Predictive Coding Produces Alpha-band Rhythmic Travelling Waves

Andrea Alamia (andrea.alamia@cnrs.fr)
CerCo, CNRS, Université de Toulouse
Toulouse, 31055 (France)

Rufin VanRullen (rufin.vanrullen@cnrs.fr)
CerCo, CNRS, Université de Toulouse
Toulouse, 31055 (France)

Abstract:

Predictive coding is a key mechanism to understand the computational processes underlying brain functioning: in a hierarchical network, higher layers predict the activity of lower layers, and the unexplained residuals (i.e. prediction errors) are sent through. At CCN'17 we showed that a simple 2-layers predictive coding model of visual cortex, with physiological communication delays between layers, gives rise to alpha-band rhythms (8-12Hz), similar to experimental observations. We now demonstrate that a multi-layer version of the same model can explain the occurrence of oscillatory travelling waves across layers, both feedforward (during visual stimulation) and backward (during rest). Remarkably, the predictions of our model are matched by the analysis of two independent EEG datasets, in which we observed oscillatory travelling waves. In conclusion, we suggest that predictive coding could be the origin of important features of neuronal activity, such as alpha rhythms and travelling waves.

Keywords: predictive coding, alpha rhythms, travelling waves, impulse-response function, neural delays, time constants

Introduction

Predictive coding is a popular computational paradigm to model sensory information processing in the brain, and it has been proposed to explain several cognitive and physiological observations (Rao & Ballard 1999). Could it also explain the emergence of alpha-band oscillations? Alpha rhythms (8-12Hz) are the most predominant oscillations in the human brain, even though their functional role remains debated. On the one hand alpha activity has been correlated with lack of visual input (i.e. closed eyes); on the other hand, alpha has been suggested to contribute to the temporal framing of sensory inputs. Importantly, a previous study (VanRullen & Macdonald, 2012) revealed a long lasting (up to ~1 s) alpha-band oscillation in the cross-correlation between EEG signal and a white-noise visual input (i.e. Impulse Response Function - IRF), suggesting a potential role in maintaining sensory information over time (fig.1A). Last year at CCN we proposed a simple 2-layers predictive coding model of visual cortex which produces alpha-band oscillations, due to the presence of physiologically plausible constraints, such as membrane time constants and communication delays between brain regions (fig 1B). Here, we address another neuronal phenomenon called travelling waves.

A recent study from our group (VanRullen & Lozano-Soldevilla 2017) showed that alpha IRF oscillations propagate across the cortex in an occipital-to-frontal direction (fig.2A). This finding is in line with other recent studies about cortical travelling waves (Muller et al. 2018, Zhang et al. 2017). We expand our predictive coding model by increasing the number of layers to investigate the occurrence of alpha band travelling waves in a hierarchical network. After defining a method to quantify travelling waves in EEG signals (fig. 2B-C), we observed that our model predictions were qualitatively similar to experimental data about travelling waves.
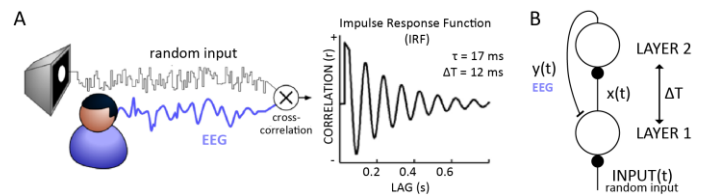

Figure 1 – Alpha-band oscillations in human IRF and in predictive coding models. A) Cross-correlating the white noise sequence of a stimulus with simultaneously recorded EEG produces an Impulse Response Function (IRF) which reverberates at 10Hz. B) A simple predictive coding model in which layer 2 makes predictions about the input received by layer 1, and the residual (prediction error) is used to update the next temporal prediction. Such a model, with physiologically plausible parameters, generates an oscillatory IRF at 10 Hz. The oscillatory IRF in A is produced by the model, with communication delay ΔT=12ms, and neural membrane time constant τ=17ms (see Figure 2B for examples of real human EEG oscillatory IRFs).
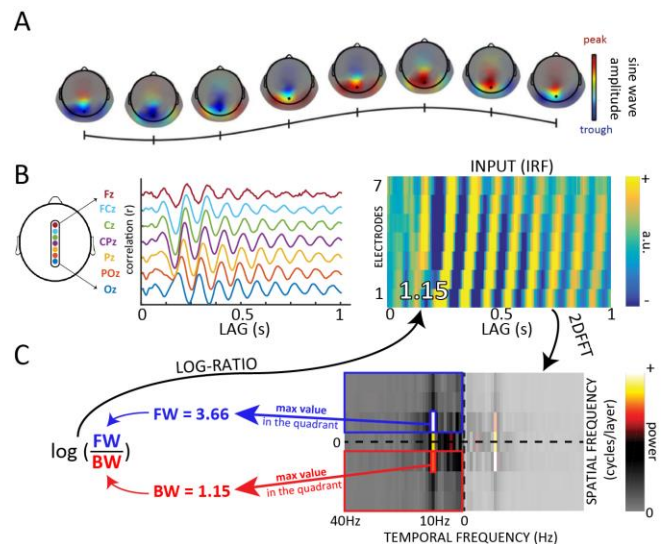

Figure 2 – Alpha-band travelling waves in EEG Impulse Response Functions. A) One average 100ms cycle of the IRF, with a scalp topography

plotted every 12.5ms. The oscillatory peaks (and troughs) propagate in a forward direction. B-C) IRF as travelling waves can be observed over the 7 midline electrodes of the 10-20 system (Oz to Fz): from the 2D-map obtained by stacking signals from the 7 electrodes, we compute a 2D-FFT and derive a log-ratio between spectral quadrants that quantifies the direction of the waves. Positive values are associated with feedforward waves (FW), and negative values with backward ones (BW).

Methods & Results

The model

Predictive coding postulates a hierarchical architecture in which higher layers predict the activity of lower layers, and the residuals (i.e the difference between the prediction and the actual activity) are carried over to update the next prediction. Our model is a greatly simplified version of the standard predictive coding circuit (Rao & Ballard, 1999), because our primary aim is to understand the temporal dynamics of brain activity in response to a simple time-varying stimulus (Figure 1). Therefore, each layer of the model contains a single neural population, with no "receptive fields" or spatial selectivity. In our model, illustrated in Fig 3A, the residual is defined as:

$$x_L(t) = y_{L-1}(t) - y_L(t-\Delta T) \tag{1}$$

where L indexes the layers, and ΔT represents the temporal delay between them. Moreover, $y_{L-1}(t) = INPUT(t)$ when L=1. The prediction y_L, as shown in equation 2, is updated based on the bottom-up residual x_L (with a delay), and on the difference between its prediction and the prediction from the next higher layer, which can be considered as a top-down prior:

$$\frac{dy_L}{dt} = \frac{1}{\tau} \cdot x_L(t-\Delta T) + \frac{1}{\tau_D} \cdot (y_{L+1}(t-\Delta T)-y_L(t)) \tag{2}$$

At the last layer, the prior y_{L+1} represents a generative endogenous process, arising from higher-level brain regions that are not part of the model; in our simulations, this prior could be imposed as a time-varying signal or set to 0 (see fig. 3A). In order to facilitate measurements of cross-correlations (IRF) with this prior and comparison with the stimulus IRF, when different than 0 we defined it as a white-noise time-varying signal with statistics similar to those of the input signals.

Besides ΔT, two other parameters play a crucial role in the model: τ and τ_D which describe the temporal dynamics (time constants) of neuronal integration and decay, weighting respectively the residual computed from the lower-layer, and the prediction from the higher layer. In all simulations, the model was composed of 7 layers, and both equations were solved numerically with a 1ms time step.

Simulations

In all simulations we tested a 7-layers model, in which τ_D (decay time constant) was fixed at 200ms, and ΔT and τ were respectively 12ms and 17ms. These parameter values are

biologically plausible, and were found to produce IRF alpha oscillations in previous simulations (VanRullen, CCN 2017; Figure 1). We simulated two conditions, in which either a white-noise input or a white-noise prior were presented to the first or last layer, respectively. When the input was present the prior was set to zero, and vice-versa. These design choices were made to facilitate comparison of input-driven and prior-driven IRFs, and do not reflect any assumption about the statistical structure of inputs or priors under natural conditions of stimulation. We treated prediction signals of each layer as the equivalent of EEG signals from distinct electrodes in our model. For each condition we computed the cross-correlation (IRF) between the EEG at each layer and either the input or the prior. We created 2D maps by stacking the EEG or IRF signals (x-axis) from the 7 layers (y-axis, see fig. 2B). To quantify the presence of waves, we extracted the maximum values in the upper and lower quadrants of the 2D-FFT of these maps, representing respectively the amount of feedforward (FW) and backward (BW) signal propagation across layers (fig. 2 B-C). Finally, the log-ratio of these two values quantified the overall direction of the waves: positive ratios indicate predominantly FW waves, whereas negative ratios reveal mostly BW waves.

The chief conclusion of these simulations is that hierarchical predictive coding gives rise to oscillatory travelling waves: as shown in figure 3B, sensory inputs generate FW waves in both the EEG and IRF signals, whereas top-down priors induce BW waves.
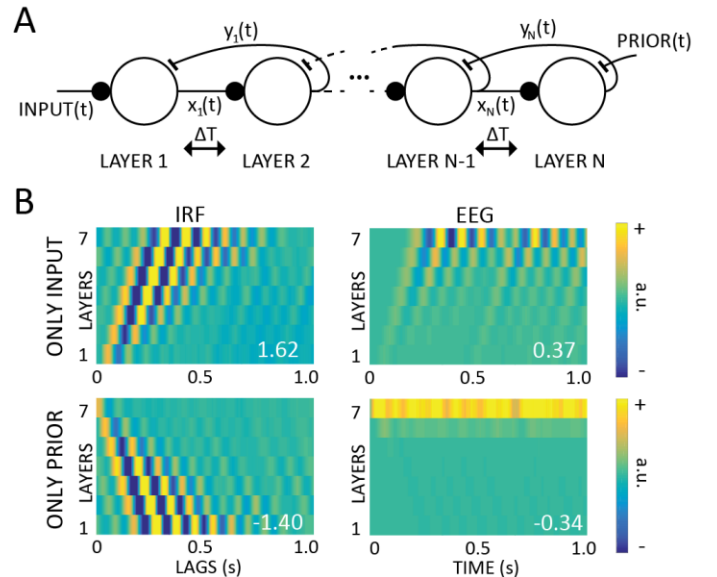

Figure 3 – Alpha-band traveling waves in a hierarchical predictive coding model. A) Multi-layer version of the model: the same parameters ΔT and τ are used throughout. The model is fed either a time-varying input (left) or a time-varying prior (right). B) 2D maps with only input (top row) or prior (bottom row): travelling waves are visible both in the IRF (left panels) and in the predictions (considered as a proxy for the EEG, right panels). The numbers in white show the log-ratio for each simulation: positive or negative numbers reveal respectively FW or BW waves.

Electrophysiological data

Finally, we turned to real human EEG data to investigate whether we could observe travelling waves as predicted by our model. We focused on 2 datasets: in the first one (INPUT), 20 participants fixated a white noise sequence of random luminance values; in the second (CLOSED EYES), 48 participants underwent a 1-minute recording with closed eyes. For both datasets, we obtained the 2D map and computed the log-ratio, as for the model, over seven midline electrodes (Oz to Fz), for a sliding window of 1 second (500ms overlap). For the INPUT dataset, we also quantified the log-ratio using the IRF (cross-correlation of visual input and EEG). We computed a null distribution of log-ratios by shuffling the electrodes' order 1000 times per subject. The difference between the distribution of the real log-ratios and this null distribution provides a statistical estimation of the proportion of reliable FW and BW wave events in each dataset. For comparison, we applied the same procedure for the model's simulations, shuffling the layers in place of the electrodes.

Concerning the results of the experimental data, as shown in figure 4A-B, more BW waves are observed when participants have closed eyes, whereas during visual stimulation both the EEG and the IRF reveal a larger amount of FW waves. Figure 4 (panels C,D) shows the same analysis applied to the model simulations: notably, in this case we can compute an IRF both with the input and with the top-down prior (for human experiments we only have access to the visual input, but the internal prior, if any, remains unknown). As in the closed eyes condition, when only the prior is present, mostly BW waves are observed, whereas when only the input is provided, we observed mostly FW waves in both the EEG and IRF signals, consistent with the results of the INPUT dataset.
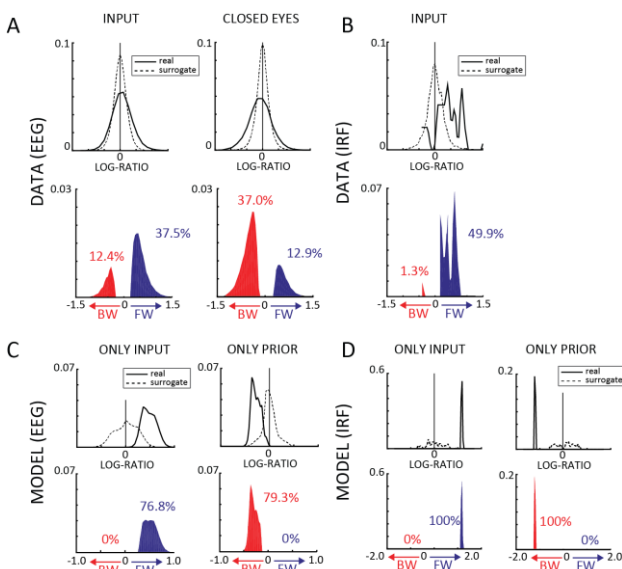
line) distributions of log-ratios computed over the 2 EEG datasets. The second row shows the difference between the two, focusing on positive values, i.e. the events in the real EEG data that occur more often than predicted by the null distribution. The proportion of significant backward and forward waves are shown in red and blue, respectively. B) Same as A, but for the IRF computed in the INPUT dataset. C,D) Same as A,B) regarding the results of the simulations. In the model we could compute the IRF in two ways, either by cross-correlating predictions with the sensory input, or with the top-down prior. Experimental EEG results (A,B) follow the same pattern as the simulations (C,D).

Discussion

Our model, despite its relatively simple design, reveals that alpha-band oscillations propagate as travelling waves in a hierarchical system based on predictive coding principles. This result is compatible with our own EEG data (fig.5 provides a summary of the results), as well as observations from other experimental studies (Zhang, Watrous, Patel, & Jacobs, 2017; Halgren et al., 2017).
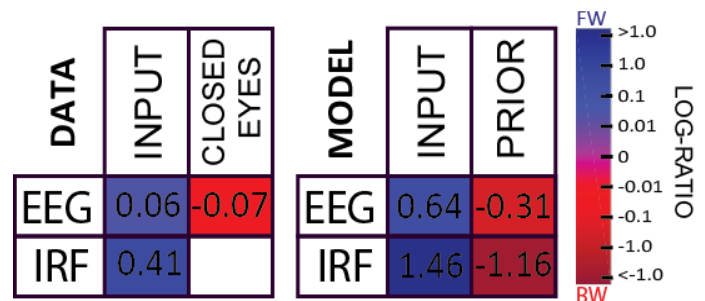
Figure 5 – Summary of the results. The tables summarize the results of the data (left) and simulation (right) for each dataset/condition. Red and blue represent respectively a higher number of BW and FW waves.

Much work has been carried out regarding travelling waves, both in humans (mostly but not exclusively intra-cortical recordings – Muller et al., 2018) and in animals (e.g. turtles, Nenadic et al., 2003). Specifically, three different types of travelling waves have been characterized in the literature. First, scalp ERPs propagating from posterior to anterior regions in response to stimulus onset have been interpreted as non-periodic travelling waves (Klimesch, Hanslmayr, Sauseng, Gruber, & Doppelmayr, 2007). Second, periodic sensory stimulations produce oscillatory travelling waves that share the same frequency as the generating stimulus (Sato, Nauhaus, & Carandini, 2012): even though they behave rhythmically, these waves are similar in principle to the first type above, as their oscillatory nature is related to the frequency characteristics of the generating stimulus. Lastly, truly periodic travelling waves have their own intrinsic frequency (e.g. within the alpha-band range, VanRullen & Lozano-Soldevilla, 2017), unrelated to the frequency content of the stimulus. In this study we investigate this last type of travelling waves.

The travelling direction of the waves is important in determining their functional role. In a recent study (Halgren et al., 2017), intra-cortical recordings from epilepsy patients during quiet wakefulness revealed alpha-band travelling waves propagating "backward" from higher-order antero-

Figure 4 – Quantification of traveling waves in human data and model simulations. A) The first row shows the real (solid line) and shuffled (dashed

superior cortex to lower-order occipital poles. Analyzing recordings simultaneously from cortex and the pulvinar, Halgren and colleagues concluded that alpha-band oscillations 1) originate in the cortex, and 2) reflect feedback processing between cortical regions. The predictions of our model are in agreement with both conclusions. Furthermore, another recent human intracranial study from Zhang and colleagues (Zhang, Watrous, Patel, & Jacobs, 2017) reported theta-alpha travelling waves (2 to 15 Hz) propagating from posterior to anterior brain regions during a visual memory task. The conclusion that feedforward travelling waves are related to stimulus processing is also in agreement with our predictions. All in all, our model reconciles both studies, positing that alpha-band travelling waves emerge as the result of layer-by-layer interactions in a hierarchical system, and their direction is related to their functional role.

Due to its simplified architecture, our predictive coding model produces a single oscillation whose frequency depends on the chosen parameters: typically in the alpha band for biologically plausible values (Figure 1). The same alpha oscillation carries top-down predictions down the hierarchy, and bottom-up prediction residuals up the hierarchy, resulting in two alpha travelling waves moving in opposite directions. As explained above, this is compatible with recent experimental reports of both feedforward and backward alpha travelling waves in human intracranial studies. However, there is also a growing number of studies reporting that faster gamma oscillations (~30-80Hz) are specifically involved in feed-forward signal transmission, while alpha- and beta-band (13-30Hz) rhythms convey top-down information (Buffalo et al, 2011; van Kerkoerle et al, 2014; Bastos et al, 2015; Michalareas et al, 2016) . These dynamics can be appropriately captured by a more detailed predictive coding model, the "canonical cortical microcircuit" model (Bastos et al, 2012; Friston, 2005), which includes different types of excitatory and inhibitory neurons, as well as detailed laminar circuitry in each brain region. An important next step would thus be to explore the existence, frequency and direction of travelling waves in a hierarchical version of the canonical microcircuit model. Finally, future versions of our model could also expand the number of neurons in each layer, together with a retinotopic organization and spatially selective receptive fields (as in Rao & Ballard, 1999), in order to process spatially as well as temporally structured inputs. This should provide a more complete understanding of predictive coding and oscillatory travelling waves in relation to essential visual functions such as object recognition or categorization.

Acknowledgments

References

Bastos, A. M., Usrey, W. M., Adams, R. a, Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, 76(4), 695–711

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J. M., Oostenveld, R., Dowdall, J. R., Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron*, 85(2), 390-401.

Buffalo, E. a, Fries, P., Landman, R., Buschman, T. J. & Desimone, R. (2011). Laminar differences in gamma and alpha coherence in the ventral stream. *Proc. Natl. Acad. Sci. USA* 108, 11262–11267.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences, 360*(1456), 815–836.

Halgren, M., Devinsky, O., Doyle, W. K., Bastuji, H., Rey, M., Mak-McCully, R. Cash, S. S. (2017). The Generation and Propagation of the Human Alpha Rhythm. *bioRxiv*, 202564.

Kilmesch,W. Hanslmayr,S., Sauseng,P., Gruber, W., Doppelmayr,M., (2007) P1 and travelling alpha waves: evidence for evoked oscillations. *J Neuroph, 97*, 1311-1318.

Michalareas, G., Vezoli, J., van Pelt, S., Schoffelen, J.-M., Kennedy, H., & Fries, P. (2016). Alpha-beta and gamma rhythms subserve feedback and feedforward influences among human visual cortical areas. *Neuron*, 89(2), 384–397.

Muller,L., Chavane,F., Reynolds,J., Sejnowski,J., (2018) Cortical travelling waves: mechianisms and computational principles. *Nature Neuroscience Review*19, 255-268

Nenandic, Z., Ghosh, B., Ulinski, P. (2003) Propagating waves in visual cortex: a large scale model of turtle visual cortex *J.Comp.Neur.* 14 161-184

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci, 2*(1), 79-87

Sato,T.K., Nauhaus,I.,Carandini,M., (2012) Traveling waves in visual cortex. *Neuron* 75, 218-229

van Kerkoerle, T. et al. Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proc. Natl. Acad. Sci. USA* 111, 14332–14341 (2014).

VanRullen, R., Lozano-Soldevilla, D.. (2017). The hidden spatial dimension of alpha: 10 Hz perceptual echoes propagate as periodic travelling waves in the human brain. *bioRxiv,* 190595

VanRullen, R., & Macdonald, J. S. P. (2012). Perceptual echoes at 10 Hz in the human brain. *Current Biology, 22*(11), 995-999.

VanRullen, R. (2017). Predictive coding & neural communication delays produce alpha-band oscillatory impulse response functions. *Cognitive Computational Neuroscience (CCN) 2017*

Zhang H., Watrous, A., Patel, A., Jacobs, J. (2017) Theta and alpha oscillations are traveling waves in the human neocortex *bioRxiv,* 218198