Predictive neural fields for improved tracking and attentional properties

Jean-Charles Quinton, Bernard Girau

Predictive neural fields for improved tracking and attentional properties

Jean-Charles Quinton and Bernard Girau

***Abstract*— Predictive capabilities are added to the competition mechanism known as the Continuum Neural Field Theory, in order to improve and extend its attentional properties. In order to respect the distributed and bio-inspired nature of the model, the prediction is introduced as an internal stimulation, directly determined by the past field activity.**
Building on mathematical developments and optimization techniques, performance is ascertained on a 2D tracking application where the system must robustly focus on a target despite rapid movements, noise and distracters. In addition to a consistent gain on previously observed capabilities, the extended model can also track stimuli with full occlusions and static obstacles on the trajectory.

I. Introduction

The Continuum Neural Field Theory (CNFT) is a particular kind of dynamic neural field model that implements lateral competition within cortical maps [2], [18]. These models rely on the hypothesis that the generic structure and topological organization of the cortex are relevant and maybe necessary to explain a wide range of cerebral phenomena. Following the classical results on the laminar structure of the cortex, the correlated activities found across cortical layers [11] and the multiscale organization deduced from cytoarchitectural and functional differences [3], [9], several computational neuroscience models have been proposed in this context [4].

The CNFT can be reduced to a single differential equation that describes the evolution of the membrane potential of neurons over entire cortical maps. The simplified version used in this paper is given by the following single-layer field equation of lateral inhibition type ($\forall \mathbf{x} \in M$):

$$u^{t+dt}(\mathbf{x}) = f\left(\left(1 - \frac{dt}{\tau}\right) u^t(\mathbf{x}) + \frac{dt}{\tau}\left(c(\mathbf{x}) + i(\mathbf{x})\right)\right) \quad (1)$$

where f is chosen such as to maintain the membrane potential u in $[0, +\infty)$. An input signal i is provided to the field and the competition term c is defined as a continuous convolution with a lateral connection kernel w by:

$$\begin{aligned} c(\mathbf{x}) &= \int_{\mathbf{x}' \in M} w(\mathbf{x} - \mathbf{x}') u^t(\mathbf{x}') \mathbf{dx}' \\ w(\mathbf{dx}) &= Ag(\mathbf{dx}, a) - Bg(\mathbf{dx}, b) \\ g(\mathbf{dx}, \sigma) &= e^{-\frac{\|\mathbf{dx}\|^2}{\sigma^2}} \end{aligned} \quad (2)$$

Jean-Charles Quinton is with the INRIA/Cortex project and Bernard Girau is with the UHP/LORIA/Cortex project, Campus Scientifique, B.P. 239, 54506 Vandoeuvre-lès-Nancy Cedex, France (email: quintonj@loria.fr).

Under adequate conditions, so-called bubbles of activity emerge on the neural field in response to external stimulation. These bubbles can be defined as spatiotemporally coherent patches of activity that exhibit robust attentional properties, such as focusing on and tracking stimuli despite the presence of noise or distracters [16]. Additionally, the non-linear dynamics of the CNFT accounts for the rapid selection of a single target even when exposed to a set of equally active stimuli.

Competition is made possible by introducing a lateral connectivity pattern following a difference of Gaussians profile. Building on mathematical developments for the conditions of convergence toward bubbles of activity [17], [1] and using genetic algorithms to select optimal parameters for the equation, we showed that there is a tradeoff between accuracy and robustness [13]. Indeed, correctly constraining the connectivity pattern to a Mexican hat profile with local excitation and global inhibition almost guarantees the emergence of an activity bubble, but the speed of the moving stimulus and the amount of noise present in the stimulation either result in a tracking lag or the system loosing its target.

This limitation can be explained by the fact that the competition term and thus the solutions to the equation are fundamentally stationary, which is quite paradoxical when considering applications such as motion tracking or spatiotemporal pattern recognition. Albeit the focus dynamically emerges from a differential equation, it is not made for dynamic inputs, it just adapts to them. As a consequence, the focus bubble resists to movement and prefers noise or distracters reinforcing its activity over any moving stimulus, even when they provide a much weaker stimulation compared to the target. Put in the broader context of cognitive agents, this phenomenon pinpoints the limits of purely reactive bottom-up attentional systems when confronted to situations that offer multiple potentialities. Some kind of "internal model" is needed to discriminate between dynamics, yet it does not need to be very complex or abstract.

When considering applications to sensorimotor systems such as robots, the controller can be modeled by a distributed network of local contingencies, each predicting the outcome of specific actions [14]. Indeed, delays are intrinsic to interactive systems and fast interactions are often required (keeping balance with a relatively slow feedback loop is a paradigmatic example), which implies that the agent should anticipate the dynamics. In that case, the CNFT can be

seen as a competition mechanism acting upon distributed anticipatory representations. Competition is then not only beneficial to perception by contributing to covert attentional mechanisms, but is also required for action selection as all anticipatory representations are simultaneously active. Furthermore, actions participate to overt attention (consider saccadic eye movements for instance), making competition a crucial and recurrent component of the sensorimotor loop [7].

Whatever the perspective taken on the potential place of the CNFT within a cognitive architecture, preferences or predictions thus need to be integrated to reactions in a single coherent model, where they can influence each other. Even though iterating over the CNFT computations or switching to more complex integration schemes might sometimes help to better converge on the stimuli and reduce the tracking lag, this has a high computational cost incompatible with real-time constraints. In this paper, we therefore evaluate the consequences of adding a predictive input signal within the CNFT equation, as the basic equation cannot cope anyway with all the other problems mentioned above.

Despite the broad scope of application the combination of a predictive system and the CNFT model may have, we here only analyze the benefits and potential drawbacks of the introduction of a single predictive term for the CNFT attentional properties (focus, accuracy, robustness). In addition to mathematical developments and qualitative results in one dimension, the performance will be quantitatively assessed on the artificial 2D visual tracking application introduced by Rougier et al. [16].

II. Proposed model

In this paper, we propose an extension of the original equation (Eq. 1). We simply decompose the input signal i as a weighted sum of a bottom-up stimulation s and a top-down prediction p:

$$i(\mathbf{x}) = \alpha p(\mathbf{x}) + (1-\alpha)s(\mathbf{x}) \tag{3}$$

where α is chosen such as to guarantee the primacy of the external stimulation on the inner prediction ($\alpha \in [0, 0.5]$). The predictive term can then be considered as a cortico-cortical modulation to be contrasted with the thalamic input from the senses, a division that is also found in neural field models of multimodal integration [10]. To remain as simple as possible relatively to the competition between the immediate external stimulation and future-oriented inner projections, α is fixed and a single predictor is considered (see Fig. 1 for a graphical representation). Under these conditions, we will determine in the following sections how the predictor should be defined to optimally improve the performance.

In more realistic settings, a variable number of predictors (associated with different potentialities, goals and values)

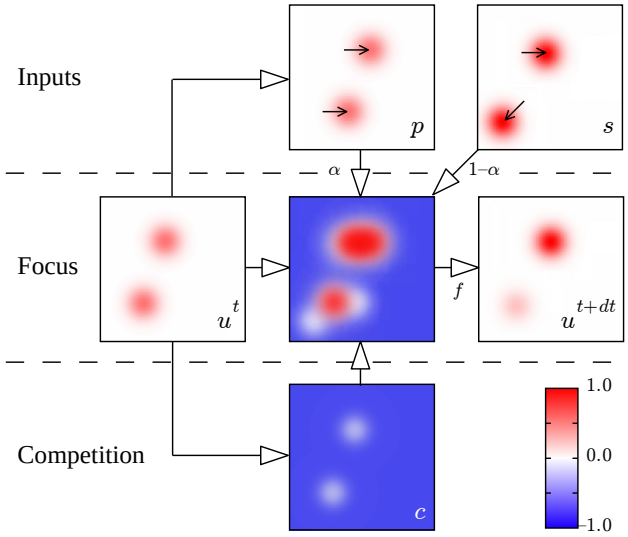

Fig. 1. Graphical illustration of the predictive CNFT dynamics. Although 2 bubbles are competing for the focus on the field u^t, a single stimuli from the stimulation s confirms the rightward translation anticipated by the predictor p (predicted and actual movements are indicated by arrows). Combining the different fields (terms in Eq. 1) and applying the non linear function f gives the new focus activity u^{t+dt}, the inhibitory field induced by the competition eliminating most non coherent activities. Convergence thus occurs on the stimulus that matches the prediction.

as well as adaptive coefficients can be introduced. The corresponding weights should then reflect the adequation of the predictors with the current needs of the agent and dynamics of the environment. The focus here is on the flexibility of the model, and a simple change in Eq. 3 facilitates the introduction of an arbitrary number of predictors that can be evaluated in an online manner (details cannot be given in this short paper).

Keeping the convolution kernel untouched and thus symmetrical makes it furthermore easy to optimize the computations, either by using linear algebra techniques (such as truncated singular value decomposition for matrix approximation) or shifting to continuous approximations of the entire field activity (Gaussian mixture models for instance) [15]. Instead of considering a linear combination of predictors, it is however possible to directly learn asymmetrical kernels that bias the competition toward predicted states as in [5]. Although this latter approach may provide an increased expressive power and the system may assimilate faster dynamics, it is also harder to evaluate the kernels correctness online. Moreover, it needs the system to learn complete and well-separated trajectories whereas our predictive neural fields simply handle local predictions and are able to cope with close trajectories.

The predictor p can be defined as a transformation of u^t in the geometric sense, using the past activity over the entire field to predict the movement of the bubble. As local linear movements can be combined to produce arbitrary tra-

jectories (as Locally Weighted Regression can approximate any function [6]), we will only consider linear predictors in presence of a single bubble. From now on, $\mathbf{v}$ represents the instantaneous speed vector of the tracked stimulus, so that a bubble centered on $\mathbf{x}$ at time t should move to location $\mathbf{x} + \mathbf{v}dt$ at time $t + dt$. This does not restrict the generality of the approach, but allows us to directly express $p(\mathbf{x})$ as:

$$p^t(\mathbf{x}) = u^t(\mathbf{x} - \gamma(\mathbf{v})dt) \tag{4}$$

Taking γ equal to the identity seems natural as it means that the prediction, if correct, will be aligned with the stimulus. It thus introduces an asymmetry between the stimulus, that benefits from both the activity of s and p, and distracters that may be equally active in the external signal but do not receive a positive influence from p. This bias in the amplitude is demonstrated on Fig. 2.

Nevertheless, due to the inertia of the CNFT equation (as the inputs get integrated with a ratio of $\frac{dt}{\tau}$), an infinite number of iterations would still be required to converge on the moving stimulus. Moreover, the stationarity and non-linear dynamics of the CNFT results in oscillations of the bubble. Indeed, it tends to stabilize at the target location, with its amplitude increasing until the stimulus and competition inputs get compensated by the leaking term. As the target moves away from the bubble, the latter will then be destabilized as the stimulation enters the inhibitory part of the lateral kernel. The bubble amplitude will decrease, allowing it to move back to the target position (see Fig. 3).

As a consequence, taking the identity for γ leads to an unbreachable lag behind moving stimuli. To quantitatively evaluate the error made in the tracking of the target, we compute the distance between the center of the stimulus $\mathbf{x}_s$ and the center of mass of the focus activity u. The resulting functional d, that depends on the γ function, is defined in Eq. 5. If convergence occurred and a single bubble emerged, it should be symmetrical and centered on the stimulus so that the distance should be null. As an aside, $\mathbf{x}_s$ is only known to the designer for computer generated inputs, where input

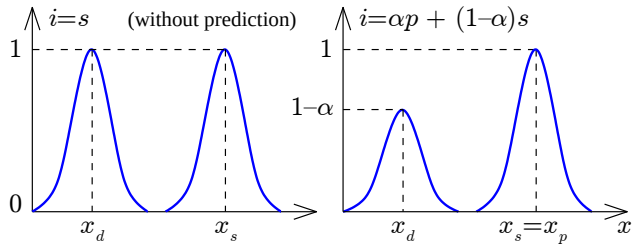

Fig. 2. Comparison of the input activity i with or without prediction. *(left)* Without prediction, a stimuli and a distracter of equal intensity are not differentiated by the CNFT. Any of them may thus act as an attractor in Eq. 1, even when the convergence already occurred but the distracter is located on the trajectory (x_d). *(right)* If the prediction matches the actual stimulus movement ($x_p = x_s$) and not the distracter, the symmetry is broken and the selection or tracking is biased.

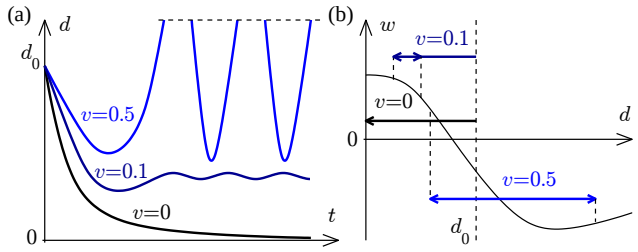

Fig. 3. *(a)* Typical convergence and tracking error curves for 3 different speeds for $\gamma(\mathbf{v}) = \mathbf{v}$. Although produced by simulation, the results have been smoothed out for a qualitative analysis of the dynamics and axis scales are thus irrelevant. Starting from a distance d_0 between the target and bubble, the tracking accuracy deteriorates as the speed increases. Oscillations described in the main text can be observed, as the bubble moves closer and away from the target. *(b)* The range of the oscillations are reproduced, and superimposed on the integrated lateral kernel function w. Oscillations result from an alternation between positive (increasing the amplitude) and negative weights (facilitating the movement).

scenarios are carefully specified.

$$d(\gamma) = \left\| \mathbf{x}_s - \frac{\int_{\mathbf{x}\in M} u(\mathbf{x})\mathbf{x}\mathbf{d}\mathbf{x}}{\int_{\mathbf{x}\in M} u(\mathbf{x})\mathbf{d}\mathbf{x}} \right\| \tag{5}$$

Given the parameters of the CNFT equation and for a single iteration of the integration scheme, we therefore try to find γ that minimizes the tracking error whatever the speed of the stimulus to follow. The following section aims at deriving such an optimal γ function by analyzing the reciprocal interactions between the stimulus and both the focus and prediction bubbles of activity.

III. Optimal prediction

According to preliminary experiments, it has appeared that γ needs to overestimate the position of the bubble so as to compensate for the intrinsic latency of the equation. Nevertheless, these experiments have shown that the optimal γ function is not constant, nor easy to find. However, due to the slow variations of the optimal value relatively to the speed of the stimulus to follow, it is possible to improve the attentional properties of the CNFT using the same parameter value for a wide range of input trajectories. We therefore propose here to analytically find such value based on the whole set of parameters.

A. Analysis of interactions

To give a rough idea of the mathematical resolution of this problem, we can get inspiration from the Gaussian mixture model introduced in [15]. Let $g_s = (\mathbf{x}_s, S)$, $g_u = (\mathbf{x}_u, U)$ and $g_p = (\mathbf{x}_p, P)$ respectively be the Gaussian components approximating the stimulus, focus and predicted bubbles. Each couple corresponds to the center and amplitude of the associated bump of activity so that $s(\mathbf{x}) \approx S \times g(\mathbf{x} - \mathbf{x}_s)$ for instance. Starting from a state where the focus bubble was perfectly aligned with the stimulus at $\mathbf{x} = \mathbf{0}$ implies that $\mathbf{x}_u = \mathbf{0}$, $\mathbf{x}_s = \mathbf{v}dt$ and $\mathbf{x}_p = \gamma(\mathbf{v})dt$. The prediction being

reduced to a translation, the predicted and focus bubbles have the same amplitude $P = U$ (see Fig. 4). By translating Eq. 1 and Eq. 3 into interactions between the various components as in [15], we can update their amplitude as follows:

$$\begin{aligned} g'_u &= (\mathbf{0}, U') & U' &= (1 + \frac{dt}{\tau}(-1 + w(0)))U \\ g'_s &= (\mathbf{v}dt, S') & S' &= \frac{dt}{\tau}((1-\alpha)S + w(\mathbf{v}dt)U) \\ g'_p &= (\gamma(\mathbf{v})dt, P') & P' &= \frac{dt}{\tau}(\alpha + w(\gamma(\mathbf{v})dt))U \end{aligned} \quad (6)$$

The new bubble position can then be computed as the barycenter of the components, where the weights correspond to the amplitudes. The optimal γ function is then obtained for:

$$d(\gamma) = 0 \Leftrightarrow \frac{U'\mathbf{x}_u + S'\mathbf{x}_s + P'\mathbf{x}_p}{U' + S' + P'} = \mathbf{x}_s \quad (7)$$

As the components considerably overlap due to the small timestep required for the competition to work, this approximation only captures the essence of the interactions that determine the efficiency of the prediction on tracking performance. This component based model should therefore not be used for precisely choosing the optimal γ function. Yet, an accurate albeit more complex analytical resolution based on definite integrals is possible.

B. Analytical resolution

To find the optimal γ analytically, we only consider a manifold $M = (-\infty, +\infty)$, take $f(u) = max(u, 0)$ in Eq. 1 and provide a single Gaussian stimulation to the system:

$$s^t(x) = Sg(x - vt, c) \quad (8)$$

We also approximate the localized bubble of activity on which this equation converges as the following Gaussian function:

$$u^t(x) = Ug(x - x_t, \sigma) \quad (9)$$

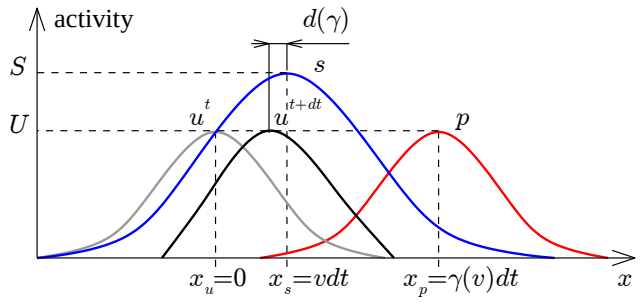

Fig. 4. Graphical illustration of the input, prediction and focus in 1D (distances between bumps on the x-axis are not to scale). The focus at time t (*gray*) is translated from position 0 by $\gamma(v)dt$ to produce the prediction (*red*). The stimulation moves at speed v, which makes its position go from 0 (previous) to vdt for this iteration (*blue*). By updating the CNFT equation, the focus activity u^{t+dt} is produced (*black*). For the tracking to be optimal, γ must be chosen such as to minimize the error distance $d(\gamma)$.

where x_t corresponds to the center of the bubble distribution at time t (parameters U and σ are dependent variables of the equation parameters and input stimulation). For simplification purpose, we suppose that the position of the bubble at time t is 0 and that the stimulus moved at position vdt (see Fig. 4). The instantaneous prediction p can then be defined by:

$$p^t(x) = Ug(x - \gamma(v)dt, \sigma) \quad (10)$$

Starting from Eq. 5 when $d(\gamma) = 0$ and $x_s = vdt$ then transforming it to eliminate the divisor, we more generally study the following function:

$$e_I(\gamma) = \int_{x\in I} u^{t+dt}(x)xdx - vdt \int_{x\in I} u^{t+dt}(x)dx \quad (11)$$

This function is ideally equal to 0 when the focus moves to the same position as the stimulus.

If the function f was absent from the equation, the integration interval I could be taken equal to $(-\infty, +\infty)$ and computations would be drastically reduced. However, this leads to a simple expression $(\gamma(v) = \frac{v}{\alpha}\left(\left(\frac{\tau}{dt} - 1\right) + \sqrt{\pi}(aA - bB)\right) + 1)$ that does not capture accurately the relationships between the parameters. The reason is that the integral of $c(x)$ is negative and far greater in absolute value than all the other components. We thus need to restrict the integration interval to $[vdt - \theta, vdt + \theta]$, with θ chosen as the minimal value to satisfy the relation $u(vdt + \theta) = 0$. The interval is symmetric around vdt, which corresponds to the ideal value of the center of the bubble at $t + dt$, and this is equivalent to only taking into account the positive values of $u(x)$. Of course θ depends on the width of the gaussian σ, that can be chosen to minimize the difference between the integral of the Gaussian and actual bubbles, which is $\left|\int_{x\in I}(u(x) - g(x, \sigma))dx\right|$.

Having thus reduced the integration interval, a factorized expression of $e_I(\gamma)$ can be found. Mathematical details are given in the appendix. The final expression does not get really simpler by differentiating it, and series developments are required for a direct approximation of the function from which we can derive the optimal γ. For example with functional parameter values $(A = 20, B = 15, a = 0.1, b = 1, dt = 0.1, \tau = 0.3, \alpha = 0.5)$ and $v = 0.1$, we obtain an optimal value $\gamma(v) = 8.04$ with $\sigma = \theta = 0.07$, which is a good approximation of the value obtained experimentally $(\gamma(v) \simeq 8)$.

C. Interpretation

This difference in the location of the anticipatory feedback $(\gamma(\mathbf{v})dt)$ and the predicted location of both the stimulus and bubble $(\mathbf{v}dt)$ can be interpreted in terms of action and perception. The injected activity p is the action performed on the CNFT by the predictor, which is partially decoupled from the actual movement observed in u. This is very similar to inner actions observed between cortical areas (the prefrontal cortex being considered for instance as part

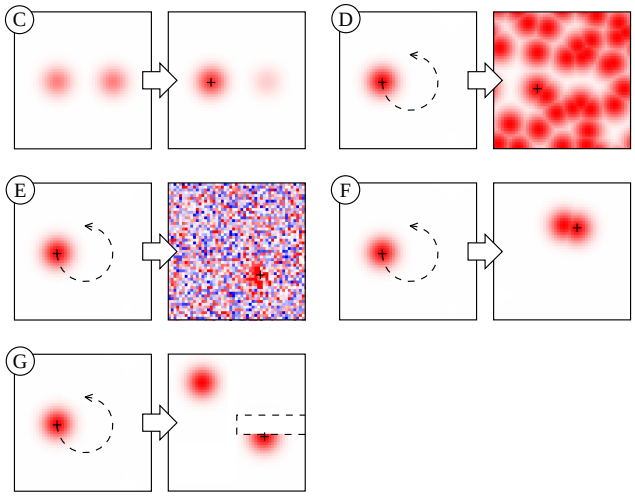

Fig. 5. Illustration of scenarios C to G (please refer to the main text for their exact description). Each scenario is represented by 2 snapshots of the input stimulation. The first snapshot always corresponds to $t = 0$, where the algorithm gets some time to converge on one of the stimuli (and there is only one but for scenario C). The second snapshot is associated to a later time, chosen to be representative of the input dynamics. The position of the target, its trajectory as well as the occluded area for scenario G are superimposed on the fields.

of the motor cortices [8]). At this internal level again, perception-action loops are required to monitor the effect of modulations between neural maps.

A parallel can also be made with the correlations we learn about signals sent from the brain to muscles (motor commands) and the sensory feedback perceived in return (proprioception). Such relationships between the afferent and efferent signals in neural maps can be used to influence the dynamics of our body and bias the decisions (for instance by modulating central pattern generators) without having a direct control over the dynamics of the muscles.

IV. Results

The discrete implementation described in [16] is used as a reference, and the same evaluation process is adopted. Choosing adequate parameters, global inhibition guarantees a single bubble will emerge, and we thus evaluate the performance of the model using the distance defined in Eq. 5. To test the robustness and attentional properties of the model, we provide 2D artificial dynamic inputs to the system. On a toric manifold mapped on $[-0.5, 0.5]^2$ and approximated by a square matrix of size 50×50, bell-shaped stimuli of standard deviation 0.1 are used for both the target and distracters. The input dynamics are defined by the following scenarios (scenarios C, D and E being taken from [13]):

- C **Competition**: 2 distant stimuli are introduced at time $t = 0$. Their intensity are governed by $S_1 = 0.5 - 0.5\sin(\pi \times (t/10))$ and $S_2(t) = 0.5 + 0.5\sin(\pi \times (t/10))$.
- D **Distracters**: 1 stimulus follows a circular trajectory of radius 0.2 around the point $(0, 0)$ at 30 deg/s from $t = 0$. From $t = 1$, 30 distracters are added and take new random positions on the field every second.
- E **Noise**: 1 stimulus (same as above). At $t = 1$, Gaussian noise with a standard deviation of 0.5 is added and refreshed every second.
- F **Fixed distracter**: 1 stimulus (same as above). At $t = 5$, a fixed distracter is added at $(0, -0.2)$, i.e. on the trajectory of the target.
- G **Occlusion**: 1 stimulus (same as above, except for a speed of 10 deg/s). At $t = 30$, a fixed distracter is added out of the trajectory, and the stimulus is then occluded on part of the trajectory for $0 < x < 0.5$ and $-0.1 < y < 0.1$.

Except for scenario C, the amplitudes of the target and distracters in the input signal are identical and arbitrarily taken to 1.0. The initial difference in intensity or the delay introduced before distracters and noise are added guarantee that the model will converge on the chosen stimulus during the first few iterations. In order to prove the efficiency and usefulness of the prediction term, optimal parameters as defined in [13] have been independently selected for each scenario. This guarantees the standard equation is not simply outperformed because of a poor calibration of its dynamics.

The mean error over entire simulations and averaged over several trials for non deterministic scenarios (noise and random distracters) is given in Table I for the model in 3 different conditions: without prediction (reference) and with both a correct and incorrect prediction. In the table, gray cells highlight the best performance for each scenario, most often when the equation is extended with an adequate predictor. It might be argued that introducing a correct prediction into the equation should of course lead to performance improvement, as we bias the dynamical attractor landscape toward the right trajectory. Another interpretation is to say that prediction stands for a prior probability distribution in the selection and tracking process. However, the efficiency of such method on an already robust yet complex dynamics is far from obvious, not even mentioning the importance of the γ function. The detailed analysis of the dynamics on the different scenarios also gives precious insights into the potential of coupled competition-prediction systems.

TABLE I
Mean tracking error for various scenarios

Prediction	Mean error on scenario				
	C	*D*	*E*	*F*	G
None	0.0066	0.179	0.047	0.090	0.082
Correct	0.0079	0.095	0.032	0.036	0.041
Incorrect	0.0407	0.156	0.081	0.123	0.174

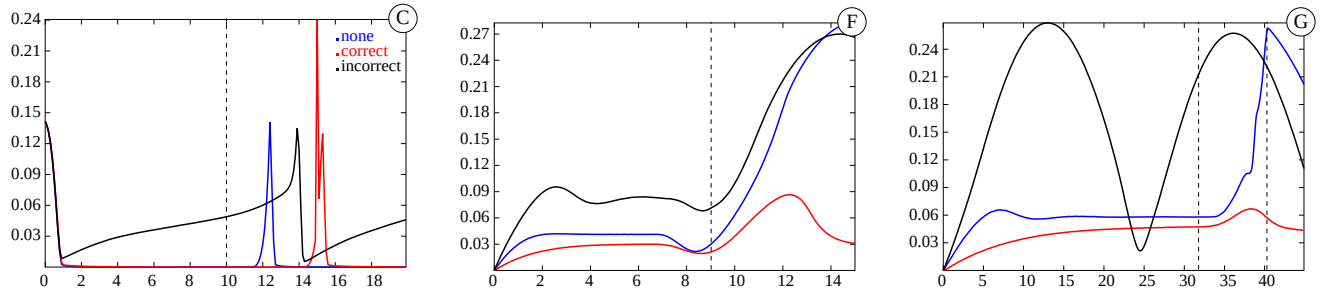

Fig. 6. Error distance as a function of time for various scenarios with no prediction (*blue*), a correct prediction (*red*) or an incorrect one (*black*). The black lines represent critical events: equal stimuli intensity for scenario C, going over the static distracter for scenario F, or entering/exiting the occluder in G. When the error goes above 0.1, one can consider the target has been lost as it entered the inhibitory part of the lateral competition kernel.

Scenario C instantiates the competition property of the CNFT between distant stimuli, and since they are not moving, the ideal prediction corresponds to a speed $\mathbf{v} = 0$. Initially, the system must choose a single target from two almost identical stimuli ($S_1 \simeq S_2 \simeq 0.5$), and the slightest asymmetry in their intensities should lead to a non-linear bifurcation and the selection of the most salient target. Fig. 6C shows that the convergence occurs at the same rate with or without prediction, which is normal since the target is not moving and stationarity is the default in the original equation. However, in the case of an incorrect prediction constantly anticipating a movement, the bubble progressively drifts. Such an anticipation that constantly conflicts with the target dynamics should not remain active when competing with better predictors, but it is not the case here as the effect of a single predictor is considered.

Also, it should be noted that the prediction increases the self-maintenance of the bubble activity relatively to the stimulation because of the reduced input weight it takes in the equation. This results in the increasing hysteresis phenomenon observed on the graph, the model keeping track of the initial target long after $t = 10$ (time at which the dominant stimulus changes). The delayed shifts visible on the figure reflect the increased attention and sustained focus activity with prediction. Such extension of the time spent focusing on a target is desired, as the system should track the target as long as it can be perceived. Whatever the difference in intensity required for the bifurcation to occur, the model is still able to relax and focus on a new stimulus when the stimulation totally disappears.

Scenario D for distracters and E for noise test the robustness of the model to random and non random perturbations in the input. Again, the parameters of the equation are chosen such as to maximize the model performance on these scenarios only, and performance is therefore higher when compared with results presented in [16], where tradeoffs are necessary to instantiate all attentional properties at once. When integrated into the lateral competition term c through a convolution with the kernel w, noise and distracters are equivalent to shifts of the input target location. Whereas the probability of a random noise to produce a large translation in any particular direction is quite low, distracters introduce clear oriented distortions in the local activity. As a consequence, an incorrect predictor simply acts as a distracter close to the target trajectory, which does not fundamentally hinder performance for scenario D as many distracters are already present on the field, in contrast to scenario E and as reflected in Table I. In addition to the results given in the table demonstrating a gain of 33% to 50% when a roughly correct prediction is used, Fig 7 also shows the mean error committed as a function of the noise level, i.e. the standard deviation of the Gaussian distribution with values in the interval $[0, 1]$. The bias in activity from the prediction results in a relative decrease of the noise level, similarly to what was illustrated on Fig 2.

More generally, testing the model with an invalid prediction shows that the system is able to fall back on a reactive tracking behavior. The effect of bad predictions on the dynamics is limited because of a conjunction of factors. First, the α coefficient in Eq. 3 gives primacy to the external stimulation. The non coherent stimulus and predicted bubble activity are spread out and their intensity is therefore reduced. As a consequence and because of the way the γ function pushes the prediction outside the excitatory range of the lateral competition kernel, predictions that are not confirmed by the external stimulation get strongly inhibited. Therefore, even without associating a variable weight to the predictor based on its level of assimilation of the dynamics, the influence of the prediction is highly constrained by the dynamics of the CNFT and stimulation.

Finally, scenario F and G correspond to tasks the standard model alone cannot perform correctly. Although the mean error values of Table I display a gain of only 50-60% with a correct prediction, this is mainly due to the experimental setup, and the model in the other conditions is in fact not able to succeed at tracking the target with a fixed distracter or an occluder on the trajectory. The inability for the standard model to discriminate between the target and distracter (Fig 6F), as well as to maintain a

moving bubble during several timesteps without stimulation (Fig 6G) is made flagrant by looking at the error curves as a function of time. Due to the toric nature of the field and the periodic trajectories of the stimuli, a large error (> 0.2) might sometimes go down progressively or oscillate, which does not always mean the algorithm was able to focus again on the correct target. It just means for instance that the focus shifted to a static distracter (the most stable attractor for the standard equation) while the target moves toward or away from it. This is the case for the no predictor and bad predictor conditions on scenario F at the end of the simulation, after they both lost the target when going over the distracter. This is also particularly striking for scenario G with the bad predictor, because the high integration constant τ required to increase the inertia of the standard model (selected by the genetic algorithms) leads to a self-maintaining bubble with its own dynamics.

V. Conclusions

We presented an extension of the inhibitory single layer CNFT model to integrate predictive signals. This new model displayed improved attentional properties and tracking capabilities over the standard version, with a fallback on the original dynamics would the prediction be incorrect. In particular, it is now possible for the stimulus to go over stationary or moving distracters without loosing track of the selected target. Whereas the excitatory projection from the prediction here amplifies the hysteresis phenomenon observed when switching between targets, inhibitory projections can be used to ease and fasten transitions to unexpected stimuli. We also showed that the relationship between the predictive modulation to introduce in the equation and its effect on the dynamics is non-trivial but coherent within the context of distributed interactive systems.

Since the paper focused on the method and benefits of combining dynamic competition and prediction within a single model, a number of key issues have been overlooked.

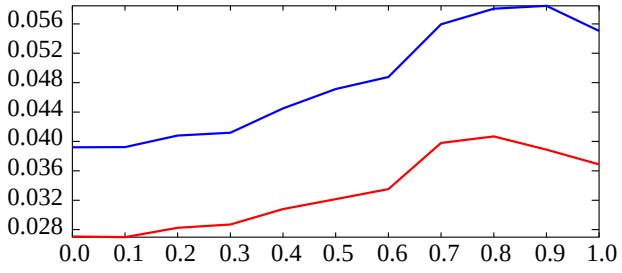

Fig. 7. Mean error distance for different levels of noise with (*red*) or without prediction (*blue*). Due to the dynamics of stabilization through activity feedback acting against the mobility of a low amplitude bubble, a large level of noise might lead to a more reactive system. With a single spatiotemporally coherent stimulus to track on the field, this explains the otherwise strange performance boost observed when the standard deviation goes over 0.8, as the system cannot spontaneously focus on another target.

Most of these questions are addressed by ongoing work, and the potential solutions should at least be introduced here. First, where does the predictor come from? As the CNFT is able to reactively track the stimuli when the prediction is absent or invalid, the resulting bubble can be used to progressively learn the predictor and improve performance. Echoing the arguments made in the previous sections, any complex trajectory can be decomposed into a set of local linear predictions, the dynamic competition smoothing out the transitions. Such linear anticipations can be generated in an ad hoc manner, by converting the observed bubble movement into anticipations whenever the accuracy drops below a given threshold. A variation/selection mechanism is conceivable as well, thus removing the need to evaluate the γ function a priori. Indeed, the "action" and its consequences could be totally decoupled and only confirmed anticipations would be reinforced.

When the system encounters a wide range of dynamics, several predictors are needed to cover the space of potentialities, but only those best adapted to the current context should be active. For this reason, a variable weight should be associated to each predictor and to the external stimulation, even though the latter should keep primacy. In the same way as local anticipations can get reinforced during learning, the immediate and relative level of assimilation of the dynamics by the predictors can be used to discriminate between situations. The presented model, by building on the normativity of the anticipation, could then be applied to online spatiotemporal pattern recognition.

References

[1] F. Alexandre, J. Fix, A. Hutt, N. Rougier, and T. Viéville, "On practical neural field parameters adjustment," in *Deuxième conférence française de Neurosciences Computationnelles*, 2008.

[2] S.-I. Amari, "Dynamics of pattern formation in lateral-inhibition type neural fields," *Biological Cybernetics*, vol. 27, no. 2, pp. 77–87, 1977.

[3] K. Brodmann, *Brodmann's 'Localisation in the Cerebral Cortex'*. Smith-Gordon, London, 1909/1994.

[4] Y. Burnod, *An adaptive neural network: The cerebral cortex*. Masson, 1989.

[5] M. Cerda and B. Girau, "Bio-inspired visual sequences classification," in *Brain Inspired Cognitive Systems 2010 Brain Inspired Cognitive Systems 2010 - BICS 2010*, Madrid Spain, 2010, p. pp. 20.

[6] W. Cleveland, "Robust locally weighted regression and smoothing scatterplots," *Journal of the American Statistical Association*, vol. 74, pp. 829–836, 1979.

[7] J. Fix, N. Rougier, and F. Alexandre, "A dynamic neural field approach to the covert and overt deployment of spatial attention," *Cognitive Computation*, pp. 1–15, 2010.

[8] J. M. Fuster, *Cortex and Mind: Unifying Cognition*. Oxford University Press, 2003.

[9] D. Hubel and T. Wiesel, "Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex," *Journal of Physiology*, vol. 160, pp. 106–54, 1962.

[10] O. Ménard and H. Frezza-Buet, "Model of multi-modal cortical processing: coherent learning in self-organizing modules," *Neural Networks*, vol. 18, no. 5-6, pp. 646–55, 2005.

[11] V. Mountcastle, "Modality and topographic properties of single neurons of cat's somatic sensory cortex," *J Neurophysiol*, vol. 20, no. 4, pp. 408–434, 1957.

[12] D. Purves, G. J. Augustine, D. Fitzpatrick, W. C. Hall, A.-S. LaMantia, J. O. McNamara, and L. E. White, *Neuroscience, Fourth Edition*, 4th ed. Sinauer Associates, Inc., 2008.
[13] J.-C. Quinton, "Exploring and optimizing dynamic neural fields parameters using genetic algorithms," in *Proceedings of IEEE World Congress on Computational Intelligence (IJCNN 2010) (Barcelona, Spain)*, 2010.
[14] J.-C. Quinton and J.-C. Buisson, "Multilevel anticipative interactions for goal oriented behaviors," in *Proceedings of EpiRob'08 - International Conference on Epigenetic Robotics (Brighton, UK)*. Lund University Cognitive Studies, 2008, pp. 103–110.
[15] J.-C. Quinton and B. Girau, "A sparse implementation of dynamic competition in continuous neural fields," in *Proceedings of Brain Inspired Cognitive Systems (BICS'2010) (Madrid, Spain)*, 2010.
[16] N. P. Rougier and J. Vitay, "Emergence of attention within a neural population," *Neural Netw.*, vol. 19, no. 5, pp. 573–581, 2006.
[17] J. Taylor, "Neural bubble dynamics in two dimensions: Foundations," *Biological Cybernetics*, vol. 80, pp. 5167–5174, 1999.
[18] H. R. Wilson and J. D. Cowan, "A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue," *Kybernetic*, vol. 13, no. 2, pp. 55–80, 1973.

APPENDIX

Let the definite integral of a Gaussian be:

$$g_{[x_1,x_2]}(\mu,\sigma) = \int_{x_1}^{x_2} g(x-\mu,\sigma)dx = \frac{\sigma\sqrt{\pi}}{2}\left(erf\left(\frac{x_2-\mu}{\sigma}\right) - erf\left(\frac{x_1-\mu}{\sigma}\right)\right) \tag{12}$$

where erf is the standard Gaussian error function defined by:

$$erf(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-x^2}dx \tag{13}$$

For the computation of the centroid of the focus, we also need the definite integral of $g(x-\mu,\sigma)x$:

$$gx_{[x_1,x_2]}(\mu,\sigma) = -\frac{\sigma^2}{2}\left(g(x2-\mu,\sigma) - g(x1-\mu,\sigma)\right) + \mu g_{[x_1,x_2]}(\mu,\sigma) \tag{14}$$

To simplify the expressions in the following sections, we also introduce a function gd:

$$\begin{aligned} gd_{[x_1,x_2]}(\mu,\sigma) &= gx_{[x_1,x_2]}(\mu,\sigma) - vdt \times g_{[x_1,x_2]}(\mu,\sigma) \\ &= -\frac{\sigma^2}{2}\left(g(x2-\mu,\sigma) - g(x1-\mu,\sigma)\right) \\ &+ (\mu - vdt)g_{[x_1,x_2]}(\mu,\sigma) \end{aligned} \tag{15}$$

The product of non-normal Gaussian distributions is as well needed for the convolution:

$$g(x-\mu_1,\sigma_1)g(x-\mu_2,\sigma_2) = e^{-\frac{(\mu_1-\mu_2)^2}{\sigma_1^2+\sigma_2^2}} g(x-\mu_m,\sigma_m)$$
$$\mu_m = \frac{\mu_1\sigma_2^2+\mu_2\sigma_1^2}{\sigma_1^2+\sigma_2^2} \qquad \sigma_m^2 = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2} \tag{16}$$

Proof: The initial expression can be written as $e^{-\omega}$, with w defined by:

$$\omega = \frac{\sigma_2^2(x-\mu_1)^2+\sigma_1^2(x-\mu_2)^2}{\sigma_1^2\sigma_2^2} = \frac{(x-\mu_m)^2}{\sigma_m^2} + \varepsilon$$

$$\varepsilon = \frac{1}{\sigma_m^2}\left(-\mu_m^2 + \frac{\sigma_2^2\mu_1^2+\sigma_1^2\mu_2^2}{\sigma_1^2+\sigma_2^2}\right) = \frac{(\mu_1-\mu_2)^2}{\sigma_1^2+\sigma_2^2}$$

■

Using Eq. 16, the convolution of a Gaussian bubble with the lateral kernel on a point is trivially obtained:

$$\begin{aligned} c(x) &= \int_{-\infty}^{+\infty} w(x,x')u(x',t)dx' \\ &= \sigma\sqrt{\pi}\left(\frac{aA}{\sqrt{a^2+\sigma^2}}g\left(x,\sqrt{a^2+\sigma^2}\right)\right. \\ &\left. -\frac{bB}{\sqrt{b^2+\sigma^2}}g\left(x,\sqrt{b^2+\sigma^2}\right)\right)U \end{aligned} \tag{17}$$

We now focus on the relationship between U and S for the optimal γ, which is given by:

$$U = \frac{1-\alpha}{\lambda}S \tag{18}$$
$$\begin{aligned} \lambda &= \frac{\tau}{dt} + \left(1-\frac{\tau}{dt}\right)g(vdt,\sigma) - \alpha g((\gamma(v)-v)dt,\sigma) \\ &- \sigma\sqrt{\pi}\left(\frac{aA}{\sqrt{a^2+\sigma^2}}g\left(vdt,\sqrt{a^2+\sigma^2}\right)\right. \\ &\left. -\frac{bB}{\sqrt{b^2+\sigma^2}}g\left(vdt,\sqrt{b^2+\sigma^2}\right)\right) \end{aligned}$$

Proof: For the optimal γ, we are looking for a stable value of U between t and $t+dt$. Because the location vdt should be the center of the new bubble, so that $u^{t+dt}(vdt) = U > 0$, we can write:

$$\begin{aligned} U &= \left(1-\frac{dt}{\tau}\right)u(vdt) \\ &+ \frac{dt}{\tau}\left(c(vdt) + \alpha p(vdt) + (1-\alpha)s(vdt)\right) \end{aligned}$$

which trivially leads to the result by using Eq. 17. ■

We now evaluate the simplified expression of $e_I = e_{[vdt-\theta,vdt+\theta]}$. Since $(1-\alpha)Sgd_I(vdt,c))$ is null, as the stimulation is symmetric around vdt (this remains true for other inputs as long as they remain symmetric in the interval I), we can factorize the expression as follows:

$$\begin{aligned} e_I(\gamma) &= \left(\frac{\alpha dt}{\tau}gd_I(\gamma(v)dt,\sigma) + \eta\right)U \\ \eta &= \left(1-\frac{dt}{\tau}\right)gd_I(0,\sigma) \\ &+ \frac{\sigma\sqrt{\pi}dt}{\tau}\left(\frac{aA}{\sqrt{a^2+\sigma^2}}gd_I\left(0,\sqrt{a^2+\sigma^2}\right)\right. \\ &\left. -\frac{bB}{\sqrt{b^2+\sigma^2}}gd_I\left(0,\sqrt{b^2+\sigma^2}\right)\right) \end{aligned}$$

With U remaining a function of γ (Eq. 18), and η almost a constant (as γ still slightly constrains the values of θ and σ).