine now that the source can try various codes whose transmission rate lies in a small neighbourhood of a chosen value, and, after feedback from the target, decide whether the chosen code is better than the previously tested ones. Then, dynamic statistical characteristics of this trial and error attempts below the asymptotic boundary will show signatures of criticality: cf. [7,8,22].

2 Neural encodings of stimulus spaces

2.1 Spatial maps in the brain I: orientation and navigation in a familiar territory

In this subsection, I briefly survey information about position and functioning of those neural networks in brain that encode the local positions and navigation of an animal in the world. For much more details and references, see [1,6], and reference therein.

It is postulated that a given type of stimuli (topographical one, or, say, semantic one as in [17,20]), can be modelled via a topological, or metric *stimuli space X*. Furthermore, brain reaction to a neighbourhood of a point in X is modelled by spiking activity of a finite set of neurons $\mathcal{N}_X$.

In turn, measurable data of such an activity are simplified in order to obtain *a bijective correspondence* between elements of a certain finite covering of X by subsets $\mathcal{U} := (U_i \,|\, i = 1, \dots, n))$ and certain subsets of $\mathcal{N}_X$, "cell groups". Namely, if the animal is in U_i, the neurons in the respective cell group "fire significantly above baseline within a broad (about 250 ms) temporal window" ([1], p. 2.)

Encoding firing neurons by 1 and remaining ones by 0, we thus construct a binary code $\mathcal{C} \subset A^n$ whose words correspond to local positions of the animal, whereas various paths of the moving animal are translated by sequences of code words.

Moreover, combinatorics of the code $\mathcal{C}$ reflects combinatorics of the mutual intersections of local neighbourhoods (U_i). Namely,

$$w \in \mathcal{C}(\mathcal{U}) \quad \text{iff} \quad \left(\bigcap_{i \in \mathrm{supp}\,(w)} U_i \right) \setminus \left(\bigcup_{j \notin \mathrm{supp}\,(w)} U_j \right) \neq \emptyset \tag{2.1}$$

where for a binary word w, $\mathrm{supp}\,(w)$ consists of all positions where the respective letter of W is 1.

See [24] or a brief account in [14], sec. 2.3, for further information on the topological content of such encoding.

2.2 The spatial maps in the brain II: study of a new territory and/or transition between two different familiar territories

In this subsection, I will briefly discuss dynamics of neurons and neural nets responsible for place recognition, and their topography in brain. Here I heavily rely upon [6].

(a) Neurons in $\mathcal{N}_X$ above are called *place cells*. They are located in hippocampus.

Moreover, in the study of new territories and passage between territories a very important role is played by additional classes of cortical cells:

(b) *Grid cells* in parahippocampal entorhinal brain regions encode coordinate grid, functionally similar to the meridians and parallels grid in the maps of earth surface. However, grid pattern is hexagonal rather than tetragonal.

Head direction cells encode the axial position of an animal, *border cells* fire in a neighbourhood of the border of the territory.

(c) When an animal moves from one local part of the environment to another local part which is already encoded by a certain group of place cells, the respective encoding is extracted from memory where it is stored.

However, if an animal enters an unexplored, or recently changed, local part of environment, a very specific type of neural activity must accommodate the exploration and encoding of geometry of this local part. In particular, signals from retina must find their way to the place cells, and be transformed into the respective code. This code

is not unique, it depends of the choice of covering of this local part, which is formed dynamically during the exploration process.

2.3 Where good error-correcting codes might be needed?

The first tentative answer to this question that comes to mind is that codes (2.1) themselves must be optimal ones: it would be reasonable to expect this as a condition for efficient space navigation.

However, it is easy to see that combinatorics of the covering can be such that minimal Hamming distance between two code words is small, for example just 1. Namely, call a code $\mathcal{C} \subset \{0, 1\}^n$ *simplicial one*, if for each $w \in \mathcal{C}$ and each $v \in A^n$ with supp $v \subset$ supp w we have $v \in \mathcal{C}$. These codes naturally appear in (2.1), and they are often quite bad by the standard criteria.

In this note, I argue that good error-correcting codes must be looked for at a higher level ("upstream"): these codes presumably are formed when a brain learns to *produce new place recognising neural codes*, and to projecting local maps into entorhinal more extended maps. Thus, the search of good error-correcting codes is an *intermediate process*, that is essential in the transfer of information from, say, retina, to hippocampal or entorhinal cells.

Moreover, I argue that the formation of such a code can be modelled by the source/channel/target scheme, and that the optimization of the code, as suggested by [GaRiVe03], starts with the stage of source compression. The main new insight is the suggestion that appropriate mathematical model of this optimization is a choice of a Kolmogorov optimal short program capable to generate possibly much longer source data.

Then, according to the philosophy explained in Sects. 1.4–1.5 above, dynamics of this choice might generate criticality signature related to the phase transition near the asymptotic bound.

2.4 Pro and contra arguments

Arguments against the conjecture that (an approximation to) the Kolmogorov complexity might be used by brain as criterium of good source compression mostly concern unusual properties of complexity and the relevant Levin's probability distribution.

Below I will briefly state two such arguments and collect the respective counter arguments.

A. Exact mathematical definition of Kolmogorov complexity can be applied only to a potentially infinite set of objects that can be effectively encoded, say, by all finite words in a given finite alphabet, or else by texts in a given finite alphabet. Then complexity produces the encoding of comparable length for most of these data, but a much shorter one for infinite and infinitely rare minority of them.

In fact, the level of compression approximating Kolmogorov's one is very clearly visible already in the experiments/data bases related to initial fragments of the potentially infinite list whose length does not exceed, say, 10^4 or 10^6: cf. [4,5] and a detailed discussion in [13].

It is less clear, how large is the volume of data needed to be encoded for efficient space navigation. In [6] where this picture based on experiments on rodents is surveyed, the question is asked whether human brains represent space navigation, or other information, using the same principles as rodents' brains. If this is true, then certainly human minds have to encode, store, and recall volumes of information sufficient to see effects of Komogorov's efficiency.

At least some research suggests that the answer must be positive. In particular the paper [6] itself refers to it in Concluding Remarks. The localisation of orientation/navigation neural networks in human hippocampus was also confirmed by the study [11] of licensed London taxi drivers. In the pre-GPS tracker's era, taxi drivers had to pass an intense two-years training for orientation in London. It turned out that the total mass of hippocampal place cells of the licensed drivers was considerably larger that in the control group. Assuming that in human's brains the same principles of encoding are used, one can see that already at this volume of data, compressed encoding might be efficiently used.

For this author, an even more convincing argument is the fact of universality of Zipf's law and suggestion in [13] that it is explained by L. Levin's probability distribution, for which probability of the use of an encoded object is a version of the inverse exponential Kolmogorov complexity.

More generally, amply discussed in the literature idea of "semantic spaces" (cf. [17] and references therein) suggests that the encoding and use of linguistic data in human brains has strong parallels with those in space navigation.

B. The model of criticality used, say, in [19] and earlier works on neural activity differs from the one used here. In particular, divergence of specific heat as function of temperature is often invoked.

My main argument defending asymptotic bound for error-correcting codes as the appropriate phase transition curve is that it supplies a strong intrinsic mathematical background for the study of the statistics of large volumes of information used and stored in brain.

Criticality as behaviour of a neural system seeking for good encoding is also intuitively very natural, if one takes in account that although inverse exponential complexity is not computable, but it is "computable from above" in a precise mathematical sense.

Finally, inverse exponential Kolmogorov complexity has very strong fractal and self-similarity properties. For all this, see [15,16]. To the contrary, intuitively natural code counting statistics, using e.g. Shannon's Random Code Ensemble, decidedly overlook good codes lying near the asymptotic bound: cf. [16], subsections 1.3–1.6.

References

1. Curto, C., Itskov, V.: Cell groups reveal structure of stimulus space. PLoS Comput. Biol. **4**(10), 1–13 (2008)
2. Curto, C., Itskov, V., Veliz-Cuba, A., Youngs, N.: The neural ring: an algebraic tool for analysing the intrinsic structure of neural codes. Bull. Math. Biol. **75**(9), 1571–1611 (2013)
3. Dehaine, S., Brannon, E. (eds.): Space, time and number in the brain. Elsevier Academic Press, Amsterdam (2011)
4. Dehaene, S., Mehler, J.: Cross-linguistic regularities in the frequency of number words. Cognition **43**, 1–29 (1992)
5. Delahaye, J.-P.: Les entiers ne naissent pas égaux. Pour la Science **421**, 80–85 (2012)
6. Derdikman, D., Moser, E.: A manifold of spatial maps in the brain. In: [3], pp. 41–57
7. Gastpar, M., Rimoldi, B., Vetterli, M.: To code or not to code: lossy source-channel communication revisited. IEEE Trans. Inf. Theory **49**(5), 1147–1158 (2003)
8. Giusti, C., Itskov, V.: A No-Go Theorem for one-layer feedforward networks. Neural Comput. **26**, 25–27 (2014). doi:10.162/NECO_a_00657
9. Levin, L.A.: Various measures of complexity for finite objects (axiomatic description). Soviet Math. Dokl. **17**(2), 522–526 (1976)
10. Levin, L.A.: A concrete way of defining measures of complexity. Soviet Math. Dokl. **18**(3), 727–731 (1976)
11. Maguire, E.A., Gadian, D.G., Jongsrude, I.S., Good, C.D., Ashburner, J., Frackowiak, R.S., Frith, C.D.: Navigation-related structural change in the hippocampi of taxi drivers. Proc. Natl. Acad. Sci. USA **97**(8), 4398–4403 (2000)
12. Manin, Y.I.: A computability challenge: asymptotic bounds and isolated error-correcting codes. In: Dinneen, M.J., et al. (eds.) WTCS 2012 (Calude Festschrift), Lecture Notes in Computer Sci. 7160, pp. 174–182 (2012). Preprint arXiv:1107.4246
13. Manin, Y.I.: Zipf's law and L. Levin's probability distributions. Funct. Anal. Appl. (2014). doi:10.107/s10688-014-0052-1. arXiv:1301.0427
14. Manin, Y.I.: Neural codes and homotopy types: mathematical models of place field recognition. Moscow Math. J. **15**, 741–748 (2015). arXiv:1501.00897
15. Manin, Y.I., Marcolli, M.: Error-correcting codes and phase transitions. Math. Comput. Sci. **5**, 133–170 (2011). arXiv:0910.5135 [math.QA]
16. Manin, Y.I., Marcolli, M.: Kolmogorov complexity and the asymptotic bound for error-correcting codes. J. Diff. Geom. **97**, 91–108 (2014). arXiv:1203.0653
17. Manin, Y.I., Marcolli, M.: Semantic spaces. 32 pp. arXiv:1605.04328
18. Meteyard, L., Cuadrado, S.R., Bahrami, B., Vigliocco, V.: Coming of age: a review of embodiment and the neuroscience of semantics. Cortex **48**, 788–804 (2012)
19. Mora, T., Deny, S., Marre, O.: Dynamical criticality in the collective activity of a population of retinal neurons. arXiv:1410.6769 [q-bio.NC]
20. Manin, D.Y.: Zipf's law and avoidance of excessive synonymy. Cognit. Sci. **32**, 1075–1098 (2008). arXiv:0710.0105 [cs.CL]
21. Nonnenmacher, M., Behrens, C., Berens, P., Bethge, M., Macke, J.: Signatures of criticality arise in simple neural models with correlations. 36 pp. arXiv:1603.00097 [q-bio.NC]
22. Tkačik, G., Mora, T., Marre, O., Amodei, D., Palmer, S., Berry, M., Bialek, W.: Thermodynamics and signatures of criticality in a network of neurons. Proc. Natl. Acad. Sci **112**(37), 11508–11517 (2015)
23. Vladut, S.G., Nogin, D.Y., Tsfasman, M.A.: Algebraic Geometric Codes: Basic Notions. Mathematical Surveys and Monographs, vol. 139. American Mathematical Society, Providence, RI (2007)
24. Youngs, N.E.: The neural ring: using algebraic geometry to analyse neural rings. arXiv:1409.2544 [q-bio.NC], 108 pp