Review Article

Annals of Behavioral Neuroscience

Are Neural Transactions in the Retina Performed by Phase Ternary Computation?

Johnson AS^{1, 2} and Winlow W^{1,3, 4 *}

¹Department of Biology, University of Naples, Federico II, Via Cintia 26, 80126 Naples, Italy

²NPC Newton, Villelongue de la Salanque, France

³Institute of Ageing and Chronic Diseases, The Apex Building, West Derby Street, University of Liverpool, Liverpool, L7 8TX, UK

⁴NPC Newton, Preston, United Kingdom

***Correspondence:** William Winlow, Department of Biology, University of Naples, Federico II, Via Cintia 26, 80126 Naples, Italy, E-mail: bill.winlow@gmail.com

Received date: August 05, 2019; Accepted date: September 25, 2019; Published date: September 30, 2019

Abstract

Substantial evidence has accumulated to show that the action potential is always accompanied by a synchronized coupled soliton pressure pulse in the cell membrane, the action potential pulse (APPulse). Furthermore, it has been postulated that, in computational terms, the action potential is a compound ternary structure consisting of two digital phases (the resting potential and the action potential) and a third-time dependent analogue variable, the refractory period. Together, with the APPulse, these phases are described as the computational action potential (CAP), which allows computation by phase. The nature of transmission, and thus computation across membranes, is dependent upon their structures, which have similar components from one neuron to another. Because perception and therefore sentience must be defined by the capabilities of the brain computational model, we propose that phase-ternary mathematics (PTM) is the native mathematical process underlying perception, consciousness and sentience. In this review, we take the CAP concept and apply it to the working of a well-defined neural network, the vertebrate retina. We propose an accurate working computational model of the retina and provide an explanation of computation of the neural transactions within it using PTM, and provide evidence that could form the basis of understanding neural computation within the entire nervous system. Evidence is presented of phase ternary computation (PTC), defined in phase ternary mathematics and shows an exact mathematical correlation between the activity of the amacrine cells, the bipolar cells and ganglion cells of the retina, once these cells have been activated by light falling on the cones. In this model, the computation of luminosity of multiple cones synapsed to a bipolar cell is performed by phase ternary mathematics at the points of convergence of CAPs. Redaction by the refractory periods of converging CAPs eliminates all but the leading APPulse resulting in sampling and averaging. In phase ternary analysis (PTA), the physiology of synapses defines their primary action as latency changers, changing the time taken for impulses to travel between points of convergence. This paper describes a novel type of computation, PTC, with evidence that it is the fundamental computational method used by the retina and by association the rest of the brain. By comparing the morphology of neurons it is now possible to explain their function singly and in networks. This has profound consequences both for our understanding of the brain and in clinical practice.

Keywords: Retinal coding, Retinal neural network, Phase ternary computation, Computational action potential, Neural transactions, Action potential pulse, Timing by synchronization, Error correction, Mean sampling, Computation of luminosity

Introduction

The basic circuit diagram and physiology of the retina are well researched and the activities various components are now known in response to light. We know from experimental evidence that computation of patterns of activity from 130 million light receptors are coded into 1000,000 visual neurons [1,2] so that the image is perceived correctly in the visual cortex at different light intensities. This is also a huge reduction from input to output requiring error free efficiency of computation, which has to be accounted for. In the retina there is no clock or method to synchronize to the natural brain neural network (BNN) but coding takes place accurately and efficiently and there is probably only a single mathematical method for this to take place for the network to remain error free.

For synchronization to occur in the retina there must be a mechanism inherent in the action potential to coordinate activity and this can only occur at the convergences of the action potentials themselves. This requires a mechanism to correct the temporal relationship of one action potential to another. This is not possible with a binary action potential based upon a spike but is possible with the computational action potential (CAP). This is a phase ternary pulse based upon the threshold, refractory and resting states. The importance of phase ternary computation (PTC) rather than binary computation within the nervous system has been dealt with elsewhere [3].

Binary computation was primarily accepted because of the inexpensive ease of manufacture of binary transistor chips and research has assumed, without evidence that the computational parameters must exist to service binary computation within a physiological BNN. Knowledge in the fields of computation and AI has diverged from neurophysiology and physiologists now have intimate knowledge of the action potential at the site of the membrane in terms of both transmission dynamics and activity of the soliton pulse. Recently we [3] deconstructed the action potential and APPulse into the computational elements of threshold (defined by the moment of opening of the Na channels) the refractory period and the resting period; these were notated in balanced phase ternary as 1,-1, and 0 respectively. The APPulse [4-6] and the Computational Action Potential CAP requires only single action potentials to compute in parallel. Hodgkin Huxley Cable Theory [7] may also be thought of as a description of PTC because in their paper they describe spike, refractory and resting phases.

However, Cable Theory cannot work in the retina because computation at the timing precision of more than 1ms is not possible in the retinal network. Therefore, timing would have to be redacted by other mechanisms. In artificial neural networks this error is adjusted for precision artificially, e.g. by introducing temporal coding of spikes into the model [8] but there is no evidence from the histology or the physiology of the retina that such an equivalent biological system exists.

Of paramount importance in consideration of computation in the nervous system is the mechanism by which individual action potentials interact. In artificial intelligence (AI) models, interaction takes place at gated nodes (Figure 1d). In a brain neural network these are replaced by the synapses, points of convergence and divergence (Figure 1a,b and c). At synapses the computational dynamics of the CAP terminate, until a new CAP is triggered postsynaptically [3]. Thus, the coding within the synaptic space is essentially binary in that an incoming spike may or may not trigger an outgoing spike. In a sense synapses play a secondary, but very important role in PTC, by altering the latency of postsynaptic responses. Long-term synaptic modifications may still occur to modify the output of the system. Contemporary understanding of computation in a computational or physiological BNN is limited to the ability of synapses to modify the properties of their connections according to the previous stimulus. Many models have been created for computing for example spike-timing-dependent plasticity [8] both in neuroscience and in AI: all use calculations of temporal positioning calculated from binary spiking neurons. PTC simplifies this process by eliminating the need for additional complexity.

Mechanical surface waves are known to accompany action potential propagation [9] in the form of solitons [10]. Coupling of the mechanical activity and Hodgkin Huxley Cable Theory have been suggested [10].

We have recently postulated that action potentials take the form of an APPulse i.e. an action potential coupled with a soliton in the nerve membrane [5,6], rather than the more orthodox view that the mechanism of regenerative action potentials can be explained by Cable Theory. In our view, the APPulse mechanically deforms membrane ion channels allowing entry of Na⁺ into the cell, which then reaches threshold to trigger a regenerative action potential. This is critical to a computational analysis of functional neural networks in the brain, which will in future be important in the development of artificial

intelligence (AI) because realistic models of neural coding in the nervous system need [11] first to be proposed, tested and understood.

Timing in a neural network

Intrinsic to AI is the belief that gated timing of synapses synchronizes computation. Where computational models have been proposed [12], models that attempt to use conventional computer theory conflict with the known neuronal morphology and function. In these models the timing of action potential and computation through the network is synchronized as a conventional computer by clock. We know that “Neuronal morphology and function are definitely interrelated” [13] indicating differing latencies and transmission behaviors for morphologically different neurons. In the retina timing of action potential from one point to another depends upon its transmission speed and the length of the neuron – both of which are variable from one neuron to the next. The neuronal connections in the retina do not permit a Turing based machine. There is no functioning clock where timing of computation can be coordinated in the retina and yet sight is extremely accurate robust and error free. The only model that will sustain computation in the retina is by synchronization and error correction using the exact physiology described by neuronal connections physiology and morphology. This is the model proposed by the CAP.

The CAP demonstrates that the action potential is capable of synchronizations if its phase ternary mathematics are calculated over the network. Thus, timing elements are unnecessary in a synchronized system. Basic computation requires that input be transferred to output in a logical manner with timing between input and output fixed. Timing is critical in a BNN and just as importantly the precision of timing of a pulse allows us to judge its constituent mechanisms. Computation is not just in phase but it is in temporal phase. In a neural network the rules of parallel processing are considerably different for PTC as compared with binary computation. Binary networks must be gated but a phase ternary network synchronizes signals (Figure 1).

In the Lateral Geniculate Nucleus 1,000,000 distinct parallel inputs from the optic nerve merge into a connected network simultaneously [1,2]. To distinguish between the activities of 1 input from 999,999 others requires a sufficient temporal gap. The combined activity forms the complete image that we see. If 1,000,000 inputs converge to the same network (effectively layer t1 of Figure 1d), each separate input must be able to be

distinguished from all others. For any individual impulse to be distinguished against all other parallel inputs temporally, each second must be divisible by the number of inputs such that the precision of timing per second can be distinguished. This produces a figure of 1,000,000 Hz, indicating an ideal precision below 1 microsecond. This is a basic calculation where error and duplication have not been considered and which would increase the computation time. Cable Theory cannot give this level of precision, but the APPulse is estimated to give a precision of 1.5 microseconds based on the distances between ion channels [6].

It is widely accepted that light reactive cells increase frequency of discharge and importantly we know that Retinal ganglion cell (RGC) frequency of discharge is proportional to mean light over RGC receptive fields [14,15]. Activity of the different light receptors has been evaluated [16] and their connections have been traced. Retinal ganglion cells are connected via bipolar cells to many light receptors. Luminosity across the light receptive area instigates activity in its respective retinal cell ganglion. Many light receptors are connected in parallel with different sensitivities to light and will have a range of activities so that as luminosity changes, then so will their combined activity. These parallel frequency differences must be collated to ensure that their mean frequencies represent the bipolar cell frequency, thereby generating meaningful information.

Figure 1: Simplified comparison of computation in a contemporary artificial neural network with a

physiological brain neural network utilizing phase ternary mathematics.

- a) and b) are illustrations of the connectivity of neurons within a physiological brain neural network modified to represent the Computational Action Potential CAP [3] with detail c). d) Is a comparison with a contemporary artificial neural network. In a) and b) action potentials have different latencies as they travel from node to node. Distances between nodes are also different. Points of connection are represented by large brown dots, these are the points of convergences. Impulses travel from left to right as illustrated in red by the threshold and in blue the refractory following the rules of the CAP. Synapses lie between points of convergences and are excluded from the diagram for simplicity; their action in computations is discussed later. There is no system of absolute timing and action potentials arrive according to the timing of their respective neurons. In figure 1 a) (i) the threshold of a CAP arrives at a convergence and passes while the threshold of (ii) is annulled resulting in the phase shift shown in b) two relative time steps later. c) The three steps are shown in detail. Successive phase shifts change the frequency of the CAP resulting in unique computation from phase. In both figure 1a and b signals must be synchronized, in an artificial neural network this is done by timing gates, however no timing exists in the brain to facilitate such a synchrony. Binary does not synchronize but phase ternary does. In a) b) c) and d) all inputs must synchronize. The precision of this input in d) depends upon the timing between t_1 and t_2 and the gating. This timing can be thought of as the precision of the network. In the brain neural network this precision is dependent upon the absolute timing of the threshold.
- c) Detail of green oval region showing three successive moments in relative time marked by small dots. Note CAP (ii) is cancelled changing the output phase of (iv) and affecting the overall frequency. Figure 1a Purple oval in contrast shows two CAP (vi) and (vii) converging together in temporal phase and pass as one. These diagrams illustrate the basis of phase ternary computation: in a neural network containing many inputs and many outputs specific phase temporal patterns are created representing input. This process is described in detail for the retina in figure 5. Computation is therefore a function of all the frequencies from all input neurons simultaneously.

- d) Diagram of a simple contemporary artificial neural network where information is gated in absolute time. In contemporary AI, inputs must code for output in a logical fashion, these are usually represented in binary as 0 and 1. AI assumes that computation is by gated synapses. In addition information passes through the network in an orderly fashion with points of computation t_1, t_2 and t_3 synchronized by timing. In such a model latencies are assumed fixed.

The basic network theory is similar in a b c and d in that unique input is expressed as a pathway through the network to unique output. The difference is that the phase ternary arrangement in a,b, and c) is that at each point of computation/convergence nerve impulses are synchronized against each other therefore redacting error from each layer. The processing of parallel information across multiple threads of information is therefore temporal, dependent upon the minute changes in phase of parallel impulses producing changes in frequency of successive individual CAP. The nodes for CAP computation are the bifurcations and convergences of the neurites, which may also be on postsynaptic membranes. Importantly inputs are not represented by binary but by frequency of impulses assessed from the temporal relationships between thresholds in phase ternary.

All computations that take place compute changing frequencies, there is no evidence of a clock. In figure 1d, a conventional network illustrates defined binary output representative of the binary input; in figure 1a action potentials that enter the network are computed to give specific frequency characteristics of outputs and are shown as CAP with a threshold in red followed by a refractory period in blue, the axons have different timings, latencies and have speeds according to the membrane components. There is no timing possible over the network – the only reliable measurement is the latency from receptors to RGCs. Error must therefore be inherently and simultaneously removed. There are no mechanisms available within the eye for error or noise correction so the only possible method is from synchronization as action potentials travel through the network. There is no present mechanism from computer science or AI that can explain this.

Phase ternary analysis of retinal coding

We postulate that synchronization of activity in the retina is via the CAP when considering the structure of

the retina (Figure 2) and the accuracy of visual processing from the receptors cells to the retinal ganglion cells. Although the action potential and the APPulse have ternary mathematics comprised of the resting threshold and refractory periods, their computation processes are different: action potential computation must be calculated from the spike, which is temporally inaccurate, while the CAP is calculated from the threshold. The CAP can be expressed mathematically in balanced phase ternary as $0 + 1$ and -1 . When two physiological action potentials collide they annul due to the two refractory periods. This is illustrated in figure 1c. If two action potential converge in the same direction the result is always addition of the ternary phase in which they arrive, this depends upon the timing of the CAP and the latencies of the neuron (Figure 1a). The timing of computation is critical and as no clock exists in the CNS capable of absolute timing, activity of impulses must be synchronized.

For computation to take place in a neural network, inputs into the network must match outputs in a logical manner (Figure 1a and d). In the retina the activity of the light receptors is represented as the output of the RGCs to the optic nerve. In contemporary binary computing these values are represented as 0 and 1 where 0 is resting and 1 is the spike but in PTM they are expressed as $0 + 1$ and -1 representing resting, point of threshold and refractory respectively, i.e. PTM is defined by the phase structure of the CAP. Note that the refractory period in the CAP does not match traditional ternary mathematics: in the CAP the refractory period always returns -1 when in collision with another CAP and is a distinct variable time related to each point of convergence.

In the visual system, the coding of parallel information commences in the connections between the light receptors and RGCs. Early studies suggested that RGCs are the first clearly spiking neurons, while none of the other retinal neurons were believed to conduct regenerative action potentials [17,18], although amacrine cells were thought to carry dendritic spikes under some circumstances. Over time this view has been called into question and it is now known that amacrine cells can conduct action potentials [19,20] as can some classes of bipolar cells [21,22]. The implication of these findings is that it is now possible to analyze the visual system all the way from the amacrine and bipolar cells using phase ternary mathematics. It should be noted that the APPulse computes using PTM [3] and the soliton will do so in non-spiking neurons, where the threshold is calculated from the soliton activating the opening of ion channels. This is

independent of spikes and is ideally suited to calculations on non-spiking neurons.

The retina uses parallel pathways to encode different features of the visual scene according to the activated state of the parallel pathways. These distinct pathways converge on circuits that mediate a distinct computation allowing for precise sight. In all likelihood, there is only one solution to how the retina codes. The prerequisites are that the changing frequency of parallel inputs from the light cells must be reflected in the changing frequency of parallel outputs in a logical manner so that error is redacted and all information is transformed synchronously. On this basis, we propose that a computational action potential pulse (CAP) [3] is the most appropriate analytical tool for evaluation of network models both in artificial physiological BNNs such as the retina, a peripheral extension of the brain.

Figure 2: The layout of the retina.

To evaluate the computational method used between the cones and the bipolar cells shown as OFF-BC and ON-BC. The rates of firing are known for both the cones and the bipolar cells and the contrast activity frequency response at the exit of the bipolar cell is known to evaluate to the average firing rate over time [15]. Using this knowledge, it is possible to mathematically model the exact mechanism that achieves this computation in the bipolar cells. (Modified from Tsukamoto and Omi [23] under the Creative Commons License (CC BY)).

Synapses and latency

Synapses are conventionally assumed to act as binary gates in neural networks with changing activity altering their transmission properties [24]. Computation has been assumed to take place at the synapses and no further investigation has taken place to examine other forms of computation. Microscopic examination shows that synapses, although closely spaced, separate neuron from neuron and their primary purpose is to modulate the activity of the post synaptic membrane affecting latency, leading in many cases to an action potential. The timing latency involved in the activation of the postsynaptic membrane is fast and is the primary function and depends upon transmitter as discussed in our previous work [3]. Any further chemical activity caused by the release of neurotransmitter will be secondary and will only affect other nearby neuronal membranes after the passage of the APPulse. Just as with CAP convergence [3] activation of the membrane is the first activity of synapses and the refractory period will block any further CAP at the point of connection between the synapse and the membrane. Either synaptically derived or neurohumoral chemical interference is considered secondary and affecting frequency due to modification of the refractory period.

In PTC the activity between two post synaptic synapses converging on the same membrane is determined by the change of phase generated by the different latencies of synapses as they pass a point of convergence and to the transmission latency of the neuron. If they are both in phase both CAPs pass, but if either is delayed the first will cancel the second. This is illustrated in figure 1a, b and c. Synapses in this model affect the latencies of the transmission and cause inhibition by changing phase of one or both CAPs. Two pathways containing different neurotransmitters will have different latencies dramatically changing computation. This computation takes place by single action potential and is precise. Neurotransmitters all have different release latencies and response latencies. A neural network with more than one neurotransmitter will have differing latencies within it. Latency changes may also occur due to external hormones, which compete non-linearly with neurotransmitters. Finally, latency is also a product of the transmission dynamics of neurons. Any chemical interaction can therefore only affect subsequent frequency after the initial CAPs have passed because the CAP is produced first and travels at a faster speed than diffusion.

The retina is a regular array of neurons set in defined

patterns and layers. In this array, synapses evidently have important functions in computation and the phase ternary model proposes that synapses act as latency changers to the phase computation as well as inhibitors. Synapses therefore have the ability to change a CAP ternary phase changing the computation at the convergences of connections and changing the direction of travel of CAP along neurites. Any error at synapses is redacted by synchronization. We have addressed the important issue of error redaction elsewhere [3,6].

Computational coding of retinal signaling

Cone photoreceptors are reactive to bright light of different wavelengths, depending upon their type, according to the structure of the visual pigments or opsins in their outer segment regions [16,25]. Cones have been categorized into many types [26]. Some are hyperpolarized by light and some also depolarized by dark [27], and a release of glutamate at the cone output synapse, the cone pedicle, transfers this signal to the postsynaptic neurons.

The molecular mechanisms of photon-transduction to release of neurotransmitter are similar in rods and cones, but the quantitative details of the function of each of the biochemical and biophysical processes in this pathway differ between receptor types [16]. Information is processed in parallel by the structured network of the retina in a feed forward from the rods and cones to RGCs from where they project to the optic nerve (Figure 2). The timing and output from the basal ganglion cells have been known for some time, but more recently the cones and the bipolar cells have been extensively studied for connectivity and activity. Information moves progressively from the cones to the RGCs. The total information represented by the RGCs over time therefore reflects the information from the rods and cones over the retina.

Cone types are reactive to contrast banding, the activity of each type depending upon the level of light within that band [14,15,28]. The organisation of the cones in relation to the bipolar cells has been studied [26,29,30]. Connectivity is complex but there are rules to the network. Bipolar cells are arranged in overlapping cone receptive fields [23,31]. The output from the cones to the bipolar cells and to the individual RGCs is known, as are the relative numbers of connections. Individual cone activity is frequency dependent upon luminosity within cone-type bands. The frequency of output of action potential from the RGCs is proportionate to light

contrast at the cone receptive fields [14,15] for each type and are similar in frequency to the output of the cones.

In the case of the retina, much of the information processing is performed through gap junctions and neurons that do not spike [32]. Gap junctions have a shorter latency than chemical synapses [33]. The CAP can therefore be represented in spiking neurons by either the action potential or APPulse but in non-spiking neurons we conjecture that graded activity may also be accompanied by the by a soliton wave to allow opening of membrane channels.

The mathematical derivation of mean sampling from phase ternary analysis

Cone-bipolar cell connections: Recordings from RGCs of the activity of single bipolar cells demonstrate computation when multiple cones are activated. The result is temporal activity related to mean field luminosity [14,15]. The ranges of frequencies detected in output were similar to the input. The calculation of luminosity change across the cone receptive field by the bipolar cell is therefore calculated by the change in frequency, at any moment, of single action potentials compared with the mean. This computation of average luminosity across the RGC receptive field must take place before the RGC as the connection between the bipolar cell and the RGC is one to one (Figure 2). This is a mathematical network template and can be used to resolve the computational functioning of the bipolar cell. Individual cones are responsive to light over the range of that type of cone [25]. There are two sites of interest on the bipolar cell where interaction may take place, within the synapses themselves or where they converge on the post-synaptic membrane, i.e. on the RGC.

Bipolar cells encode diverse temporal image signaling from the cones in a subtype-dependent manner to initiate temporal visual information-processing pathways [14]. Cone bipolar cells individually react to luminosity by increasing frequency (f) of action potentials (i) [14,15,34] within the bandwidth of the cone type [14]. Figure 2 illustrates the distance between cone-bipolar cell (BC) activation points and the point of convergence. Computation cannot take place from one synapse that would affect another except by CAP due to the distances between multiple cone/bipolar synapses and their interaction points. A graded potential or action potential, depending on bipolar cell type, precedes any possible chemical interference between the separated cone-bipolar cell synapses that converge and this will proceed

towards the RGC before any intra-synapse activity. The point of convergence and computation is therefore the site at which the outputs from the cone/bipolar synapses connect (Figure 2 Point of convergence). This membrane is situated some distance proximal to the synapses. In this case, the action of synapses in PTM is to depolarize the membrane changing it to a refractory state and simultaneously adding latency to the action potential. Latency is defined as the amount of time necessary to pass between convergent/divergent points in the network. The result during convergence of CAPs is to continuously sample the frequency change from the mean (Figure 3).

Figure 3: Mean sampling. The output frequencies from two cones converge at the point of convergence of a bipolar cell. a) Represents the CAP travelling from the Cone BC synapse and the point of convergence (red) on the membrane as shown in b. Each red line represents the same point of convergence as successive CAP converges. The first CAP inactivates the whole of the membrane for the refractory period annulling any within the sampled refractory time. Here activity is sampled into groups of time t . Time t is a set period of time after a CAP reaches the convergence of the bipolar cells where the activity of all the cones are combined. (t) Represents a grouping of Cone BP 1, cone BP 2, etc. Any impulses arriving within the grey refractory area are inactivated by the membrane refractory period. c.) Shows the interpreted result after the pulses within the selection area are redacted and represents the information passed to the RGC. The pulses in the interpreted result have a fixed time t and each has a variable time indicated as t_1 t_2 t_3 samples of frequency

change from the previous. $t_1 + t_2 + t_3$ etc. sampled by time represents the mean of light focused on the cone receptive field. If many cones are arranged in parallel Cone BP 1, cone BP 2, Cone BP 3 etc. The mean of, $t_1 t_2 t_3$ will be the mean luminosity of light over that time period. The accuracy therefore of calculating the mean depends upon the timing of the refractory period. From this we can see that the output from the bipolar cells to the retinal ganglion cells is a mean value of the added frequencies from its input, achieved by mean sampling at the point of convergence. The first CAP/ action potential will pass this point and leave a refractory period blocking any subsequent activity. This results in the mean sampling of successive CAPs originating from different light receptors. The refractory period is a product of the ion channels and membrane components and is thus stable. This refractory period measured in absolute time deselects successive bands of CAPs, which results in the time difference between these selective bands being a mean sample of the input. This can be expressed as:

$$m^r = \frac{(\sum^r f^r)}{n}$$

Where m^r is the resultant sampled mean from the collisions, f^r is the count of impulses arriving from each light receptor during the refractory period r . within each suppressed refractory period; $\sum^r f^r$ is the summation of this number, n is the number of convergent light receptors. The time between r and the next impulse is dependent only upon the timing of the next light receptor t_1, t_2, t_3 as shown in this change in time t^n between refractory periods is a measure of mean frequency change.

BP = bipolar cell.

Computation of mean frequency by the bipolar cells revolves around there being only a threshold followed by a refractory period. It is equivalent to mean sampling of the frequencies as demonstrated in figure 3, i.e. the output from the bipolar cells to the retinal ganglion cells is a mean value of the added frequencies from its input. This is achieved by mean sampling at the point of convergence. This method of evaluating mean sampling fits with the precise definition of phase ternary described in [3,4,6]. This way of computation is only possible between single action potentials in parallel, because there is a defined threshold with precise timing of less duration than the refractory period. It is impossible to use a spike-timed action potential for this computation, as accurate timing of the spike to less time than the refractory period

would need concurrent computation for which there is no evidence. Phase ternary mean sampling is expressed mathematically as the grouping of phase ternary activity at an interaction point followed by addition. As yet there is no other known mechanism that can explain the mathematics as resolved from experimental data, the physiology and connections of the retina. Although the CAP can be formed from the Hodgkin-Huxley Cable Theory, for frequency mean sampling to take place each CAP must have a defined beginning such as a measurable threshold.

The Computational Action Potential [3] simplifies the mathematics by recognizing that each CAP may be in three states, resting, threshold, or refractory. For a CAP to pass during convergence the threshold timing on convergence must be synchronized with another CAP in phase or does not encounter a refractory period for the CAP to progress. At a point of convergence (Figure 2), the mean output can be calculated from the sampling of phase ternary representations of the converging cone action potential as they interfere if we assign each cone activity the correct phase ternary relationship. There are many types of bipolar cell [25] and the refractory periods at convergence will almost certainly be different between types, this will give a different imprint of the total average over time, marking each bipolar cell as unique in its pattern of activity. This is reflected in the exact phase tuning of the CAP. Information from each cone receptive field is also distinct being matched to the time of distinct refractory periods and mean activity.

The cone RGC receptive fields: Cones are grouped into overlapping receptive fields [14,25,26] with about 12 cones connected directly to a bipolar cell by synapses (Figure 2) Overlapping fields allow receptive field activity to be compared and individual cone-receptive fields are mapped into the complete retina with each RGC representative of the positioning of its receptive fields. Activity from ganglion cells connecting via bipolar cells to adjacent cone-receptive fields will therefore represent multiple gradients of activity where the information from multiple-cone receptive fields forms a map of activity. As this is a direct connection it can be assumed that this information is transferred intact to the optic nerve and could be read to some extent by the central BNN. If our assumption that the output from the optic nerve to the rest of the brain is phase ternary coded is correct, this would imply that the rest of the brain also computes using phase ternary coding.

The distribution of cones and cone types within the retina ensures that shifts of external activity are registered between the cone receptive fields as light changes. This may be deconstructed to provide data expressing continuous activity gradients across the entire retina. Assimilation and coding of gradients permits shapes to be coded by comparison to other activity on the retina. By sampling and averaging from a number of cones across a receptive field the bipolar cells are able to provide information according to the reactivity of the cones and are collectively able to compute across other receptive fields.

Gross retinal coding: Using knowledge of how individual neurons function, and given the diversity of cones and their connections, it is possible to use the known connectivity and responses of neurons to elucidate and predict the functions of the rest of the retinal network and thus produce exact computational coding. Parallel computation occurs in a network when distinct inputs into the network produce distinct outputs, which are representative of the inputs.

In figure 4, we present a collection of simplified Feynman type diagrams [35] which represent the passage and collisions of CAPs over two points of convergence. This type of diagram allows us to illustrate multiple collisions and relationships over time. Time travels from left to right with computation points marked. Time has been simplified so that each phase is the same time and is equal to the refractory period at each computational step. The diagram illustrates that unique inputs into the system produce unique outputs with unique patterns of activity produced: a simple example of PTM.

The structure and coding within the retina

Assuming that the mathematical relationship is the same between the computation of the first connection, cones-bipolar cells and the rest of the retina, the synchronization and coding of the retina can be easily elicited and conclusions of what is being coded can be drawn. Using the network connections from Tsukamoto and Omi [23] frequency activity can be plotted on a diagram (Figure 5). Furthermore, knowing the resultant output and theory of the parallel coding will eventually provide an opportunity to examine the parallel data streams as they pass to the LGN, through the optic chiasm (OC) and on to the visual cortex [36].

Figure 4: Passage and collisions of CAPs over two points of convergence. Numbers 0 +1 and -1 represent the initial ternary states, resting threshold and refractory of the computation action potential CAP [3]. The network is simplified such that over each connective distance each CAP does not change phase: Timing or latency between points of convergence and are written beside the network. The timing for the latency of the CAP at each point of synchronization is set to whole phase to simplify the illustration; fractional latencies will produce similar computation conformity. For the illustration latency is represented as marked in small type and is simplified to 1 phase between points of convergence. The numbers on the left represent the CAP split into its three ternary phases. The resultant value for any combination at convergence is the grouping then summation of the ternary numbers. In the retina where 12 cones converge onto a bipolar cell phases are grouped before addition indicating the resultant phase at that point. t = time. This does not depend upon neurons spiking but only upon the existence of a threshold and of a refractory period at point of convergence. The components for both exist in bipolar cells and the existence of a spike is not relevant. Phase ternary predicts that there is a point of threshold and refractory period from the mathematics of mean sampling to produce an explanation of the output of the light receptors to the retinal ganglion cells.

The maths of phase ternary convergence is to group the ternary numbers and then correct for a refractory

-1. Presence of the refractory always results in -1 except where the result is no preceding +1. A refractory -1 in any group of phase ternary competing CAP resolves to -1.

+1 (+) -1 (+) -1 = -1 = refractory,
+1 (+) +1 (+) -1 = -1 = refractory
0 (+) 0 (+) 1 = 1 = threshold,
-1 (+) -1 (+) -1 = -1 = refractory.

a.) Inputs all have defined phases. The top two inputs 0 and +1 converge producing activity shown in red. At each point of convergence phase is either changed or not. c.) Demonstrates synchronicity within the network when parallel CAPs converge in a complex system stabilising and correcting error. If either top or bottom pulse is desynchronised a correction is made as in Figure 3c where a point of synchrony is indicated. If there is spontaneous activity it will be redacted. Every CAP internal to this network is similarly corrected for phase by temporal realignment to attain synchronicity over the whole network this synchronisation redacts error. b. Alternative patterns demonstrating distinct computation. c) Demonstrates a point of synchrony and d) is another computation.

(Modified from Tsukamoto and Omi [21] under the Creative Commons License (CC BY)). This is same type of diagram in figure 4. Activity starts at the top of the diagram and moves along the network in any subsequently connected direction with interaction occurring at convergences. Timing is by consecutive and synchronised between points of convergence, in a real retina the timing of the CAP are independent and progress according to the speed of the axon and latency of the synapses. For the illustration latency is represented as a zero phase shift, in reality there might be two or more phase shifts for example where there are slow synapses and neurons. For this example the threshold refractory periods and the resting period are equal and equivalent to each phase. The length (a) to (b) represents 0 shift in the model - should the latency of this neuron be any fractional or complete phase it will make no difference to the demonstration of phase ternary computation. The diagram shows points of convergence and represents the phase collisions taking place with the resultant coded phase changes reflected in the output.

The box at the top represents action potential split into the three ternary phases, the four rows represent four successive phases starting at the bottom of the row and the numbers represent their respective phase.

Coloured numbers represent the number of phase changes from each coloured row from the initial. Numbers in brackets represent the computational moves from the initial cone activity. At point (a), in the first row (Purple), two cones converge to a bipolar cell their values are +1+0=+1. This result is written below each node in the network with the order of computation in brackets:1(1).

The resultant value for any combination of phase ternary is grouping then addition of the ternary phases turning all results to -1 where present. In the retina where 12 cones converge onto a bipolar cell phases are grouped before addition indicating the resultant phase at that point.

Notable points in the network: This is a synchronised system where the precision of resultant phase change is dependent upon the precision of the threshold and its ability to change the membrane to refractory.

Point (b) represents the input into the horizontal cell. Each time this neuron fires all connected neurons fire, this will synchronise the output of the cones to each

Figure 5: Representation of the retina superimposed on figure 2 as a phase ternary connective network.

other and to the rods.

At point (c) phase input from the cones (-1 -1 0) resolves to 0. The continuous feed from the cones synchronises the rod output by changing phase originating from the cones and passing along the HC. This will synchronise the sampling initiation times of the cones and rods over three phases and permit precise error redaction and phase tuning as pulses pass to the RGC.

At point (d) is a refractory synapse -1 that changes +1 to 0. This synapse may not produce an action potential only a refractory period, for illustration we have added this to our model although it is not consequential. Point (c) both inhibits in cone reactive light and in darkness when the rods are active but not the cones. The rods therefore inhibit the resting frequency of the OFF BC when cones are inactive. Point (e) is the convergence of the amacrine cell.

Differences in synapses and the length of neurons will change the coding uniquely but will contain the same information. The exact nature of each element can be accounted for in this model.

Horizontal cells: (HCs) synapse with both cones and rods directly and through triads (Figure 2). HCs are often electrically coupled by gap junctions [31]. In PTM the ON-OFF response [23,37] is dependent upon the phase annulment of succeeding CAP. Activation of a horizontal cell by a cone cell will reset all bipolar cells within its distal connections to the same phase. This synchronization between bipolar cells permits error redaction as described above but more importantly permits phase synchronization between bipolar cells. Any change in synchronicity between ON and OFF bipolar cells due to horizontal cell reduction will therefore lead to a sporadic ON-OFF response. This effect is demonstrated by Chaya et al. [38] where the receptive field was changed when horizontal cells were removed. The horizontal cells therefore synchronize the sampling of mean luminosity over the distance of its connections.

Amacrine cells: Amacrine cells (Figure 2) synchronize and correct for error [39]. They also connect the On-Off basal ganglion permitting phase ternary computation (Figure 5).

What is encoded into these parallel information streams and how?

As streams of parallel phase ternary CAPs successively converge, collide and compute, phase changes will occur across the output of RGC reflecting minute changes in

frequency from the cones/rods. These synchronized phase changes represent gradients of information relating to hue and colour and changing as the light to the retina changes.

Each RGC contains direct information not just about the direct link to the receptive field to which it is closest, but also all other connected cone receptive fields as they interact to complete the image with other types of receptor. Contextually each RGC contains the information on changing luminosity from its directly connected bipolar cell. This is the primary phase determinant, which can be modified in phase by all the other cells connected to that RGC. As convergence continues on the membrane of the RGC, phase changes are modified and corrected by further connectivity as the number of parallel connections increases. The resulting activity in figure 5 is repeated simultaneously across all synchronized RGCs producing coding by tuning the phase of the RGC at the point of parallel temporal computation. This synchronicity produces relative timing over distinct regions of the retina [21,40] and the premotor cortex [41]. It is this synchronization that permits phase ternary CAPs to redact error and function without absolute timing. Individual bipolar cells in the connected retina therefore contain the primary information of their own direct connected bipolar cells but then have convergences with other bipolar cells where information is progressively added by tuning with the result that each layer represents that above. This further integration of activity codes the geometric relationship of each light receptive field of all types into the output to the optic nerve. Thus, at the level of the RGCs the relevant information relating all connected bipolar cells to each other is encoded.

The network is lossless. Error and noise are redacted at each consecutive convergence. No information is lost between the convergence of the cones and the bipolar cells. Adjacent cone receptive fields whatever their sensitivity to hue, position or luminosity will be reflected in the combined output of the RGC in the form of overlapping gradients from the receptive fields.

Conclusions

In this paper we have described a novel type of computation, PTC, with evidence that it is the fundamental computational method used by the retina and by association the rest of the brain. PTC is computation by synchronization of action potentials in a ternary system where, in quantum computational terms, qutrits would be used as default information carriers, rather than the

qubits used in binary systems [42].

Using PTM, we are confident that there is a clear relationship between the CAP/phase ternary association and meaningful retinal computation. In this model, computation takes place without interruption at the convergences of neurons.

Given that the intrinsic morphology and physiology of neuron membranes are similar, it is logical to assume that all neurons act as phase ternary computational components, implying that PTM is the primary form of computation within physiological BNNs.

The computation of luminosity of multiple cones synapsing on a bipolar cell is performed by PTM at the point of convergence of CAPs and not at the synapses. Redaction by the refractory periods of converging CAPs eliminates all but the leading impulse resulting in sampling and averaging. The physiology of synapses defines their action in PTC as latency changers by extension of the refractory period.

The discovery of the phase-ternary CAP may signify a paradigm shift in our understanding of how neurons transmit information and also provides us with an alternative computation theory that is empirically quantifiable in all neurons from invertebrates to humans.

Declarations

Ethical approval

None required

Consent for publication

Both authors consent to publication.

Availability of data and material

All contained within the cited references.

Competing interests

None

Funding

No external funding.

Authors' contributions

The original idea came from ASJ. Writing and construction of the article was shared 50/50 between the authors.

Acknowledgements

Help and support from Dr. Salvatore Cozzolino, Dr. Annal Behav Neurosci, 2(1): 223-236 (2019)

Gianluca Polese and Prof. Anna Di Cosmo, all from the Department of Biology, University of Naples, Federico II, Naples, Italy.

References

1. Jonas JB, Schneider U, Naumann GOH. Count and density of human retinal photoreceptors. *Graefe's Arch Clin Exp Ophthalmol*. 1992;230(6):505-510. Doi: <https://doi.org/10.1007/BF00181769>
2. Kolb H. Facts and figures concerning the human retina. In: Kolb H, Fernandez E, Nelson R, eds. *Webvision: The Organization of the Retina and Visual System*. Salt Lake City (UT): University of Utah Health Sciences Center; 1995. <http://www.ncbi.nlm.nih.gov/books/NBK11556/>.
3. Johnson AS, Winlow W. Computing action potentials by phase interference in realistic neural networks. *EC Neurology*. 2017;5:123-134.
4. Johnson AS. The coupled action potential pulse (APPulse)-neural network efficiency from a synchronised oscillating lipid pulse Hodgkin Huxley action potential. *EC Neurology*. 2015;2:94-101. https://www.researchgate.net/publication/282861120_The_Coupled_Action_Potential_Pulse_Appulse-Neural_Network_Efficiency_from_a_Synchronised_Oscillating_Lipid_Pulse_Hodgkin_Huxley_Action_Potential
5. Johnson AS, Winlow W. Mysteries of the action potential - From 1952 to infinity and beyond. *Physiology News*. 2018;111(38-41). https://www.researchgate.net/publication/326698288_Mysteries_of_the_Action_Potential_-_From_1952_to_infinity_and_beyond
6. Johnson AS, Winlow W. The soliton and the action potential – primary elements underlying sentience. *Front Physiol*. 2018;9:779. Doi: <https://doi.org/10.3389/fphys.2018.00779>
7. Hodgkin AL, Huxley AF. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol*. 1952;117(4):500-544. Doi: <https://doi.org/10.1113/jphysiol.1952.sp004764>
8. Kheradpisheh SR, Ganjtabesh M, Thorpe SJ, et al. STDP-based spiking deep convolutional neural networks for object recognition. *Neural Networks*. 2018;99:56-67. Doi: <https://doi.org/10.1016/j.neunet.2017.12.005>
9. El Hady A, Machta BB. Mechanical surface waves accompany action potential propagation. *Nat Commun*. 2015;6:6697. Doi: <https://doi.org/10.1038/ncomms7697>

10. Engelbrecht J, Tamm K, Peets T. On mathematical modelling of solitary pulses in cylindrical biomembranes. *Biomech Model Mechanobiol.* 2015;14(1):159-167. Doi: <https://doi.org/10.1007/s10237-014-0596-2>
11. Johnson AS, Winlow W. Shortcomings of current artificial nodal neural network models. *EC Neurol.* 2017;4:198-200. <https://www.econicon.com/ecne/pdf/ECNE-04-0000113.pdf>
12. Guo T, Tsai D, Bai S, et al. Understanding the retina: A review of computational models of the retina from the single cell to the network level. *Crit Rev Biomed Eng.* 2014;42(5):419-436. <https://www.ncbi.nlm.nih.gov/pubmed/25745804>
13. Grbatinić I, Krstonošić B, Marić D, et al. Computational RSM modeling of neuromorphofunctional relations of dentate nuclear neurons and dentatostriate inter-cluster mapping with the dentatostriate neural network reconstruction: RLSR/PCR regression and canonical correlation analysis. *ABNE.* April 2019:168-196. Doi: <https://doi.org/10.18314/abne.v2i1.1674>
14. Ichinose T, Fyk-Kolodziej B, Cohn J. Roles of on cone bipolar cell subtypes in temporal coding in the mouse retina. *J Neurosci.* 2014;34(26):8761-8771. Doi: <https://doi.org/10.1523/JNEUROSCI.3965-13.2014>
15. Grimes WN, Schwartz GW, Rieke F. The synaptic and circuit mechanisms underlying a change in spatial encoding in the retina. *Neuron.* 2014;82(2):460-473. Doi: <https://doi.org/10.1016/j.neuron.2014.02.037>
16. Korenbrot JI. Speed, sensitivity, and stability of the light response in rod and cone photoreceptors: Facts and models. *Prog Retin Eye Res.* 2012;31(5):442-466. Doi: <https://doi.org/10.1016/j.preteyeres.2012.05.002>
17. Dowling JE. LCNs in the vertebrate retina. In: Rakic P, ed. *Local Circuit Neurons.* Vol 13. Neurosci Res Program Bull; 1975:334.
18. Dowling JE. Neurons and networks: An introduction to neuroscience. Belknap Press, Harvard University, Cambridge Mass, USA. 1992:447.
19. Roska B, Nemeth E, Werblin FS. Response to change is facilitated by a three-neuron disinhibitory pathway in the tiger salamander retina. *J Neurosci.* 1998;18(9):3451-3459. Doi: <https://doi.org/10.1523/JNEUROSCI.18-09-03451.1998>
20. Heflin SJ, Cook PB. Narrow and wide field amacrine cells fire action potentials in response to depolarization and light stimulation. *Vis Neurosci.* 2007;24(2):197-206. Doi: <https://doi.org/10.1017/S095252380707040X>
21. Baden T, Berens P, Bethge M, et al. Spikes in mammalian bipolar cells support temporal layering of the inner retina. *Curr Biol.* 2013;23(1):48-52. Doi: <https://doi.org/10.1016/j.cub.2012.11.006>
22. Cui J, Pan ZH. Two types of cone bipolar cells express voltage-gated Na⁺ channels in the rat retina. *Vis Neurosci.* 2008;25(5-6):635-645. Doi: <https://doi.org/10.1017/S0952523808080851>
23. Tsukamoto Y, Omi N. Classification of mouse retinal bipolar cells: Type-specific connectivity with special reference to rod-driven All amacrine pathways. *Front Neuroanat.* 2017;11:92. Doi: <https://doi.org/10.3389/fnana.2017.00092>
24. Gisiger T, Boukadoum M. Mechanisms gating the flow of information in the cortex: What they might look like and what their uses may be. *Front Comput Neurosci.* 2011;5. Doi: <https://doi.org/10.3389/fncom.2011.00001>
25. Behrens C, Schubert T, Haverkamp S, et al. Types basal cell ganglia length. Connectivity map of bipolar cells and photoreceptors in the mouse retina. *eLife.* 2016;5:e20041. Doi: <https://doi.org/10.7554/eLife.20041>
26. Strettoi E, Novelli E, Mazzoni F, et al. Complexity of retinal cone bipolar cells. *Prog Retin Eye Res.* 2010;29(4):272-283. Doi: <https://doi.org/10.1016/j.preteyeres.2010.03.005>
27. Cao LH, Luo DG, Yau KW. Light responses of primate and other mammalian cones. *Proc Natl Acad Sci U S A.* 2014;111(7):2752-2757. Doi: <https://doi.org/10.1073/pnas.1400268111>
28. Euler T, Haverkamp S, Schubert T, et al. Retinal bipolar cells: Elementary building blocks of vision. *Nat Rev Neurosci.* 2014;15(8):507-519. <https://www.ncbi.nlm.nih.gov/pubmed/25158357>
29. Dacey D, Packer OS, Diller L, et al. Center surround receptive field structure of cone bipolar cells in primate retina. *Vision Res.* 2000;40(14):1801-1811. Doi: [https://doi.org/10.1016/S0042-6989\(00\)00039-0](https://doi.org/10.1016/S0042-6989(00)00039-0)
30. Chen M, Lee S, Park SJH, et al. Receptive field properties of bipolar cell axon terminals in direction-selective sublaminae of the mouse retina. *J Neurophysiol.* 2014;112(8):1950-1962. Doi: <https://doi.org/10.1152/jn.00283.2014>
31. Tsukamoto Y, Omi N. OFF bipolar cells in macaque retina: Type-specific connectivity in the outer and inner synaptic layers. *Front Neuroanat.* 2015;9:122. Doi: <https://doi.org/10.3389/fnana.2015.00122>

32. Arai I, Tanaka M, Tachibana M. Active roles of electrically coupled bipolar cell network in the adult retina. *J Neurosci*. 2010;30(27):9260-9270. Doi: <https://doi.org/10.1523/JNEUROSCI.1590-10.2010>
33. Winlow W, Qazzaz M, Johnson A. Bridging the Gap – The Ubiquity and Plasticity of Electrical Synapses. *EC Neurology*. 2017;7:7-12. <https://pdfs.semanticscholar.org/a92/3a311e0bf14004771e41f823e4610b90ecaf.pdf>
34. Odermatt B, Nikolaev A, Lagnado L. Encoding of luminance and contrast by linear and nonlinear synapses in the retina. *Neuron*. 2012;73(4):758-773. Doi: <https://doi.org/10.1016/j.neuron.2011.12.023>
35. Feynman RP. The theory of positrons. *Phys Rev*. 1949;76(6):749-759. Doi: <https://doi.org/10.1103/PhysRev.76.749>
36. Larsson ML. Binocular vision, the optic chiasm, and their associations with vertebrate motor behavior. *Front Ecol Evol*. 2015;3. Doi: <https://doi.org/10.3389/fevo.2015.00089>
37. Rosa JM, Morrie RD, Baertsch HC, et al. Contributions of rod and cone pathways to retinal direction selectivity through development. *J Neurosci*. 2016;36(37):9683-9695. Doi: <https://doi.org/10.1523/JNEUROSCI.3824-15.2016>
38. Chaya T, Matsumoto A, Sugita Y, et al. Versatile functional roles of horizontal cells in the retinal circuit. *Sci Rep*. 2017;7(1):5540. Doi: <https://doi.org/10.1038/s41598-017-05543-2>
39. Demb JB, Singer JH. Intrinsic properties and functional circuitry of the All amacrine cell. *Vis Neurosci*. 2012;29(1):51-60. Doi: <https://doi.org/10.1017/S0952523811000368>
40. Shlens J, Field GD, Gauthier JL, et al. The structure of large-scale synchronized firing in primate retina. *J Neurosci*. 2009;29(15):5022-5031. Doi: <https://doi.org/10.1523/JNEUROSCI.5187-08.2009>
41. Protopapa F, Hayashi MJ, Kulashekhar S, et al. Chronotopic maps in human supplementary motor area. Zatorre R, ed. *PLoS Biol*. 2019;17(3):e3000026. Doi: <https://doi.org/10.1371/journal.pbio.3000026>
42. Nisbet-Jones PBR, Dilley J, Holleczek A, et al. Photonic qubits, qutrits and ququads accurately prepared and delivered on demand. *New J Phys*. 2013;15(5):053007. Doi: <https://doi.org/10.1088/1367-2630/15/5/053007>

Copyright: © **Johnson et al.** This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.