

Exploring and Optimizing Dynamic Neural Fields Parameters Using Genetic Algorithms

Jean-Charles Quinton

► **To cite this version:**

Jean-Charles Quinton. Exploring and Optimizing Dynamic Neural Fields Parameters Using Genetic Algorithms. IEEE World Congress on Computational Intelligence 2010 - WCCI 2010, Jul 2010, Barcelona, Spain. 2010. <inria-00488914>

HAL Id: inria-00488914

<https://hal.inria.fr/inria-00488914>

Submitted on 3 Jun 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Exploring and Optimizing Dynamic Neural Fields Parameters Using Genetic Algorithms

Jean-Charles Quinton

Abstract—The Continuous Neural Field Theory introduces biologically-inspired competition mechanisms in computational models of perception and action. This paper deals with the use of Genetic Algorithms to optimize its parameters, as to guarantee the emergence of robust cognitive properties. Such properties include the tracking of initially salient stimuli despite strong noise and distracters. Interactions between the parameter values, input dynamics and accuracy of model, as well as their implications for Genetic Algorithms are discussed. The fitness function and set of scenarios used to evaluate the parameters through simulation must be carefully chosen. Experimental results reflect an ineluctable tradeoff between generality and performance.

I. INTRODUCTION

A. Continuous Neural Field Theory (CNFT)

The cortex can be roughly approximated as a two dimensional structure made of myriads of elementary units called cortical columns [5]. Despite this apparent homogeneity, the organization of the body and innate structure of the brain constrain the afferent connections in different parts of the cortex. Self-organization and competition within the cortex lead to differentiated areas dedicated to specific modalities, communicating through long-range connections and associative areas. But this process furthermore introduces a local topological organization within the cortical sheet, where close units tend to respond to similar input configurations [10]. This local correlation reinforces local excitatory connections, while balanced by global inhibition. Whether the connectivity patterns and input driven organization are partly innate or engaged in a form of circular causality, this phenomenon leads to the emergence of important perceptual and behavioral properties.

Extensive research has been done on how lateral interactions might evolve through the course of development [9], [8], the Continuous Neural Field Theory (CNFT) also examines the dynamic properties of such neural maps [4]. Properties include the emergence of bubbles of activity able to robustly track input stimuli in spite of strong noise or numerous distracters [11]. This kind of attentional property combined with the dynamic competition resulting from the large scale inhibition is one of the characteristics needed to adopt complex goal-oriented behaviors. Bubbles can indeed shape the attractor landscape and constrain the fast

dynamics of sensorimotor maps over an extended period of time. Radical changes constantly occur on low-level sensory maps during the course of many common actions and would lead to chaotic behavior without such modulations (consider for instance the optical flow when temporarily turning away from a visual target).

Although general convergence results have been demonstrated and parameters constrained to allow the emergence of such bubbles of activity [12], many parameters need to be tuned through simulation. Mathematical analysis indeed provides simple rules to determine functional parameters for which numerical simulations can be run [3], and experimentation then allows to further test the robustness of the model, but the search for generic and optimal parameter values is limited for complex input dynamics. Parameters tightly interact in the equations and a slight change of a single parameter sometimes lead to dramatic changes in the behavior of the model on specific simulation scenarios. With the assumption that correct sets of parameters should lead to the emergence of bubbles able to robustly track salient input stimuli, many values are simply non functional: they either lead to the saturation of the field, the lack of any coherent activity, or the self-maintenance of static bubbles independently of the input dynamics. In this paper, we therefore focus on the selection of optimal parameters using Genetic Algorithms.

B. Genetic Algorithms (GA)

There are several reasons for choosing Genetic Algorithms (GA) as a metaheuristic for optimizing the attentional properties of the CNFT. Since we are interested in the robustness of the model, randomness is introduced in the input as to increase its variability and limit the assumptions to the kind of noise, shape and trajectory of distracters used. Whereas no random term is introduced in the equations governing the model and that its inner dynamics is entirely deterministic, several subsequent simulations may lead to totally different results. With inadequate parameter values and noisy inputs, the initial phase leading to the potential emergence of a bubble is highly chaotic, the reinforcement of the bubble requiring spatiotemporally coherent stimulations. In this sense, GA allow to test parameter values again and again for each generation, not simply relying on a single successful trial for a single individual. Purely coincidental survival of a model configuration will lead to its elimination and that of its progeny with a high probability.

TABLE I
DESCRIPTION OF THE CNFT PARAMETERS USED TO OPTIMIZE THE MODEL BEHAVIOR.

Parameter	Description
τ	Inertial parameter. How fast can the activity change, whatever the dynamics of the stimulation.
A	Excitatory amplitude. How much close units should form coherent patches of activity.
a	Excitatory range. This constrains the size of the emerging bubbles.
B	Inhibitory amplitude. How much distant units should inhibit each others (relatively to A).
b	Inhibitory range. If $b \geq 1$ (size of the neural field), we can consider the inhibition to be global.

From experimental results and hand-tuning of parameters, there may be several distant and local optima, wide plateaus and even ridges for some subregions of the parameter space. This makes simple methods like the steepest descent inadequate for the problem, since a convex function was not found to evaluate the performance of the model. Even if it must be underlined that many local minima lead to acceptable results, the search for optimal parameters also make simulated annealing techniques not very efficient though perfectly adapted to large search spaces, as restarts are often needed. Once groups within the global population have converged to local minima, GA keep the adaptive pressure increasing while still allowing short and long jumps within the search space respectively through mutations and cross-overs.

Finally, it is not obvious that fixed parameter values can generically provide a good functional basis on which ontogenetically reinforced synaptic connections might self-organize. Using GA to evolve the parameters is therefore a way to test the biological plausibility of having fixed global inhibition and local excitation, instead of having them co-emerge with or simply adapt to other cortical and thalamic connections through learning and development.

II. METHOD

A. CNFT equations and parameters to optimize

Adopting the notations introduced by Amari [4], we will present a version of the CNFT using a single-layer field equation of lateral inhibition type. The focus neural field on which bubbles may form is represented by a manifold M in bijection with $[-0.5, 0.5]^2$ and the membrane potential at the position vector \mathbf{x} and time t on this field by $u(\mathbf{x}, t)$. Similar notations are used for the stimulation $s(\mathbf{x}, t)$, that is the only external input to the system. The dynamics of the membrane potential is described by Equation (1).

$$\tau \frac{\partial u(\mathbf{x}, t)}{\partial t} = -u(\mathbf{x}, t) + \int_{x' \in M} w(\mathbf{x}, \mathbf{x}') u(\mathbf{x}', t) d\mathbf{x}' + s(\mathbf{x}, t) + h \quad (1)$$

where h is the resting potential (taken equal to 0) and $w(\mathbf{x}, \mathbf{x}')$ the lateral connection weight function satisfying the

following equation:

$$w(\mathbf{x}, \mathbf{x}') = Ae^{-\frac{|\mathbf{x}-\mathbf{x}'|^2}{a^2}} - Be^{-\frac{|\mathbf{x}-\mathbf{x}'|^2}{b^2}} \quad (2)$$

At first glance, Equation (1) may seem to lack degrees of freedom, as the relaxation term $-u(\mathbf{x}, t)$ and stimulation $s(\mathbf{x}, t)$ share the same coefficient. However, whereas introducing non-linearities in the lateral competition term (such as a sigmoidal function) or integration of receptive fields in the stimulation term might totally change the dynamics of the CNFT, introducing a simple multiplicative coefficient for the input will only scale the focus (as long as h is null). Indeed, if we replace $s(\mathbf{x}, t)$ by $\alpha * s(\mathbf{x}, t)$ in Equation (1) and let $u' = u/\alpha$, u' then satisfies the original equation (i.e. without α).

When discretizing this equation on a regular mesh of units for simulation purpose, a spatial resolution parameter n is additionally introduced. However, the dynamics of the model is theoretically independent of both the spatial and temporal discretization parameters, except when chosen too coarse (taking a low n or high dt). This was confirmed in experiments with n ranging from 10 to 100, with $n = 30$ being sufficient for the expected properties to emerge. In the following sections, let $n = 50$ units and $dt = 0.1$ seconds. These parameters will therefore not be optimized and should not be referred to anymore.

Remaining parameters of Equation (1) and (2) that need to be tuned are synthesized in Table I. Although most parameters are limited to positive values, mathematical analysis of the equations further constrains the parameters to allow the emergence and disappearance of bubbles of activity, avoiding at the same time the saturation of the field [12]. A Mexican hat shape for the lateral weight function is thus adopted here, implying the conditions $A > B$ and $a < b$ to be satisfied. As to introduce such constraints in the evolutionary process, we let $K = B/A$ and $k = a/b$ with $(K, k) \in [0, 1] \times]0, 1]$. This does not reduce the number of parameters to be optimized but limits the dramatic effect that slight changes in parameters might have on the dynamics (for instance if a mutation leads from $A > B$ to $A + dA < B$).

B. GA implementation

GA are applied to evolve a population of models, each individual genotype being encoded by a parameter vector $\mathbf{v}_i = (p_i^k)_{k \in [1..s]}$, s being the number of parameters. In our case, this vector must account for all CNFT parameters and is thus defined by $\mathbf{v}_i = (A_i, k_i, K_i, b_i, \tau_i)$. The phenotype translation corresponds to the straightforward generation of the lateral connectivity pattern and the use of differential equation parameter τ_i in Equation (1).

A new population $g + 1$ is then composed of the individuals that were the most successful at generation g (see the definition of the fitness below) as well as new individuals obtained through the cross-over and mutation of two individuals taken randomly from population g . This guarantees that the best results ever obtained during the artificial evolutionary process will never be discarded from the population. The size of the population remains the same over generations as to maintain a constant evaluation cost. The cross-over operation produces a new vector where each parameter is assigned to a random weighted sum of the parent values. Mathematically, a genotype $\mathbf{v}_c = (\gamma^k p_1^k + (1 - \gamma^k) p_2^k)$ is created where \mathbf{v}_1 and \mathbf{v}_2 would be the genotypes of the parents and (γ^k) random coefficients taken in $(0; 1)^s$. Mutations are random variations Δp_c^k applied to random components of the child vector (probability p_{mut}). Let $\Delta p_c^k = \pi^k |p_{max}^k - p_{min}^k|$ with π^k randomly taken in $[-m, m]$ and m a factor introduced to constrain the mutation maximum amplitude. However, the algorithm also ensures that the associated parameter value remains within its imposed limits $[p_{min}^k, p_{max}^k]$, as $K \in [0, 1]$ for instance. Although dependent on the chosen mutation range, these operations allow fast jumps over the entire parameter space, expansion through mutations and contraction through cross-overs.

Whereas GA are quite generic, the difficulty generally lies in the choice of the right genes and correct fitness function. In our case, the genes selection and encoding is intuitive and straightforward, but the fitness function goes through the evaluation of complex dynamical simulations. The overall process and its exact formulation must therefore be specified with great care. The following two sections are dedicated to the underlying assumptions and resulting consequences of the choice of fitness and associated scenarios used for its evaluation.

C. Fitness thoughts

Since the focus is on the distributed and emergent properties of the CNFT, a static evaluation of the parameters is not sufficient and simulations of the model must be run. The problem is now the high dependence of performance on the chosen scenario. If the goal is to optimize the entire spectrum of emergent properties displayed by the CNFT models, all these properties must be reflected in the selected scenarios.

As the model might easily bifurcate in the first few seconds of simulation based on initial conditions and amount of noise introduced, the fitness function is computed on a sequence of n stereotyped scenarios. This allows to easily account for the various properties of the model, but also for the inter-simulation variability. The global fitness $f(\mathbf{v})$ for a given parameter vector \mathbf{v} is defined by:

$$f(\mathbf{v}) = \frac{1}{n} \sum_{k=1}^n f(s_k, \mathbf{v})$$

where $f(s_k, \mathbf{v})$ corresponds to the evaluation of \mathbf{v} for a single scenario s_k . This subfunction must be defined as to synthesize the accuracy and efficiency of the model over the entire simulation in a unique scalar value, defined as follows.

$$f(s_k, \mathbf{v}) = error(s_k, \mathbf{v}) \times conv(s_k, \mathbf{v}) \times shape(s_k, \mathbf{v}) \quad (3)$$

In Equation (3), *error* corresponds to the mean normalized distance between the center of the focus bubble and the center of the input stimulus the model is supposed to track and focus on. An instantaneous error (e_t) is thus computed for each iteration in the simulation, and then averaged on the last seconds of simulation to rule out the initial chaotic dynamics. This computation already assumes that the quality of the model is correlated to its ability to make a single stereotyped bubble emerge. For the details of the approximation of the center of the focus bubble, please refer to [11]. This *error* measure also supposes that we a priori know or can easily determine which input stimulus is attended, and that so-called input stimuli merely exist. To make all these points clear, we determine the input flow independently of the model dynamics, as designers and evaluators. The model has indeed no mean of changing the input dynamics by acting on it (as would a typical sensorimotor system engaged in closed loop interactions with its environment). We therefore define what is considered to be coherent stimuli for the model and optimize the parameters in relation to this definition. As observers of the model dynamics, we can also compute

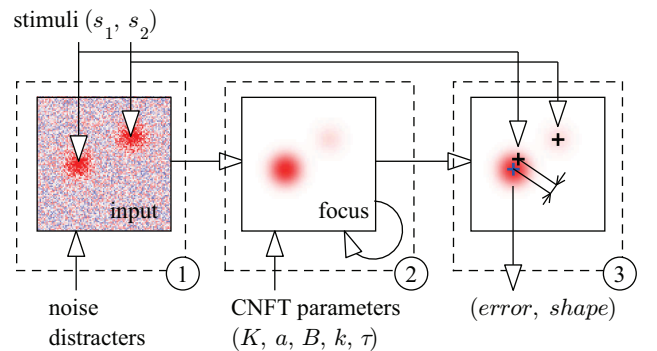


Fig. 1. Simulation steps for one iteration. ① The input map is generated from the current scenario parameters. ② The model is updated, only using the input map and CNFT parameters. ③ Statistics are computed based on the known input positions and focus map produced by the model.

statistics on the model and determine the closest stimulus to the focus bubble, information the model has of course no access to. Figure 1 proposes an illustration of these considerations.

Again in Equation (3), $conv$ is defined as the convergence time, i.e. the time needed for e_t to drop below a given relative threshold. In other words, let $conv$ be the minimal time such that whatever $t_c \in [conv, t_{end}]$, $e^{t_c} < \alpha \times \min_t(e_t) + (1 - \alpha) \times \max_t(e_t)$, where α is the threshold parameter set to 0.2 (see Figure 2). If the condition cannot be satisfied, $conv$ is set to the ending time of the simulation t_{end} . An additional variable $shape$ is also taken into consideration to constrain the profile of the focus bubble. A lack of inhibition on the focus field or a limited coherence of the activity under the bubble may lead to a focus field with a wide support (in the sense of measure theory) and non stereotyped bubbles. Although the computation of the bubble center is not altered by such distortions, spread out activations may be inadequate for decision or motor control. At a given instant t of simulation, the instantaneous shape computation satisfies:

$$shape_{e_t} = \int_{x \in M} |u^*(\mathbf{x}, t) - u(\mathbf{x}, t)| d\mathbf{x}$$

To put it briefly, this equation computes a similarity between the actual focus u and an hypothetical ideal focus field u^* where a single stereotyped bubble would emerge in the presence of the tracked stimulus, defined by:

$$u^*(\mathbf{x}, t) = \frac{I^t}{A - B} w^+(\mathbf{c}^t, \mathbf{x})^2$$

where w^+ is the positive part of the lateral weight function w defined in Equation (2), \mathbf{c}^t the bubble center at time t and I^t the tracked stimulus amplitude (used as a normalization factor for the CNFT profile whose maximum value is $A - B$).

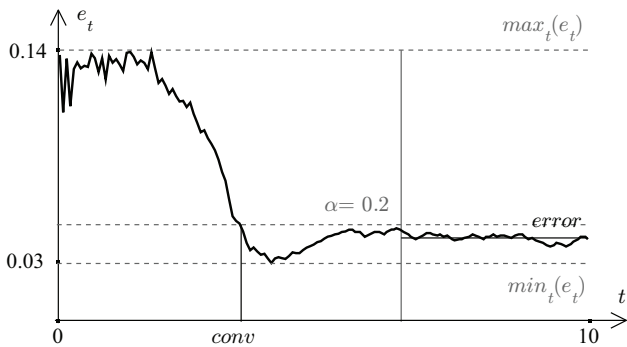


Fig. 2. Typical profile of the error when stimuli are competing for the focus. After a few iterations where noise dominates and the focus center is close to the center of the map, the non-linearity introduced by the lateral competition induces a rapid convergence to one of the input stimulus. The instantaneous error e_t drastically decreases and definitely fall below the threshold used for the convergence time $conv$. Once the initial chaotic behavior has settled, the mean $error$ is computed on the last seconds of simulation.

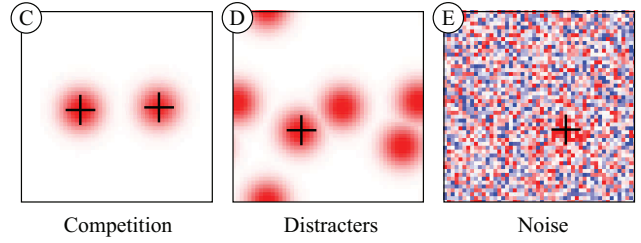


Fig. 3. Stereotypical input for each scenario. The field is approximated by a grid of $n \times n$ units, where $n = 50$. Although not accessible to the model, the input stimuli are marked with *black crosses* for the reader to easily locate them.

D. Scenarios

As stated in the previous section, several complementary scenarios are needed to account for the various emergent properties of the CNFT (already described in [4], [11] and [6]). For instance, the a value is highly correlated to the distance between the stimuli to focus on: a large value may improve the results only when unambiguous distant stimuli are presented. If they are too close, the CNFT will merge their activity under a single bubble that will not track either of the stimuli. But the model should not be dependent on the input dynamics within a given functional range, even if a multi-scale, multi-resolution approach should be adopted for a wider range of conditions. The parameters are thus optimized as to produce the best results for all three following scenarios, although non optimal for each scenario taken separately:

- C** 2 bell-shaped distant stimuli s_1 and s_2 are introduced at time $t = 0$. Their intensity are refreshed every 5 seconds and governed by $I_1 = 0.9$ and $I_2(t) = 0.5 + 0.5 \cos(\pi \times (t/5))$.
- D** 1 bell-shaped stimulus of standard deviation 0.1 and intensity 1.0 follow a circular trajectory of radius 0.2 around the point $(0, 0)$ at 10 deg/s from $t = 0$. From $t = 1$, 5 distracters are added and take new random positions on the field every 1 s.
- E** 1 bell-shaped stimulus (same as above). At $t = 0$, Gaussian noise with a standard deviation of 0.5 is added.

Scenario **C** instantiates the competition property of the CNFT. The model should initially focus on the most salient target s_2 before switching when its intensity will drop to 0 for $t = 5$. The shifts in attention introduce an hysteresis phenomenon in the dynamics, as the resilience to noise and distracters partially comes from the combined effect of the feedback and input. For $t > 5$, the focus should remain on s_1 as s_2 will never reach a significantly higher intensity (at most 0.1 higher). This scenario also tests the ability of the model to relax when the stimulation disappears, i.e. not to sustain saturated bubbles independently of the input.

Scenario **D** tests the capability of the model to track a single stimulus despite distracters of the same intensity.

TABLE II

OPTIMAL PARAMETERS AND SIMULATION RESULTS OBTAINED WITH DIFFERENT SETS OF FITNESS VARIABLES AND SCENARIOS
(CELLS IN GRAY CORRESPOND TO THE SIMULATIONS WITH QUALITATIVELY LOWER PERFORMANCE THAN THE OTHERS)

GA scenarios	Parameters					<i>error</i> for test scenarios			
	A	a	B	b	τ	C	D	E	E'
$C, D, E', \neg shape$	0.135	0.34	0.128	1.00	0.29	0.05×10^{-4}	0.019	0.270	0.009
$C, D, E', shape$	0.069	0.14	0.041	1.25	0.27	0.12×10^{-4}	0.019	0.256	0.015
$C, D, E, shape$	0.074	0.28	0.062	0.88	0.45	0.75×10^{-4}	0.040	0.016	0.017
$E, shape$	0.065	0.38	0.061	0.93	0.64	2.08×10^{-4}	0.071	0.016	0.017

As to guarantee their spatiotemporal coherence, distracters are only randomly moved every second, allowing them to destabilize the focus for several iterations. Since the model cannot distinguish between the stimulus and the distracters, these are introduced with a delay of 1 second after the beginning of the experiment, as to allow the convergence of the focus field on the desired input. Notice that movement is introduced to make the experiment valid. Indeed, staticity is the default in Equation (1), as the feedback at \mathbf{x} and time $t + dt$ comes from the same field position \mathbf{x} at time t . The speed of the stimulus is thus relatively high in order to force the model dynamics to be reactive.

Finally, scenario [E] checks if the model is robust to random noise on the entire field. The bubble is therefore highly distorted as shown on Figure 3. The noise is added from the start, as to ensure the selection of parameters allowing the convergence of the CNFT model under such conditions, and not simply its resilience to noise. When interacting with real environments, noise is constantly present and must be taken into account at any time. It would therefore be highly artificial to take its absence for granted during the first iterations. In order to show the qualitative differences in results obtained when removing this constraint, let [E'] be a new scenario identical to [E] except for the introduction of noise at $t = 1$ instead of $t = 0$. Results using different combinations of scenarios are detailed and discussed in the following section.

III. RESULTS

The GA described in the previous section has been applied to a population of 20 to 100 individuals. Their number did not quantitatively change the results as long as enough generations were produced to allow the convergence of the genome (>20 generations) and that mutation had a sufficient effect on the parameters ($m = 0.1$, $p_{mut} = 0.1$). The parameters were randomly chosen with a uniform distribution in the following ranges: $A \in [0.1, 2]$, $K \in [0.1, 1]$, $b \in [0.01, 2]$, $k \in [0.1, 1]$, $\tau \in [0.1, 2]$. Some of these boundaries were selected to avoid divisions by zero, the others to get rid of trivial or absurd dynamics. For each new generation, the 40% top parents were kept unchanged (the ones with the highest fitness), while 60%

of the individuals were randomly selected as parents and randomly combined to produce the children.

The discrete implementation and equations used for the computer simulations were inspired by Rougier and Vitay [11]. A singular value decomposition (SVD) of the lateral connection kernel leads to the decomposition of the 2D convolution into two orthogonal 1D convolutions. The resulting drop in complexity allows a ten fold speed-up on 50×50 maps. On an old Intel Pentium M 735 based laptop configuration, each update of the CNFT model takes approximately 37ms, each simulation therefore lasting for 3.7 seconds on average with $dt = 0.1$ and a time limit of 10 seconds. The optimization of the parameters in 20 generations using the GA on 20 individuals then takes about 74 minutes for one set of 3 scenarios and one fitness definition.

As stated previously, GA run on different sets of scenarios lead to different optimal parameters. To also show the importance of the expression used for the fitness function, results are given whether the *shape* variable is used. Optimal sets of parameters after convergence of the GA are then used to test the associated model on all four scenarios [C], [D], [E] and [E']. The results are synthesized in Table II and discussed below.

A. Discussion

When looking at Table II, one can observe the large variations in optimal parameters and performance on the different scenarios (several orders of magnitude). The parameters obtained when using neither [E] nor the *shape* variable are the most efficient in terms of reactivity, convergence speed and mean error, except for [E] where the model never converges (mean *error* = 0.27). These measures are however correlated as a low value for τ and high values for A and B lead to large variations of the potential when no bubble is present to inhibit the rest of the field. This initial chaotic behavior, not suitable when noise is introduced at the beginning of the simulation, is illustrated on Figure 4. Thus, although the GA produces a lower error on several scenarios when compared to the hand tuned parameters proposed in [11], the optimal solution

does not reflect all the constraints an expert might take into account.

The *shape* variable introduces such an additional constraint (bubble stereotyping), whose effect can be seen on Figure 5. The increased spatial coherence of the bubble requires an excitatory core on the lateral connection kernel, therefore resulting in changes in all its parameters. This leads to a more "Mexican hat like" lateral weight function compared to the mainly inhibitory function of the previous set of GA scenarios. The dynamics is then smoother and the system is still quite reactive to input changes, but still not able to cope with noise introduced from start (mean error = 0.26).

When using scenario **E** in the GA, τ jumps from 0.29/0.27 to 0.45 and is accompanied by a proportional change of A and B to compensate the slower temporal integration. The results well display the tradeoff between robustness and reactivity, as to follow rapidly moving inputs, the model needs to take into account any stimulations,

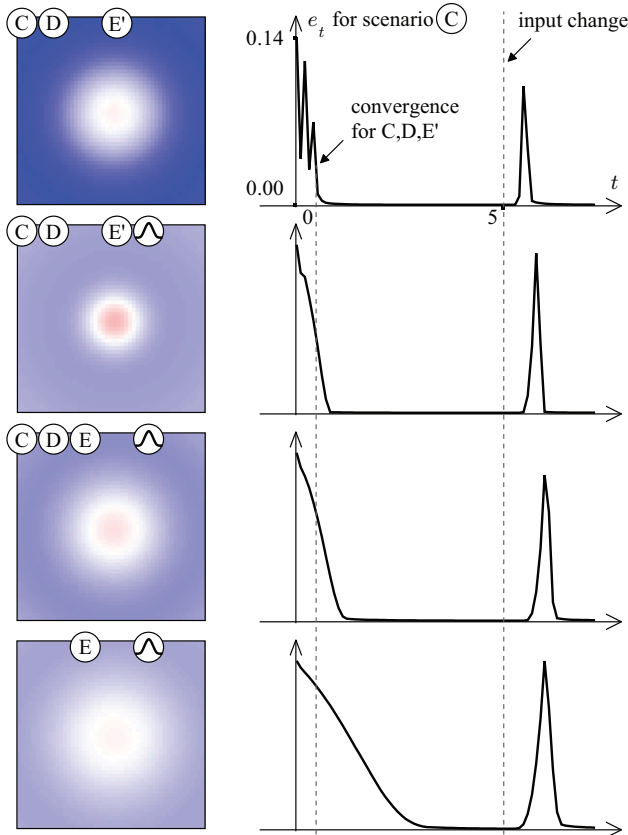


Fig. 4. Comparison of the lateral connection kernels and error dynamics on scenario **C**. This scenario is the only one without a random component in it, allowing reproducible deterministic dynamics without using statistics over a large number of trials. Each row represents the results on a different set of scenarios and variables used to evolve the parameter vector population. When using the *shape* variable and scenario **E**, the model loses in reactivity but gains in smoothness and stability.

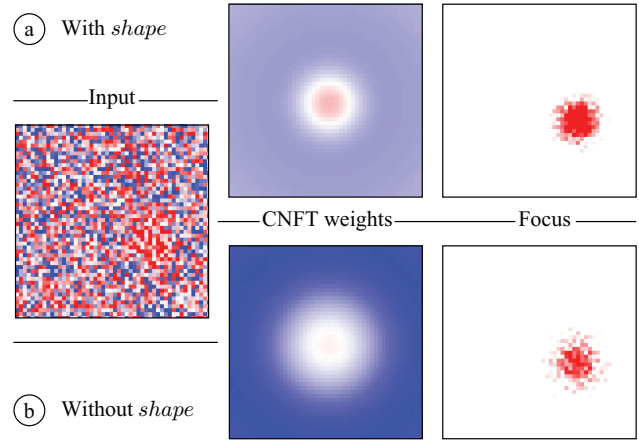


Fig. 5. Qualitative difference in the optimal solution and CNFT lateral kernel relatively to the use of *shape* in the fitness evaluation. Although the stimulus is hard to distinguish on the noisy input (of standard deviation 1.0), both sets of optimal parameters found by the GA accurately track it. The GA however converge to different parameter values due to the tradeoff between fast convergence (*conv*) and resilience/lack of distortions (*shape*). **(a)** The top row screenshots show the CNFT weights and focus field obtained when integrating the *shape* variable. The excitatory core produces a well-formed bubble by reinforcing its spatial coherence. **(b)** The bottom row shows the same elements without the use of *shape* in the fitness.

including noise and distracters. Setting τ thus constrains the ability to switch between targets and the amount of spatiotemporally coherent stimulations required to destabilize the bubble.

At this point, a last parameters optimization is performed to check if the *shape* variable and scenario **E** would not be sufficient to account for all the CNFT emergent properties. Although still functional in most cases, the GA converges to a different optimal solution, where τ takes a high value of 0.64. The model is of course optimal on scenarios **E** and **E'**, but is slow to switch between targets or simply to converge on one. Whatever the set of scenarios chosen, the results are however quite robust, as correlated variations up to 20% in the parameters values relatively to the optimal result only lead to mean error fluctuations lower than 0.03 (the results are obtained using the best parameter sets found for the last generations). This has to be contrasted with random changes of the same amplitude which can produce non converging simulations.

As a concluding remark for this section, whatever values the parameters may take in Equation (1), the error cannot converge to zero with moving stimuli. This is true except if there is only one symmetric stimulus, no noise and no distracter, in which case the degenerated model with $\tau=dt$, $A=B=0$ will no longer hold any CNFT property as $u(\mathbf{x}, t) = s(\mathbf{x}, t)$. This explains the efficiency of the model on scenario **C**, where the staticity of the competitive term in the equation (the lateral connectivity kernel for position \mathbf{x} being centered on \mathbf{x}) is reflected by the staticity

of the input. As a consequence, the model is unable to track a stimulus passing over a static distracter, as staticity is the default. Whether we include such a scenario in the evaluation of the fitness function or simply test the optimal solution with it, no set of parameter values can succeed as the lack of anticipation is inherent to the equation.

IV. CONCLUSIONS

The optimization of the Continuous Neural Field Theory parameters using Genetic Algorithms produces improved results compared to hand tuned values. Even with expertise and intuition, the effect of parameter changes are hard to predict as the system involves a large number of interconnected units, feedback and non linear dynamics. All these ingredients of deterministic chaos indeed generate complex dynamics prior to the emergence of a bubble. In addition to the robustness of the model emergent properties, this optimization process improves the understanding of the interactions between the parameters. The analysis of the results and inclusion of new scenarios also points to the limits of the considered equation and tradeoffs between the reactivity, robustness and generality of the model.

This not only opens the possibility to change the equation when limits are encountered, but the automatic exploration of the parameter space also allows a fair comparison of different implementations, as the number of parameters and their optimal values might greatly differ. Such extensions to classical Dynamic Neural Field equations include the introduction of delayed inhibition [2], the use of spiking units rather than rate coding [6] or the introduction of a predictive term (to address the limitation presented at the end of the discussion). Other differential equations governing the neural fields have been proposed [7] and the same metaheuristic applies. The method may also be useful to compare different implementations of the same equations, for instance approximating the maps with sparse representations for increased performance.

Experimental simulations are required as no mathematical framework is yet able to formally characterize the dynamics of an arbitrary DNF equation on arbitrary scenarios. Such an exploration is thus a way to determine if the behavioral differences observed during simulations come from side effects, badly tuned parameters or fundamental functional differences. Others have already acknowledged the need to optimize the parameters [1] and in this perspective, we presented a low cost alternative to brute force approaches. For these later ones, coarsely sampling the 5-dimensional space explored in this paper ensures completeness, but already requires many times more simulations than Genetic Algorithms.

REFERENCES

- [1] L. Alecu and H. Frezza-Buet, "An empirical evaluation framework for qualifying dynamic neural fields," in *Deuxième conférence française de Neurosciences Computationnelles*, 2008.
- [2] —, "A dynamic neural field mechanism for self-organization," in *BMC Neuroscience 2009, 10(Suppl 1)*, 2009, p. 273.
- [3] 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.
- [4] S.-I. Amari, "Dynamics of pattern formation in lateral-inhibition type neural fields," *Biological Cybernetics*, vol. 27, no. 2, pp. 77–87, 1977.
- [5] Y. Burnod, *An adaptive neural network: The cerebral cortex*. Masson, 1989.
- [6] S. Chevallier and P. Tarroux, "Visual focus with spiking neurons," in *European Symposium on Artificial Neural Networks*, 2008, pp. 385–389.
- [7] S. E. Folias and P. C. Bressloff, "Breathers in two-dimensional neural media," *Physical Review Letters*.
- [8] L. C. Katz and E. M. Callaway, "Development of local circuits in mammalian visual cortex," *Annual Review of Neurosciences*.
- [9] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biological Cybernetics*.
- [10] C. v. d. Malsburg, "Self-organization of orientation-sensitive cells in the striate cortex," *Kybernetics*.
- [11] N. P. Rougier and J. Vitay, "Emergence of attention within a neural population," *Neural Netw.*, vol. 19, no. 5, pp. 573–581, 2006.
- [12] J. Taylor, "Neural bubble dynamics in two dimensions: Foundations," *Biological Cybernetics*, vol. 80, pp. 5167–5174, 1999.