# Structure of receptive fields in a computational model of area 3b of primary sensory cortex

Georgios Is Detorakis, Nicolas P. Rougier

# Structure of Receptive Fields in a Computational Model of Area 3b of Primary Sensory Cortex

Georgios Is. Detorakis[1] and Nicolas P. Rougier[2,3,4,*]

[1]Laboratoire des signaux et systèmes, Supélec, Gif-sur-Yvette, France
[2]INRIA Bordeaux Sud-Ouest, Bordeaux, France
[3]LaBRI, Université de Bordeaux, Institut Polytechnique de Bordeaux, Centre National de la Recherche Scientifique, UMR 5800, Talence, France
[4]Institut des Maladies Neurodégénératives, Université de Bordeaux, Centre National de la Recherche Scientifique, UMR 5293, Bordeaux, France
[*]Corresponding author: Nicolas.Rougier@inria.fr

**Abstract**

In a previous work, we introduced a computational model of area 3b which is built upon the neural field theory and receives input from a simplified model of the index distal finger pad populated by a random set of touch receptors (Merkell cells). This model has been shown to be able to self-organize following the random stimulation of the finger pad model and to cope, to some extent, with cortical or skin lesions. The main hypothesis of the model is that learning of skin representations occurs at the thalamo-cortical level while cortico-cortical connections serve a stereotyped competition mechanism that shapes the receptive fields. To further assess this hypothesis and the validity of the model, we reproduced in this article the exact experimental protocol of DiCarlo et al. that has been used to examine the structure of receptive fields in area 3b of the primary somatosensory cortex. Using the same analysis toolset, the model yields consistent results, having most of the receptive fields to contain a single region of excitation and one to several regions of inhibition. We further proceeded our study using a dynamic competition that deeply influences the formation of the receptive fields. We hypothesized this dynamic competition to correspond to some form of somatosensory attention that may help to precisely shape the receptive fields. To test this hypothesis, we designed a protocol where an arbitrary region of interest is delineated on the index distal finger pad and we either (1) instructed explicitly the model to attend to this region (simulating an attentional signal) (2) preferentially trained the model on this region or (3) combined the two aforementioned protocols simultaneously. Results tend to confirm that dynamic competition leads to shrunken receptive fields and its joint interaction with intensive training promotes a massive receptive fields migration and shrinkage.

**Keywords:** Receptive Field — Neural Field — Somatosensory Cortex — Area 3b — SI — Computational Model — Self-Organization

# Contents

# Introduction

In a previous work (**Detorakis and Rougier** (2012)), we proposed a computational model of the somatosensory cortex based on neural field theory (**Amari** (1977); **Bressloff** (2011)). This model allowed us to investigate formation and maintenance of ordered topographic maps in the primary somatosensory cortex during the critical period of development (postnatal), where representations are shaped, and the post-critical period, where representations are maintained and possibly reorganized in face of cortical or sensory lesions or dynamic changes of the environment. The main hypothesis of the model is that feed-forward thalamo-cortical connections are an adequate site of plasticity while cortico-cortical connections drive a competitive mechanism that is central in the learning process. The model relies functionally on the balance between lateral excitation and inhibition, allowing to widen or sharpen the response of the model and plays a critical role in the shaping of the receptive fields during development. This modulation of the balance may originate from at least two distinct processes at two different time scales. In the long-term, neurogenesis/neuronal death and synaptogenesis/synaptic degeneration **Edelman** (1987) are ontogenetic factors that shape cortical connectivity during development as explained in **Bressler and Tognoli** (2006). Synaptic density spikes during the childhood followed by a decline during adolescence and adulthood **Feinberg et al.** (1990).

To further support this hypothesis, we first reproduced in this article the experimental protocol of **DiCarlo et al.** (1998) that has been used to characterize the structure of receptive fields (RFs) in area 3b of primary somatosensory cortex in three alert monkeys. This protocol is based on the passive stimulation of the distal finger pad using a rotating drum. This allowed the authors to show that most RFs contain a single, central region of excitation and one or more regions of inhibition. In this work, we adapted this protocol to our model and validated our results using the same modified linear regression algorithm to characterize excitatory and inhibitory components of each RF. This helped us to tune the model and we found very consistent results using a stereotyped profile for lateral connections, resulting from a fixed balance between the amount of excitation and inhibition.

We further processed our analysis by considering the dynamic modulation of the competition following a top-down signal that is supposed to originate from higher order cortical areas and has been implemented as a gain multiplication at the level of the lateral intra-cortical connections. In the short-term perspective, such modulation allows the model to give a sharper and stronger response to any stimulus. In the long term perspective, the repeated modulation of the response has a long-lasting influence onto the structure of the RFs. We hypothesized such a modulation to represent a form of somatosensory attention (spatial attention) because such modulation has been already proposed in the visual dimension as a possible mechanism for spatial attention, more specifically in area V4 **Salinas et al.** (1997); **Salinas and Sejnowski** (2001). Indeed, attention has been mostly studied in the visual system and can be defined as a mechanism that enhances the processing of interesting (understood as behaviorally relevant) locations (spatial or featural) while darkening the rest **Posner** (1980); **Treisman** (1988). The first neural correlate of that phenomenon has been discovered by **Moran and Desimone** (1985) in V4 where neurons respond preferentially for a given feature in their receptive fields. Since then, attentional effects have been found in each map of the ventral stream but also in the dorsal stream (area MT encoding for stimulus movement, LIP representing stimuli in a head-centered reference frame). Such attentional effects have also been identified in other modalities as well: auditory **Fritz et al.** (2007); **Picton and Hillyard** (1974), motor **D.A. Norman** (1980) and somatosensory to a much lesser extent, **Hsiao et al.** (1993). In fact, even if the somatosensory system has been extensively studied in monkeys and rats, the nature of attentional mechanisms and how they may affect neocortical maps of somatosensory cortices remain largely unknown.

Our main hypothesis is that the modulation of a response in area 3b may be one of the core mechanism, even though the origin of the modulation signal is not detailed in this article. To test this hypothesis, we developed a specific protocol where modulation occurs only if a presented stimulus is located within a region of interest (RoI) that corresponds to the attended region and we compared results with a protocol where the region of interest is specifically trained. Results tends to highlight a prominent role of the modulation into

the shrinkage of the RFs even if only the joint interaction of training and attention lead to maximal effects.

# Material & Methods

## Model

### Finger pad

We modeled a skin patch of the index distal finger pad where Merkel's ending complex (MEC) density is known to be the highest and to convey information about touch and pressure, **Pare et al.** (2002). These receptors have been shown to have a sustained response to any mechanical deflection of the skin tissue. We thus considered a set of 256 receptors uniformly spread over the skin patch. When a stimulus is applied at a given position $\mathbf{z}$ of the skin patch, its mechanic property extends the pressure level to nearby locations (**Goodwin et al.**, 1995). More formally, the response $s_i$ of any receptor $i$ located at $\mathbf{r_i}$ is given by the following equation:

$$s_i(\mathbf{z}) = \exp(-\frac{1}{2}\|\mathbf{z} - \mathbf{r_i}\|) \tag{1}$$

It is apparent that when a stimulus is present and its distance from the corresponding receptor tends to zero, the activity is the highest possible. On the contrary, when there is no stimulus present, the activity is zero. This model assumes a very simple correlation between the distance of the receptor to the stimulus center and its level of activity. We chose such a simple model because it eases the mathematical analysis of the model and we are not interested in the full modeling of the finger pad. More accurate models can be found in **Srinivasan** (1989) (waterbed model), **Dandekar et al.** (2003) (finite elements) and in **Sripati et al.** (2006b) (continuum mechanics) but we do not think using these models would fundamentally change the properties of our model (see figure 1 for a comparison of the waterbed and Gaussian surface deflection models) since the set of 256 receptors encode a two-dimensional quantity that corresponds to the position of the stimulus.

### Dorsal pathway

The dorsal column-medial lemniscus (DCML) pathway is the major afferent pathway for mechanosensory information and mediate tactile discrimination as well as proprioception (**Purves et al.**, 2001). There exist several relays along this path (dorsal root ganglion, gracile and cuneate nuclei of caudal medulla and ventral posterior lateral nucleus of the thalamus) that convey information from first order neurons up to the somatosensory cortex. We modeled this complex pathway as a direct transformation of the MEC activity corresponding to the mean distance between receptors activity and the corresponding feed-forward weights. Consequently, and considering a stimulus at position $\mathbf{z}$ on the skin patch, the input $I(\mathbf{x}, \mathbf{z}, t)$ received by a neuron $\mathbf{x}$ of SI is given by equation:

$$I(\mathbf{x}, \mathbf{z}, t) = 1 - \frac{1}{n}\sum_{i=0}^{n} |s_i(\mathbf{z}) - w_f^i(\mathbf{x}, t)| \tag{2}$$

where $i$ designates a specific skin receptor and $w_f^i(\mathbf{x}, t)$ is the feed-forward weight at time $t$ linking receptor $i$ to neuron $\mathbf{x}$. This equation implies that any SI neuron receives input from all the skin receptors. From a neurophysiological point of view, such an assumption is valid to the extent that we considered only a small skin patch on distal finger pad. The transformation itself can be considered as the complement of the normalized distance between the set of receptors and the set of feed-forward weights. Such transformation is maximal ($I(\mathbf{x}, \mathbf{z}) = 1$) for a given stimulus $\mathbf{z}$ if $\forall i, s_i(\mathbf{z}) = w_f^i(\mathbf{x}, t)$. This is true because equation (1) implies that the maximum amplitude of a stimulus is equal to one and we assumed that the feed-forward weights, $w_f$, are bound between 0 and 1 and therefore the maximal value of $I(\mathbf{x}, \mathbf{z}) = 1$ and the minimum value can be $I(\mathbf{x}, \mathbf{z}) = 0$.

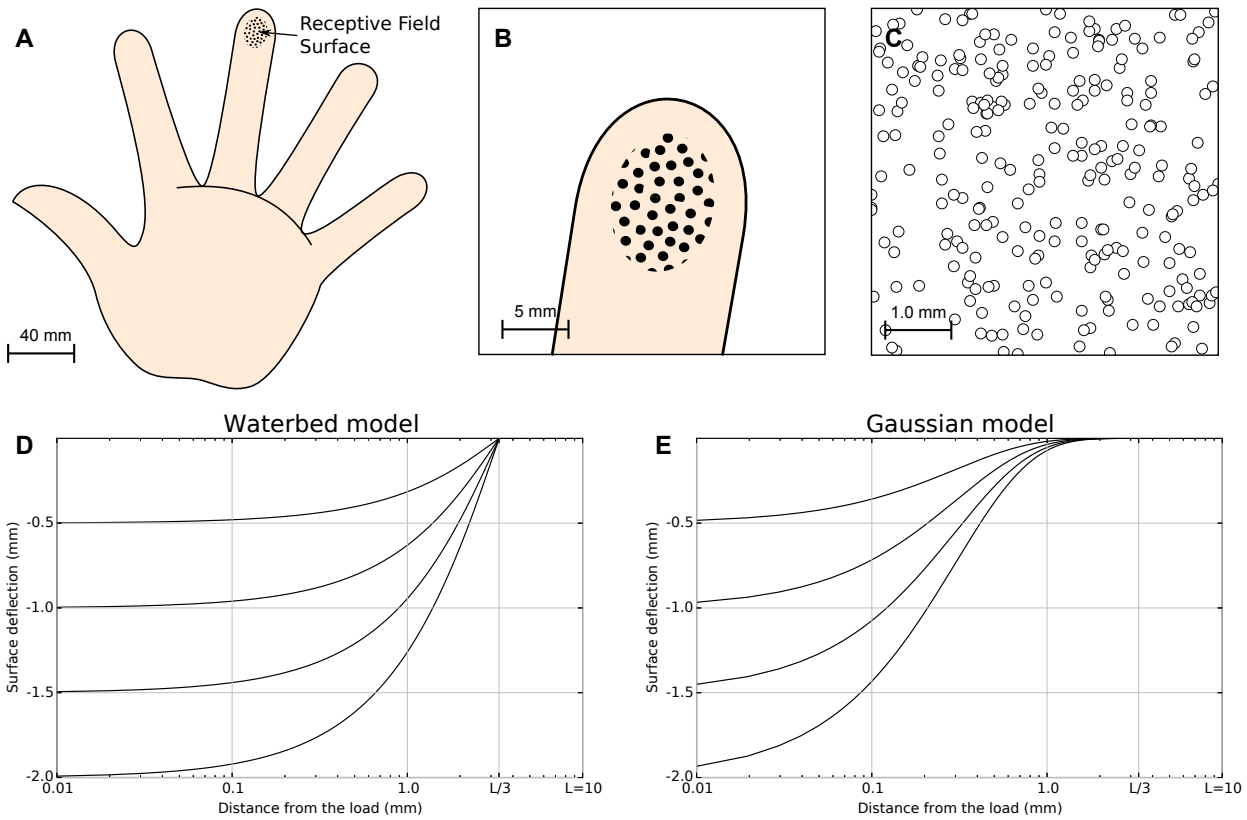


Figure 1: **Skin model.** The finger pad skin patch is approximately of size 25mm$^2$, using a receptor density of 10/mm$^2$. It has been modeled as a planar surface and we considered 256 MEC's that are arranged in a regular grid over the whole surface with a position jitter of 5%. This results in a quasi-uniform distribution consistent with actual distribution of MEC as reported in **Pare et al.** (2002). **A.** Schematic diagram of the hand. **B.** Position and relative size of the skin patch. **C.** Magnification of the skin patch showing MECs distribution. **D.** Waterbed surface deflection model from **Srinivasan** (1989). **E.** Gaussian surface deflection model from **Detorakis and Rougier** (2012). Each model predicts smaller deflection as a function of the distance from the load.

## Area 3b

Area 3b of the somatosensory cortex has been modeled using neural field theory (**Amari**, 1977; **Taylor**, 1999; **Wilson and Cowan**, 1973) which considers the cortex as a continuous surface $\Omega$. Considering a stimulus $\mathbf{z}$, the dynamic of the field is given by equation:

$$\tau \frac{\partial u(\mathbf{x},t)}{\partial t} = \underbrace{-u(\mathbf{x},t)}_{\text{decay term}} + \underbrace{\int_\Omega w_l(\mathbf{x},\mathbf{y}) f(u(\mathbf{y},t)) d\mathbf{y}}_{\text{lateral interaction}} + \underbrace{I(\mathbf{x},\mathbf{z},t)}_{\text{feed-forward input}} \tag{3}$$

where $u(\mathbf{x},t)$ is the membrane potential at position $\mathbf{x}$, $\tau$ is the membrane time constant, $f$ is the firing rate function, $w_l$ is the lateral connections function and $I(\mathbf{x},\mathbf{z},t)$ is the output from the DCLM pathway as defined in previous section. The dynamic of the field is tightly linked to the lateral connections function $w_l$ that defines the behavior of the field (traveling waves, spiral waves, bump solutions, see **Bressloff** (2011) for extensive review). In **Detorakis and Rougier** (2012), we defined $w_l$ as a difference of Gaussian functions such as to obtain bump solutions. More precisely, we assume $w_l$ is both isotropic and homogeneous (i.e. $w_l(\mathbf{x},\mathbf{y}) = w_l(|\mathbf{x}-\mathbf{y}|)$) and defined as

$$w_l(x) = w_e(x) - w_i(x) = \underbrace{K_e \exp\left(\frac{-x^2}{2\sigma_e^2}\right)}_{\text{excitatory part}} - \underbrace{K_i \exp\left(\frac{-x^2}{2\sigma_i^2}\right)}_{\text{inhibitory part}} \tag{4}$$

where $(K_e, \sigma_e)$ and $(K_i, \sigma_i)$ are constants that describe the extent and the strength of short-range excitation and long-range inhibition ($\sigma_i \gg \sigma_e$).

Learning occurs at the thalamo-cortical level using an Oja-like learning rule (proportional to a pre-synaptic measure multiplied by a post-synaptic quantity) which solves stability problems that is known to exist in the standard Hebbian learning rule (see **Oja** (1982)). It reads:

$$\frac{\partial w_f(\mathbf{x},t)}{\partial t} = \gamma \underbrace{(s(\mathbf{z}) - w_f(\mathbf{x},t))}_{\text{pre-synaptic term}} \underbrace{\int_\Omega w_e(|\mathbf{x}-\mathbf{y}|) f(u(\mathbf{y},t)) d\mathbf{y}}_{\text{post-synaptic term}} \tag{5}$$

where $\gamma$ is a constant learning rate. We showed in **Detorakis and Rougier** (2012) how this learning rule, coupled with the neural field, allow the model to self-organize and develop topological representations of the skin patch. All the details are given in **Detorakis and Rougier** (2012) but briefly, equation (3) allows the model to exhibit a single bump of activity (for any input) and the learning rule (5) exploits this bump solution to promote learning at position where the excitatory part of the lateral connections function is maximal. It is to be noted that because of the pre-synaptic term and the boundedness of receptors values (i.e. are bounded between 0 and 1), feed-forward weights are also bounded between 0 and 1.

## Gain modulation

As explained earlier, the shape of the bump solution of the neural field can be controlled via lateral connections function $w_l$. We have been using until now a stereotyped profile defined by the extent and the strength of short-range excitation $(K_e, \sigma_e)$ and long-range inhibition $(K_i, \sigma_i)$. This profile is used for the whole duration of the initial training protocol and has a direct influence on the self-organization process. We could have used instead a wider/weaker or thiner/stronger profile as shown in figure 3 but more importantly, we can also modify it *online*, provided a signal is sent to indicate which profile is to be used for processing the next stimulus. This is what we refer as the attentional signal, originating from higher cortical areas. More precisely, we can use two parameters sets , $(K_e', K_i')$ and $(K_e'', K_i'')$, and use the first set when no attentional signal is present and the second one, when an attentional signal is present.

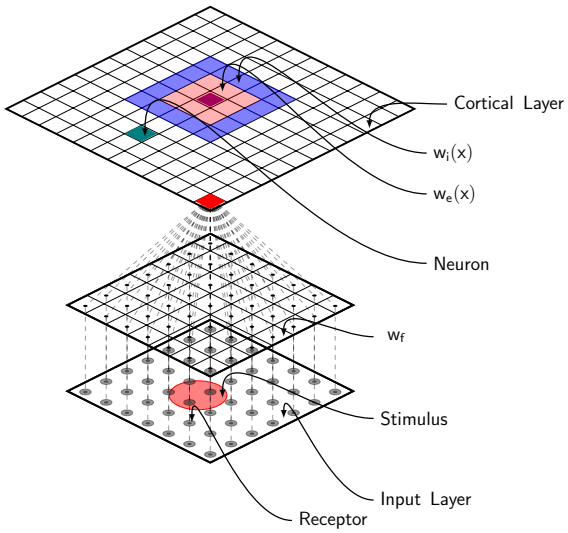


Figure 2: **Schematic of the full model.** Area 3b has been modeled using a neural field with lateral short-range excitation ($w_e$) and long-range inhibition ($w_i$). Each unit is fed with the information from all the 256 MEC receptors via feed-forward connections ($w_f$).

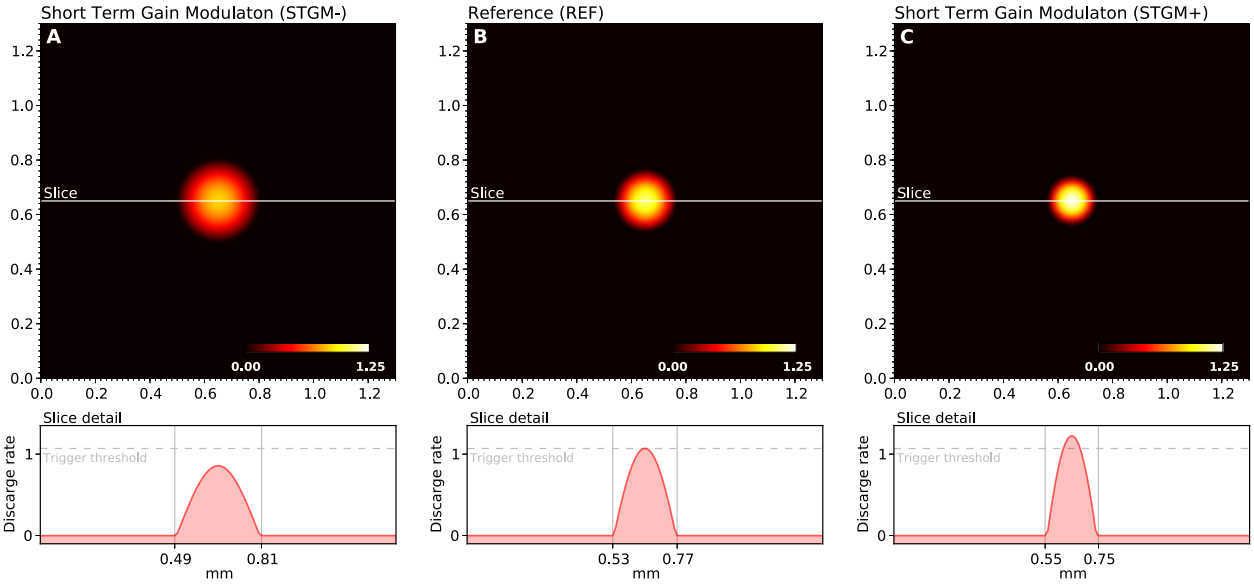


Figure 3: **Gain modulation.** The response of the model depends functionally on the balance between lateral excitation (gain $K_e$) and inhibition (gain $K_i$), allowing to widen (panel **A**) or sharpen (panel **C**) the peak of activity when a stimulus is presented. If we consider the trigger threshold to be the peak of nominal response (panel **B**), the same stimulus can either trigger a sharp response or not trigger any response at all, depending on the modulation. This modulation is considered in this work as a form of somatosensory attention.

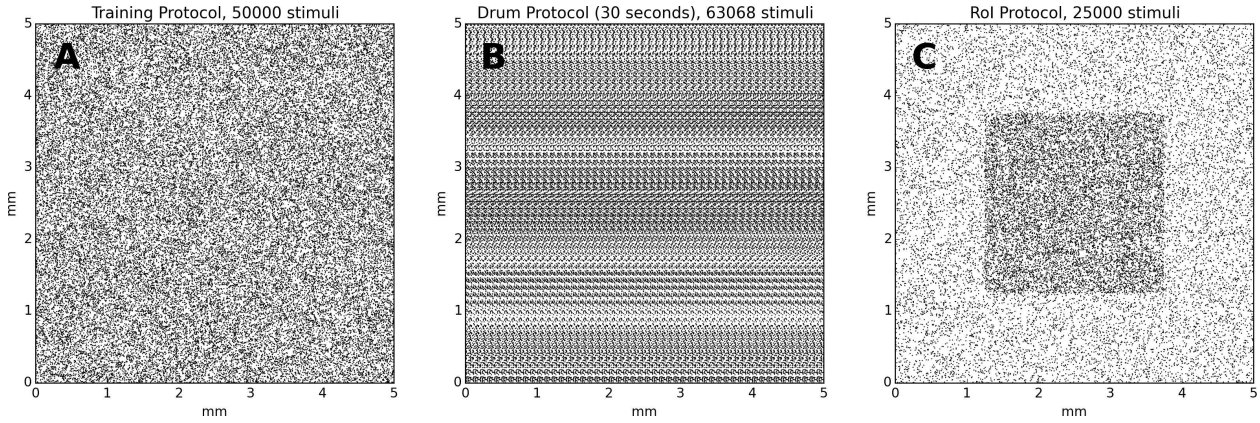


Figure 4: **Protocol stimuli sets. A** The training protocol set is made of 50000 stimuli distributed uniformly over the whole skin patch. At any moment, only one stimulus is presented to the model. **B** The drum protocol is based on a rotating drum made of 750 dots spread over the surface of the drum. The rotation of the drum makes stimuli to enter on the left side and exit on the right side of the skin patch, leading to temporal correlation between the different trials. At any moment, one to several stimuli can simultaneously stimulate the skin patch. **C** The RoI protocol, in the case of intensive training, is made of two sets of equal size (12500) for a total of 25000 stimuli. One set is made of stimuli exclusively located in the center of the skin patch and the other set is made of stimuli located outside this central region. This results in a higher (twofold) stimulus density in the central region. At any moment, only one stimulus is presented to the model.

# Protocols

## Initial training

Since the model initially possesses random weights, it is firstly necessary to train it in order to develop topological representations of the skin patch. We thus re-implemented the training protocol that has been used in **Detorakis and Rougier** (2012) and the training set is made of 50000 stimuli with random positions uniformly distributed over the whole skin patch. Each stimulus is presented once to the model and equations (3) and (5) are evaluated simultaneously until stability is achieved, i.e. there is no noticeable difference between $u(t)$ and $u(t + \delta t)$. The model is reset and another stimulus is picked up until there is no more available stimuli. A significant difference with the original model is the toric implementation of skin patch and cortical model as well. This means that any part of a stimulus that lay outside the skin patch reappears at the opposite side of the skin patch. The same holds true for the cortical sheet. This has been done to avoid any boundary effects that are known to exist in self-organization models. Once the training ends, the model has developed a topological representation of the skin patch such that two neighbor neurons on area 3b represent two neighbor location on the skin patch.

## Drum Protocol

The drum protocol is a direct adaptation of the protocol that has been used in **DiCarlo et al.** (1998). Authors used a cylindrical drum covered with a plastic sheet (28 mm $\times$ 250 mm) that possesses raised dots pattern (with a density of 10 dots per square centimeter for a total of 750 dots). The drum was mounted on a rotating drum stimulator and the orientation and the angular velocity of the drum were adjusted to produce proximal-to-distal stimulus movement at 40mm/sec across the skin surface. The drum completed 100 revolutions and the total time of simulation was 14 minutes and stepped a total distance of 20mm. We adapted the drum protocol as a planar surface of size 250mm$\times$30mm and moved the skin patch over the full length (40mm/second) before jumping back to the start and shifting up the patch by 200 $\mu$m. The drum

surface is made of 750 uniformly distributed dots, achieving a mean density of 10 dots/cm$^2$. Using a sample time step of 5ms, the model has been fed with 120000 samples for a complete sweep of the drum surface. Activity of all neurons are recorded at once without centering the drum onto each individual receptive field.

### RoI Protocol

We first defined an arbitrary region of interest (RoI) on the surface of skin patch whose size is one quarter of the total skin patch surface (see figure 4 panel C, the shaded squared area in the middle of the skin patch). For the intensive training session, we used a set of 25000 stimuli such that one out of two stimuli landed into the RoI (1 in / 1 out ratio, (see figure 4)). This means that the RoI, was twice more stimulated compared to the rest part of the skin patch. We presented each stimuli once to the model until no more stimuli were available. Learning occurs for the whole duration of the protocol. For the attentional experiment, we used 25000 uniformly spread stimuli over the whole skin patch. We presented each stimulus once to the model until no more stimuli were available. If a stimulus position was within the RoI, (i.e. the center of the stimulus, which is the most active zone of a stimulus) we explicitly instructed the model to attend to this stimulus by modifying the gain of the lateral connections ($K_e$ and $K_i$) as explained in the gain modulation section. Learning occurs for the whole duration of the protocol.

# Results

## Characterization of the RFs

We first report results concerning the characterization of RF structures observed in area 3b following the exact protocol of **DiCarlo et al.** (1998), that was used to investigate the two-dimensional structure of area 3b neuronal non-classical receptive fields (ncRFs) in three alert monkeys (non-classical receptive fields are defined in Appendix B). As explained earlier, this protocol has been slightly adapted to meet the constraints of the proposed computational model architecture. Following the initial training protocol were cortical representations have been shaped (see **Detorakis and Rougier** (2012)), we applied the drum protocol for a total of 120000 samples (that can have multiple sites of skin patch stimulation because of the raised dot patterns). From these data, we applied the exact same linear regression algorithm proposed and used by **DiCarlo et al.** (1998) for the characterization of the excitatory and inhibitory components of each ncRF. More precisely, for each matrix representing an non-classical receptive field, we first convolved it with a Gaussian filter ($\mu = 0$ and $\sigma = 1.7$) and applied a thresholding (10% of the absolute peak value) on every value. If a value was below the threshold, it was set to zero. We let each pixel of the non-classical receptive field to have at least two of the four neighbors non-zero and of the same sign such that isolated islands of positive or negative values were not allowed if they had a total area less than $0.7mm^2$. Each time we computed a ncRF, we also computed the signal-to-noise ratio (SNR) as well as the noise index, in order to constraint them to low values (see appendix B). After this preprocessing stage, we measured the respective size of excitatory (positive) and inhibitory (negative) areas. The minimum and maximum values of excitatory ncRFs were 9.12mm$^2$, 25.92mm$^2$, respectively for a mean size of $14.14^2$. The minimum and maximum values of inhibitory ncRFs were 5.92mm$^2$, 26.56mm$^2$, respectively for a mean size of $14.4^2$. Figure 5 shows the bivariate plot of excitatory versus inhibitory area (similar results have been found by **DiCarlo et al.** (1998)). Furthermore, a k-means classification of the ncRFs was performed on the ncRFs in order to compare the number of ncRFs classes from the model with the number of classes in **DiCarlo et al.** (1998). The k-means classification separated 16 different classes accodring to the topology of the excitatory and inhibitory areas (homogeneity = 0.39, completeness = 1.0, V-measure = 0.56). We found non-classical receptive fields whose excitatory area was surrounded by the inhibitory one as well as non-classical receptive fields whose excitatory area was facing the inhibitory area (see figure 5). It is to be noted that figure 5 shows a remarkable similarity with physiological results of DiCarlo et al., where most of the ncRFs are centered around a central point of 15mm$^2$ (excitatory) / 15mm$^2$ (inhibitory). The spread is larger in the case of DiCarlo but this was expected since we used a toric stimulation.

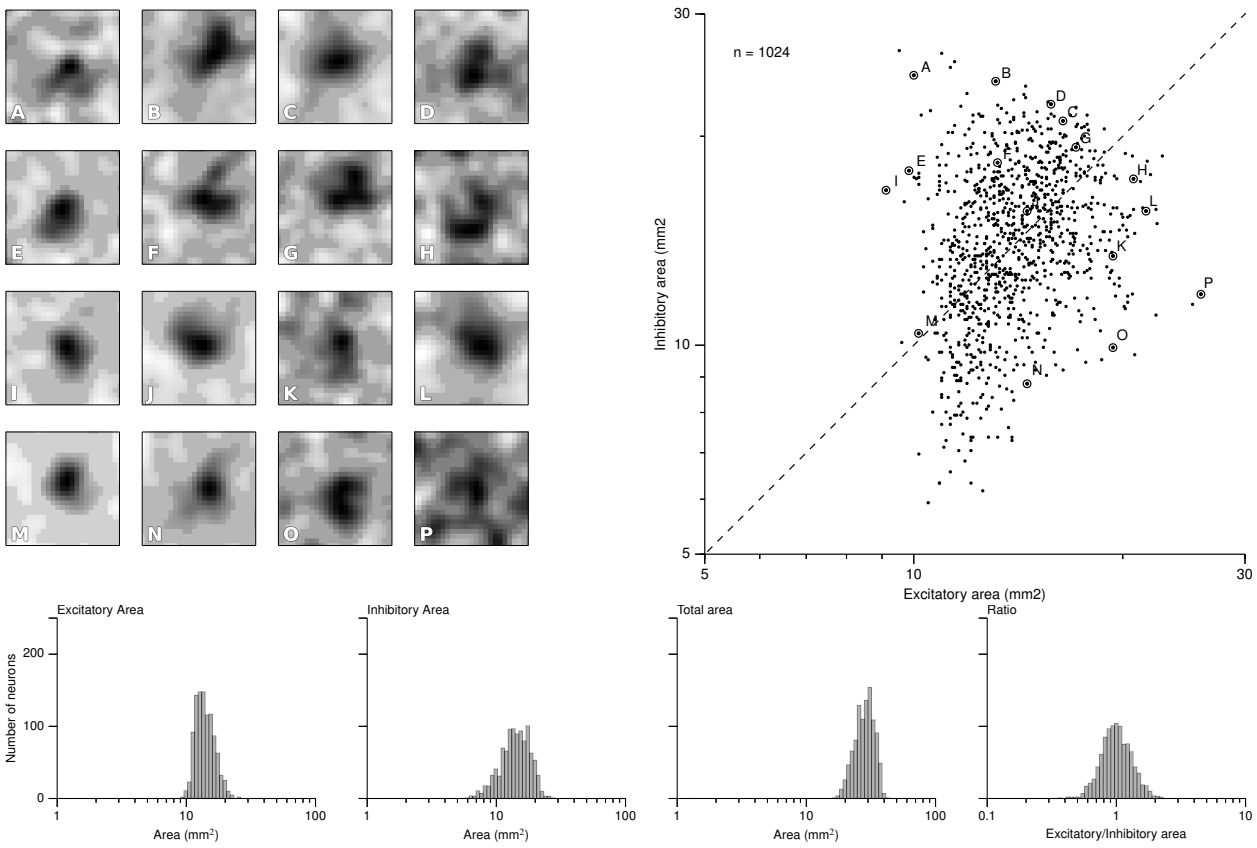


Figure 5: **Characterization of the ncRFs.** From the experimental drum protocol of **DiCarlo et al.** (1998), we recorded 120000 responses for each of the 1024 neurons of the model and we subsequently applied the same analysis with **DiCarlo et al.** (1998) to obtain the respective ncRF. The scatter plot on the *right* displays the balance between excitatory and inhibitory components of each ncRF. Excitatory area was measured as the total positive area in the thresholded ncRF (positive ncRF regions with values $\geq 10\%$ of the peak absolute ncRF value, see Materials and Methods and Appendix B). Inhibitory area was measured as the total negative-thresholded ncRF area (negative ncRF regions with absolute values $\leq 10\%$ of the peak absolute ncRF value). The left part of the figure illustrates the diversity of ncRFs and the letter corresponds to a point in the scatter plot. The bottom row shows the distributions of the sizes of ncRFs. The y-axis indicate the number of neurons (n=1024) and the x-axis, from left to right displays the excitatory area of ncRFs, the inhibitory area, the total area (is the sum of the excitatory and inhibitory areas) and the ratio of excitatory area to inhibitory one in logarithmic scale.

### Training the RoI

During the specific training of the RoI, we considered a set of 25000 stimuli, half of them being located in the RoI. We will later refer to this as the intensive protocol. At the end of the protocol, we measured the location and the size of the classical receptive fields (cRF) or simply receptive fields (RF) (see appendix B for details) and compared them to the control setup, that corresponds to the end of the nominal training period (or the start of this protocol). Figure 6 (panel B) reveals a strong migration of most RF towards the RoI with an overall final density being higher in the center of the RoI. We also measured RFs size at the end of the protocol and compared them with control. The control setup shows a normal distribution of sizes around a central value ($2.1\text{mm}^2$ , SD=0.42) while the intensive training setup leads to a significant reduction of the RFs ($1.6\text{mm}^2$ , SD=0.48). Overall, there has been a significant decrease in the mean size of RFs (see figure 7, panel B'). Such results are consistent with **Xerri et al.** (1994) that shows that intensive training over a skin area can cause the corresponding cortical territory expansion with a simultaneous shrink of receptive fields of neurons of the somatosensory cortex.

### Modulating the RoI

In order to make the model to attend to the RoI, we considered a set of 25000 stimuli, uniformly spread over the whole skin patch and we instructed the model to attend to a stimulus if this was within the RoI, i.e. using different gains for the lateral connections. The major difference compared to the intensive training experiment is the non-migration of the RFs towards the center of the RoI as shown in figure 6, panel C. The distribution remains actually quasi-uniform and the RoI does not benefit from significant higher density. However, the sizes of the RFs have shrunk by 33%, leading to a mean size of $1.4\text{mm}^2$ (SD=0.37). This demonstrates that migration and shrinkage of RFs are actually two distinct processes that can be (partly) separated.

### Joint effect of training and modulation

For studying the joint effect of training and modulation, we mixed the two RoI protocols and considered both a non-uniform set of 25000 stimuli, half of them being located in the RoI and we instructed the model to attend to a stimulus if it was located in the RoI. The final density of RFs shown in figure 6D reveals a massive migration of the RFs towards the RoI with a simultaneous shrinkage in their sizes compared to the control conditions ($0.71\text{mm}^2$, SD=0.04). These results point out that the combined effects of intensive training and modulation actually sum up, leading to both a massive migration and a dramatic shrinkage of RFs, down to half the nominal size.

## Discussion

Using the model presented in **Detorakis and Rougier** (2012), we first validated it using the protocol and neurophysiological data from **DiCarlo et al.** (1998). We adapted the protocol to meet computational constraints and relevant recorded data. Results clearly indicate that the model is able to capture the main aspects of the original data recorded on three alert monkeys with most of ncRFs to contain a single region of excitation and one or more regions of inhibition located on one, two, three, or all four sides of the excitatory center. This is the first, to the best of our knowledge, computational model of area 3b that is able to replicate real neurophysiological data with such accuracy even though we used a very simple model for the distal finger pad and the dorsal column-medial lemniscus, as well. This tends to confirm that the thalamo-cortical feed-forward connections are an adequate site of plasticity while cortico-cortical connections drive the competition mechanism. Furthermore, even if the present study has been circumscribed to the spatial characteristics of the receptive fields, (**Sripati et al.**, 2006a) have shown spatio-temporal receptive fields (STRF) in area 3b tend to have early excitatory region followed by in-field (replacing) greater inhibition. Authors conclude that such *greater inhibition observed in cortical STRFs points to the existence of underlying intracortical mechanisms* that is very consistent with our own hypothesis. To go further in this direction, we would need

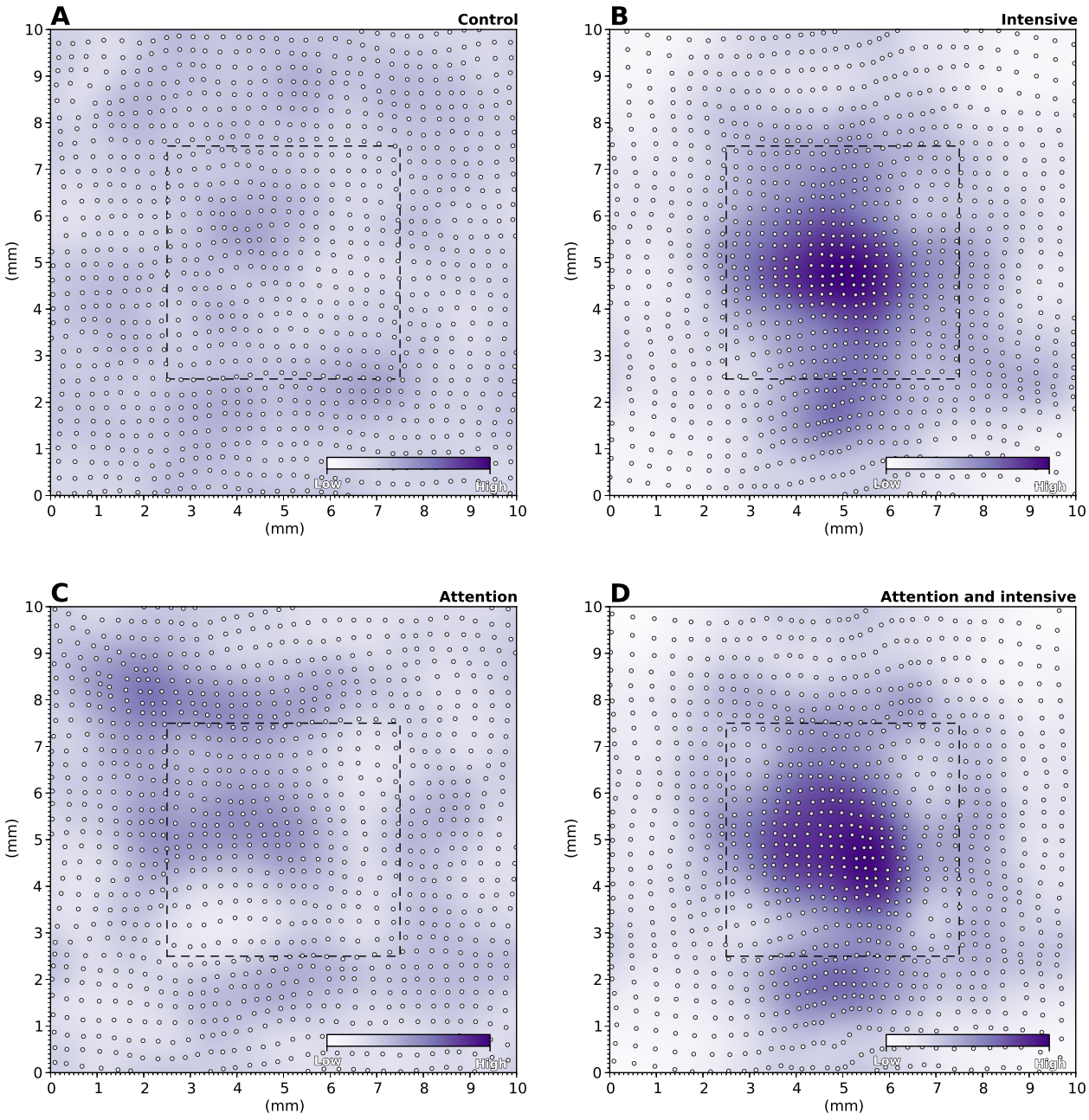


Figure 6: **RF Migrations. A.** The distribution of RFs over the skin patch is quasi-uniformly distributed for the control. **B.** Intensive training onto the RoI makes RFs to migrate towards the RoI leading to a higher density of RFs within the RoI. **C.** Explicitly attending the RoI modifies only marginally the distribution of RFs that tend to remain quasi-uniformly distributed over the whole skin patch. **D.** The joint effect of intensive training and attention leads to an even greater migration of RFs towards the RoI (compared to intensive training only).

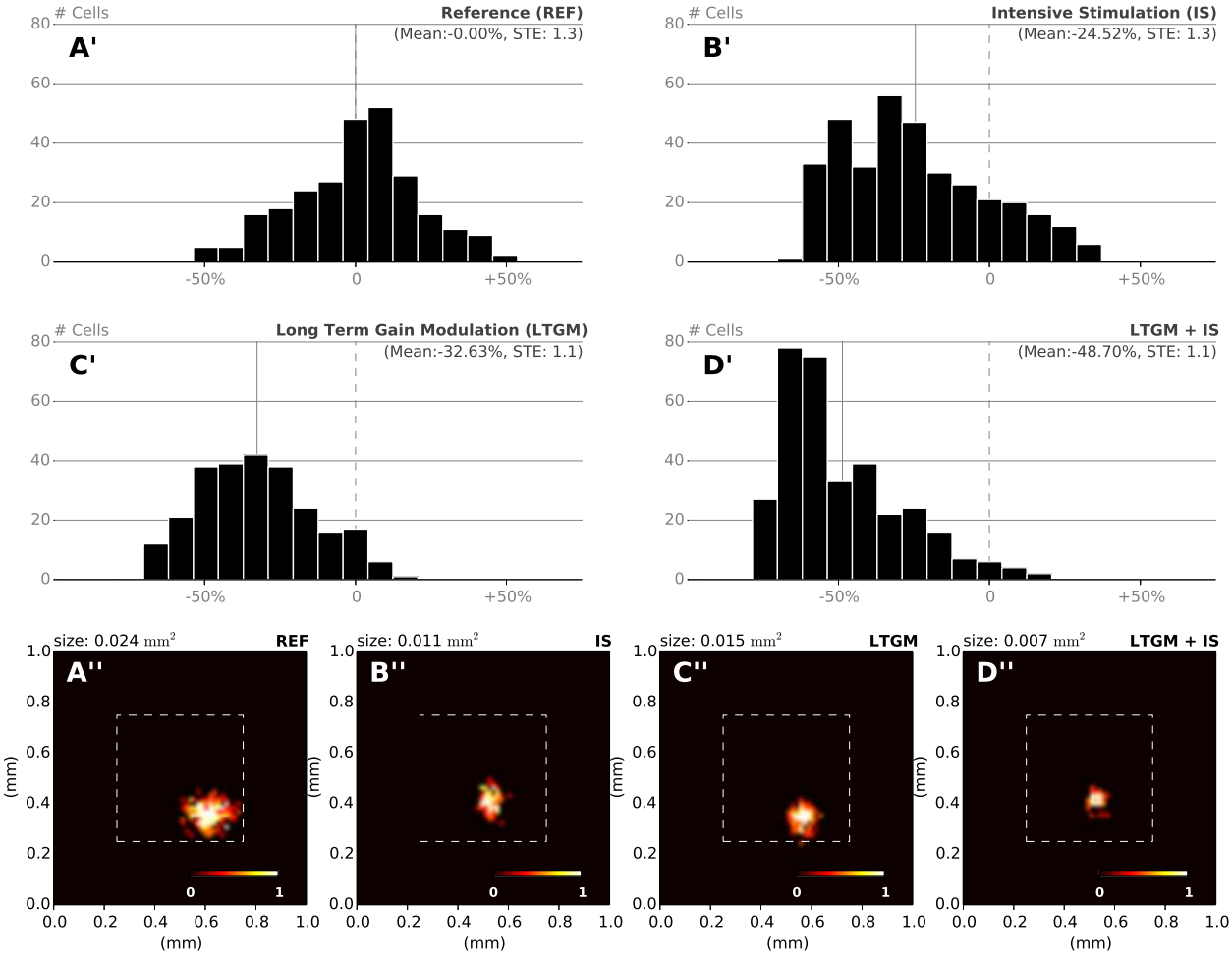


Figure 7: **RF Shrinkage. A'.** The relative histograms of RFs sizes after initial training (50000 samples) follows a normal distribution. **B'.** After model training, specifically in the RoI (with a 1/1 ratio) using 25000 extra samples, the mean RF size has been reduced by 25% compared to the nominal mean size. **C'.** By sharpening the model response when a stimulus is presented within the RoI (25000 samples), the mean RF size has been reduced by 33% compared to the nominal mean size. **D'.** The joint effect of training and modulation (25000 samples) leads to a dramatic shift in relative size of RF, with a mean size being half of the nominal mean size. **A" D".** Receptive field of a single cell recorded at the end of each of the aforementioned experiments. The receptive field size in the attentional/intensive condition (0.007 $mm^2$) has shrunk to one third of the control size (0.024 $mm^2$).

to consider finite transmission speed in cortico-cortical connections instead of instantenous connection **Hutt and Rougier** (2010).

We have also shown how this competition mechanism can be explicitly modulated by the modification of the gain at both the excitatory and inhibitory lateral connection levels. Such instructed modulation leads to receptive fields shrinkage in the region of interest while keeping intact the overall organization, with no noticeable migration of RFs. We identified such modulation as a form of spatial attention that is believed to be deployed selectively on this or that part of the body. Interestingly enough, these effects are known to occur in the visual system and a number of recent studies have identified such effects in area MT (**Anton-Erxleben et al.**, 2009; **Womelsdorf et al.**, 2008). More precisely, authors have shown how attention inside the cRF shrinks it, whereas directing attention next to the cRF expands it. Authors uses in their modeling work a bell shaped attentional signal while we have been using a constant attentional signal, modifying the gain for the whole population at once.

However, in the literature, the evidence for the effects of such spatial attention on SI are still contradictory. **Hsiao and Vega-Bermudez** (2001) has shown that attention is engaged in the modification of RFs in both primary and secondary somatosensory cortices and **Braun et al.** (2002) have confirmed such engagement of attention in the primary somatosensory cortex using neuroimaging techniques. However, **Godde et al.** (2000) claimed that attention is not critical in enhancing performance during a discrimination task even though, consecutive training or pairing stimulation (leading to co-activation) can affect the RFs. In the case of intensive training of the RoI, our model tends to suggest a large expansion of the cortical territory with a simultaneous shrinkage of the receptive fields as well as strong migration of their centers towards the RoI. These findings are still contradictory with psychophysical and neurophysiological studies such as **Godde et al.** (2000); **Pilz et al.** (2004); **Recanzone et al.** (1992), where authors noticed that the cortical representations undergo an expansion but at the same time RFs undergo a similar expansion. However, other neurophysiological and neuroimaging studies have shown that when cortical representations expand, RFs sizes seem to decrease **Elbert et al.** (1995); **Xerri et al.** (1994). These latter results are also consistent with early findings of **Sur et al.** (1980). They found, from neurophysiological recordings and mappings that the magnification factor of cortical representations is related to the size of RFs. More precisely, the magnification factor is proportional to the size of RFs (the smaller the RFs the larger the cortical representation). Our findings tend to confirm that cortical representations in the case of intensive stimulation increase their relative size with a simultaneous RFs shrinkage. These findings indicate that there are two distinct processes at work, namely modulation and training, that are believed to be present simultaneously in most cases, while there may exist a few cases where only one process is active. This may reconcile the aforementioned contradictory results. To confirm these findings, it would thus be necessary to setup new experiments where modulation and training needs to be carefully dissociated. This can be done, for example, by precisely controlling the amount of training received by a subject and by distracting the subject such as drifting attentional process away from the primary task.

Even if our model suggests a hypothesis on how somatosensory spatial attention may modify the processing of stimulus and promote reshaping of RFs, nothing has been said so far about the exact nature, the origin and the selectivity of such attentional signal. **Sarter et al.** (2005) have proposed a possible circuitry involving the basal forebrain corticopetal cholinergic system since it has been observed in several studies **Donoghue and Carroll** (1987); **Jimenez-Capdeville et al.** (1997) that the loss of cortical cholinergic system directly impacts attentional functions. Furthermore, **Juliano et al.** (1991) have shown that the cholinergic depletion prevents expansion of somatosensory topographic maps, suggesting that cholinergic neurotransmitters are critical in the structure of cortical representations. Similarly, **Rasmusson and Dykes** (1988); **Tremblay et al.** (1990a,b) proposed that a cholinergic signal is responsible for the gain modulation of neuronal populations and that the co-activation of basal forebrain and the somatosensory cortex by cutaneous stimulation lead to enhanced cortical activity. Overall, such a signal may originate from a complex network involving the insular cortex, the dorsolateral prefrontal cortex, the posterior parietal cortex, the ventromedial prefrontal cortex, the posterior cingulate cortex and the anterior cingulate cortex

as proposed by **Menon and Uddin** (2010). The main point is that the insular cortex acts as a switch between two different prefrontal networks leading to an attentional effect through saliency occurring in the anterior insular cortex. We can thus speculate that such a cholinergic signal may affect the gain of intracortical lateral connections and the explicit signal that has been used during the attended RoI protocol may originate from a frontal decision.

Finally, even though we hardly notice it in our everyday life, somatosensory attention plays a critical role in our perception of the outer world. For example, the contact of clothes on the skin can be largely unattended even though all body receptors are activated at once. This results from habituation and yet, it is still possible to concentrate on a specific part of the body to actually experience the contact. Such spatial selectivity is very similar to the concept of the spotlight of attention proposed by **Posner** (1980) in the eighties for the visual perception. At that time, authors were hypothesizing for the existence of a dedicated control mechanism even though this view was later challenged by the premotor theory of attention proposed by **Rizzolatti and Craighero** (1988). This later theory postulates instead that there is no need for two different mechanisms (attention and action) and has received support from several electrophysiological and brain imaging studies. However, how this theory can be adapted to somatosensory attention remains unclear. Our model cannot answer the question on the selectivity since we only used a broad and constant modulation of the model. This choice has been made because we consider a small part of SI cortex where exactly one bump of activity can exist anytime. If we were to consider a larger part of SI, where for example several digit representations would co-exist, we would need a selective attentional signal to be able to direct gain modulations to the relevant population involved in the representation of the RoI. This is quite a complex problem, since this would involve not only a sensory representation of several digits (sensory homunculus), but also a motor representation (motor homunculus) and visual information as well. This is far beyond the scope of the present work but we think this might allow for a better understanding of somatosensory attention.

# References


Amari, S. (1977), Dynamics of pattern formation in lateral-inhibition type neural fields, *Biological cybernetics*, 27, 2, 77–87

Anton-Erxleben, K., Stephan, V., and Treue, S. (2009), Attention reshapes center-surround receptive field structure in macaque cortical area mt, *Cerebral Cortex*, 19, 10, 2466–2478

Braun, C., Haug, M., Wiech, K., Birbaumer, N., Elbert, T., Roberts, L., et al. (2002), Functional organization of primary somatosensory cortex depends on the focus of attention, *Neuroimage*, 17, 3, 1451–1458

Bressler, S. L. and Tognoli, E. (2006), Operational principles of neurocognitive networks, *International Journal of Psychophysiology*, 60, 139–148

Bressloff, P. (2011), Spatiotemporal dynamics of continuum neural fields, *Journal of Physics A: Mathematical and Theoretical*, 45, 3, 033001

D.A. Norman, T. S. (1980), Consciousness and self-regulation: advances in research and theory (Kluwer / Plenum Publ), chapter Attention to action: Willed and automatic control of behavior, 1–18

Dandekar, K., Raju, B. I., and Srinivasan, M. A. (2003), 3-d finite-element models of human and monkey fingertips to investigate the mechanics of tactile sense, *Journal of biomechanical engineering*, 125, 5, 682–691

Detorakis, G. and Rougier, N. (2012), A neural field model of the somatosensory cortex: Formation, maintenance and reorganization of ordered topographic maps, *PloS one*, 7, 7, e40257

DiCarlo, J. J., Johnson, K. O., and Hsiao, S. S. (1998), Structure of receptive fields in area 3b of primary somatosensory cortex in the alert monkey., *Journal of neuroscience*, 18, 7, 2626–2645

Donoghue, J. and Carroll, K. (1987), Cholinergic modulation of sensory responses in rat primary somatic sensory cortex, *Brain Research*, 408, 367 – 371

Edelman, G. (1987), Neural Darwinism (Basic Books, New York)

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., Taub, E., et al. (1995), Increased cortical representation of the fingers of the left hand in string players, *Science*, 270, 5234, 305–307

Feinberg, I., Thode, H. J., and Chugani, J., H.T. an March (1990), Gamma distribution model describes maturational curves for delta wave amplitude, cortical metabolic rate and synaptic density, *Journal of Theoretical Biology*, 142, 2, 149–161

Fritz, J., Elhilali, M., David, S., and Shamma, S. (2007), Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in a1?, *Hearing research*, 229, 1-2, 186

Godde, B., Stauffenberg, B., Spengler, F., and Dinse, H. (2000), Tactile coactivation-induced changes in spatial discrimination performance, *The Journal of Neuroscience*, 20, 4, 1597–1604

Goodwin, A., Browning, A., and Wheat, H. (1995), Representation of curved surfaces in responses of mechanoreceptive afferent fibers innervating the monkeys fingerpad, *The Journal of Neuroscience*, 15, 1, 798–810

Hsiao, S. and Vega-Bermudez, F. (2001), 8 attention in the somatosensory system, in R. Nelson, ed., The somatosensory system: Deciphering the brain's own body image (CRC)

Hsiao, S. S., O'shaughnessy, D., and Johnson, K. O. (1993), Effects of selective attention on spatial form processing in monkey primary and secondary somatosensory cortex, *Journal of Neurophysiology*, 70, 1, 444–447

Hutt, A. and Rougier, N. P. (2010), Activity spread and breathers induced by finite transmission speeds in two-dimensional neural fields, *Physical Review E*, 82, 5, 055701

Jimenez-Capdeville, M., Dykes, R., and Myasnikov, A. (1997), Differential control of cortical activity by the basal forebrain in rats: a role for both cholinergic and inhibitory influences, *Journal of Computational Neuroscience*, 381, 53–67

Juliano, S., Ma, W., and Eslin, D. (1991), Cholinergic depletion prevents expansion of topographic maps in somatosensory cortex, *Proceedings of the National Academy of Sciences*, 88, 3, 780–784

Menon, V. and Uddin, L. (2010), Saliency, switching, attention and control: a network model of insula function, *Brain Structure and Function*, 214, 5, 655–667

Moran, J. and Desimone, R. (1985), Selective attention gates visual processing in the extrastriate cortex, *Science*, 229, 782–784

Nordlie, E., Gewaltig, M.-O., and Plesser, H. E. (2009), Towards reproducible descriptions of neuronal network models, *PLoS Comput Biol*, 5, 8, e1000456, doi:10.1371/journal.pcbi.1000456

Oja, E. (1982), Simplified neuron model as a principal component analyzer, *Journal of mathematical biology*, 15, 3, 267–273

Pare, M., Smith, A., and Rice, F. (2002), Distribution and terminal arborizations of cutaneous mechanoreceptors in the glabrous finger pads of the monkey, *The Journal of Comparative Neurology*, 445, 347–359

Phillips, J., Johnson, K., and Hsiao, S. (1988), Spatial pattern representation and transformation in monkey somatosensory cortex, *Proceedings of the National Academy of Sciences*, 85, 4, 1317–1321

Picton, T. and Hillyard, S. (1974), Human auditory evoked potentials. ii: Effects of attention, *Electroencephalography and Clinical Neurophysiology*, 36, 191–200

Pilz, K., Veit, R., Braun, C., and Godde, B. (2004), Effects of co-activation on cortical organization and discrimination performance, *Neuroreport*, 15, 17, 2669

Posner, M. I. (1980), Orienting of attention, *Quarterly Journal of Experimental Psychology*, 32, 3–25

Purves, D., Augustine, G. J., Fitzpatrick, D., Katz, L. C., LaMantia, A.-S., McNamara, J. O., et al. (2001), The Major Afferent Pathway for Mechanosensory Information: The Dorsal Column-Medial Lemniscus System (Sinauer Associates), chapter The somatic sensory system, 199 – 202

Rasmusson, D. and Dykes, R. (1988), Long-term enhancement of evoked potentials in cat somatosensory cortex produced by co-activation of the basal forebrain and cutaneous receptors, *Experimental Brain Research*, 70, 2, 276–286

Recanzone, G., Merzenich, M., Jenkins, W., Grajski, K., and Dinse, H. (1992), Topographic reorganization of the hand representation in cortical area 3b owl monkeys trained in a frequency-discrimination task, *Journal of Neurophysiology*, 67, 5, 1031–1056

Rizzolatti, G. and Craighero, L. (1988), Advances in psychological science: Vol.2. Biological and cognitive aspects (Sabourin, M. and Craik, F. and Robert, M.), chapter Spatial attention: Mechanisms and theories, 171 – 198

Salinas, E., Abbott, L., et al. (1997), Invariant visual responses from attentional gain fields, *Journal of Neurophysiology*, 77, 6, 3267–3272

Salinas, E. and Sejnowski, T. J. (2001), Book review: Gain modulation in the central nervous system: Where behavior, neurophysiology, and computation meet, *The Neuroscientist*, 7, 5, 430–440

Sarter, M., Hasselmo, M., Bruno, J., Givens, B., et al. (2005), Unraveling the attentional functions of cortical cholinergic inputs: interactions between signal-driven and cognitive modulation of signal detection, *Brain Research Reviews*, 48, 1, 98–111

Srinivasan, M. (1989), Surface deflection of primate fingertip under line load, *Journal of Biomechanics*, 22, 4, 343–349

Sripati, A., Yoshioka, T., Denchev, P., S.S., H., and Johnson, K. (2006a), Spatiotemporal receptive fields of peripheral afferents and cortical area 3b and 1 neurons in the primate somatosensory system, *The Journal of Neuroscience*, 26, 7, 2101–2114

Sripati, A. P., Bensmaia, S. J., and Johnson, K. O. (2006b), A continuum mechanical model of mechanoreceptive afferent responses to indented spatial patterns, *Journal of neurophysiology*, 95, 6, 3852–3864

Sur, M., Merzenich, M., Kaas, J., et al. (1980), Magnification, receptive-field area, and hypercolumn size in areas 3b and 1 of somatosensory cortex in owl monkeys, *J Neurophysiol*, 44, 2, 295–311

Taylor, J. G. (1999), Neural bubble dynamics in two dimensions: foundations, *Biological Cybernetics*, 80, 5167–5174

Treisman, A. (1988), Features and objects: The bartlett memorial lecture, *The Quarterly Journal of Experimental Psychology*, 40, 201–237

Tremblay, N., Warren, R., and Dykes, R. (1990a), Electrophysiological studies of acetylcholine and the role of the basal forebrain in the somatosensory cortex of the cat. i. cortical neurons excited by glutamate, *Journal of neurophysiology*, 64, 4, 1199–1211

Tremblay, N., Warren, R., and Dykes, R. (1990b), Electrophysiological studies of acetylcholine and the role of the basal forebrain in the somatosensory cortex of the cat. ii. cortical neurons excited by somatic stimuli, *Journal of neurophysiology*, 64, 4, 1212–1222

Wilson, H. R. and Cowan, J. D. (1973), A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue, *Kybernetic*, 13, 55–80

Womelsdorf, T., Anton-Erxleben, K., and Treue, S. (2008), Receptive field shift and shrinkage in macaque middle temporal area through attentional gain modulation, *The Journal of Neuroscience*, 28, 36, 8934–8944

Xerri, C., Stern, J., and Merzenich, M. (1994), Alterations of the cortical representation of the rat ventrum induced by nursing behavior, *The Journal of neuroscience*, 14, 3, 1710–1721

# Appendix A - Detailed description of the model

## Tabular description

**A Model Summary**

| | |
|---|---|
| Populations | Three: receptors, thalamus and cortex |
| Topology | Two-dimensional, toric |
| Connectivity | Feed forward: one-to-all, lateral: all-to-all (including self-connections) |
| Neuron model | Dynamic rate model |
| Channel model | – |
| Synapse model | – |
| Plasticity | Oja-like learning rule |
| Input | Touch pressure from skin receptors |
| Measurements | Classical and non-classical receptive fields, topographic maps |

**B Topology**

| *Name* | *Type* |
|---|---|
| Receptors | Two-dimensional toric regular grid with jitter (5%) |
| Thalamus | None |
| Cortex | Two-dimensional toric regular grid |

**C Populations**

| *Name* | *Elements* | *Size* |
|---|---|---|
| Receptors | Mechanic input | 256 |
| Thalamus | Thalamic neuron | 256 |
| Cortex | Cortical neuron | $32 \times 32$ |

**D Connectivity**

| *Name* | *Source* | *Target* | *Pattern* |
|---|---|---|---|
| – | Receptors | Thalamus | Excitatory, fixed, one to one, non-plastic |
| $W_f$ | Thalamus | Cortex | Excitatory, random uniform, one to all, **plastic** $W_f(\mathbf{x}, t) = U(0, 1)$ |
| $W_e$ | Cortex | Cortex | Excitatory, Gaussian, toric, all to all, non-plastic $W_e(\lvert\mathbf{x} - \mathbf{y})\rvert = K_e \exp\left(\frac{-\lVert\mathbf{x}-\mathbf{y}\rVert^2}{2\sigma_e^2}\right)$ |
| $W_i$ | Cortex | Cortex | Inhibitory, Gaussian, toric, all to all, non-plastic $W_i(\lvert\mathbf{x} - \mathbf{y}\rvert) = K_i \exp\left(\frac{-\lVert\mathbf{x}-\mathbf{y}\rVert^2}{2\sigma_i^2}\right)$ |

**E1 Neuron Model**

| | |
|---|---|
| *Name* | Thalamic |
| *Type* | Rate model |
| *Membrane potential* | $I(\mathbf{x},\mathbf{z},t) = 1 - \frac{1}{k}\sum_{i=0}^{k} \lvert s_i(\mathbf{z}) - w_f^i(\mathbf{x},t)\rvert$ |

**E2 Neuron Model**

| | |
|---|---|
| *Name* | Cortical |
| *Type* | Dynamic rate model |
| *Membrane potential* | $\frac{1}{\tau}\frac{\partial U(\mathbf{x},t)}{\partial t} = -U(\mathbf{x},t) + \alpha\int_\Omega W_l(\lvert\mathbf{x}-\mathbf{y}\rvert) f(U(\mathbf{y},t))d\mathbf{y} + \alpha I(\mathbf{x},\mathbf{z},t)$ |

**F Plasticity**

| *Name* | *Description* |
|---|---|
| $W_f$ | $\frac{1}{\gamma}\frac{\partial W_f(\mathbf{x},t)}{\partial t} = \gamma(s(\mathbf{z},t) - w_f(\mathbf{x},t))\int_\Omega W_e(\lvert\mathbf{x}-\mathbf{y}\rvert) f(u(\mathbf{y},t))d\mathbf{y}$ |

**G Input**

| *Type* | *Description* |
|---|---|
| Mechanic | $s_i(\mathbf{z}) = \exp(-\frac{1}{2}\lVert\mathbf{z}-\mathbf{r_i}\rVert)$ |

**H Measurements**

Classical and extended receptive fields of all cortical neurons

Table 1: Tabular description of the model following the prescription of **Nordlie et al.** (2009)

# Algorithms

## Training protocol

The initial training protocol described by equations (1) to (5) can be described using the following algorithm. The full description of the learning process is given in Materials and Methods section and in **Detorakis and Rougier** (2012) as well.

**Algorithm 1** Learning protocol

**Require:** $K_e, K_i, \sigma_e, \sigma_i, \gamma, \alpha, \tau, dt, S, \varepsilon$
  Compute $w_e$, $w_i$ and $w_l$ according to equation (4)
  **for** each stimulus $\mathbf{z}$ in S **do**
    Reset the activity $u(\mathbf{x}, 0)$
    Compute receptors activity $s(\mathbf{z})$ according to equation (1)
    Compute input activity $I(\mathbf{x}, t)$ according to equation (2)
    **while** $\lvert u(t) - u(t+dt)\rvert < \varepsilon$ **do**
      Update $u(\mathbf{x}, t)$ according to equation (3)
      Update $w_f(\mathbf{x}, t)$ according to equation (5)
    **end while**
  **end for**

### RoI/modulation protocol

The RoI protocol with modulation requires a slightly different alorithm since $w_e, w_i$ and $w_l$ depend on the position of the stimulus.

**Algorithm 2** RoI/modulation protocol

**Require:** $K'_e, K'_i, K"_e, K"_i, \sigma_e, \sigma_i, \gamma, \alpha, \tau, dt, S, \varepsilon$
  Compute $w'_e$, $w'_i$ and $w'_l$ according to equation (4)
  Compute $w''_e$, $w''_i$ and $w''_l$ according to equation (4)
  **for** each stimulus $\mathbf{z}$ in S **do**
    Reset the activity $u(\mathbf{x}, 0)$
    Compute receptors activity $s(\mathbf{z})$ according to equation (1)
    Compute input activity $I(\mathbf{x}, t)$ according to equation (2)
    **if** stimulus is inside the RoI **then**
      $w_e, w_i, w_l \leftarrow w'_e, w'_i, w'_l$
    **else**
      $w_e, w_i, w_l \leftarrow w''_e, w''_i, w''_l$
    **end if**
    **while** $|u(t) - u(t + dt)| < \varepsilon$ **do**
      Update $u(\mathbf{x}, t)$ according to equation (3)
      Update $w_f(\mathbf{x}, t)$ according to equation (5)
    **end while**
  **end for**

## Convergence

In order to measure the convergence of the learning process, we measured:

1. the evolution over time of the root mean square error (RMSE) of the feed-forward weights
2. the evolution over time of the receptive fields during the development of the topographic map and during the RoI protocol

RMSE has been measured using the following equation:

$$\text{RMSE}[w_f] = \frac{1}{M}\sum_{i=0}^{M}(\hat{\mathbf{w}}_f^i - \mathbf{w}_f^i)^2, \tag{6}$$

where $\hat{\mathbf{w}}_f^i$ dsignate the final feed-forward weights and $\mathbf{w}_f^i$ the feed-forward weight at epoch $i$. Data is recorded every 50 epochs and RMSE is computed at the end of the simulation since it requires the final weights. Figure 8 displays RMSE for initial training and clearly shows the rapid convergence of the model on the final set of weights.

## Parameters

All the parameters of the model are given in table 2 below, where $n$ is the number of neurons of the cortex model, $K_e$ is the excitatory gain, $\sigma_e$ is the variance of the excitatory lateral connections, $K_i$ is the inhibitory gain, $\sigma_i$ is the variance of inhibitory lateral connections, $\alpha$ is a free scaling parameter, $\tau$ is the synapses temporal decay constant, and $\gamma$ is the learning rate. For the RoI protocol, we use two sets of parameter, the first set (RoI out) is used whenever the center of a stimulus is not within the RoI while the second set (RoI in) is used each time a the center of a stimulus is within the RoI. The parameters of the model have been tuned manually and a more detailed description of the role of each parameter can be found in **Detorakis and Rougier** (2012).

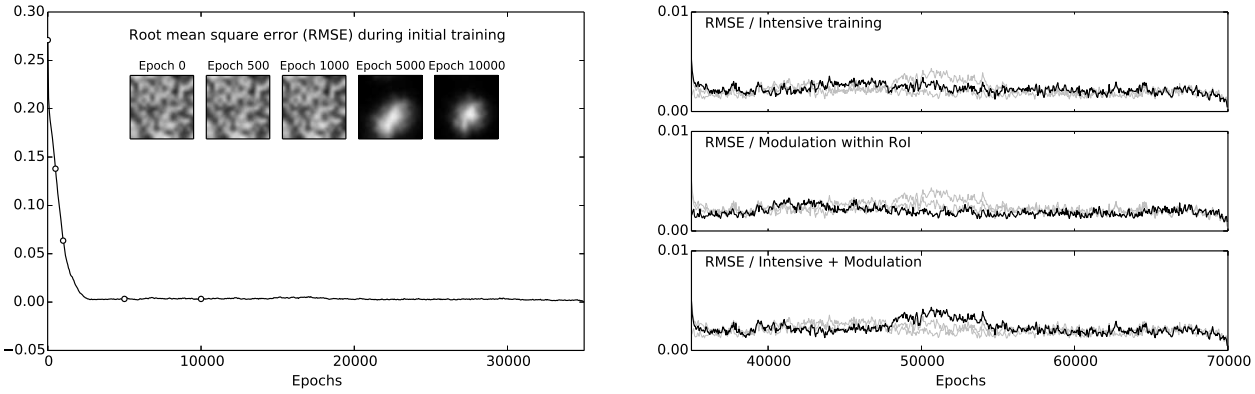


Figure 8: **Root-mean-square-error over thalamo-cortical weights.** Left. RMSE has been computed every 50 epochs during the initial training protocol and show a rapid convergence after 5000 epochs Right. RMSE has been computed for all RoI protocols to ensure no divergence occurs during training.

| *Protocol* | $n$ | $K_e$ | $\sigma_e$ | $K_i$ | $\sigma_i$ | $\alpha$ | $\tau$ | $\gamma$ |
|---|---|---|---|---|---|---|---|---|
| Training | 32×32 | 3.72 | 0.1 | 2.40 | 1.0 | 0.1 | 1.0 | 0.03 |
| Drum | 32×32 | 3.72 | 0.1 | 2.40 | 1.0 | 0.1 | 1.0 | – |
| RoI out | 32×32 | 8.02 | 0.1 | 6.10 | 1.0 | 0.1 | 1.0 | 0.03 |
| RoI in | 32×32 | 3.72 | 0.1 | 2.40 | 1.0 | 0.1 | 1.0 | 0.03 |

Table 2: Model parameters

## Simulation details

Simulations were performed on a HP Z800 Workstation with 8Gb of memory. The source code of all simulations is written in Python using Numpy and Scipy scientific libraries, and it can be found at http://webpages.lss.supelec.fr/perso/georgios.detorakis/software/index.html. The training simulation consumes $\sim$ 45MB of physical memory and requires $\sim$ 30 minutes of real time (measured with *time* Unix command). In all simulations, we used parameters given in table 2.

# Appendix B - Data analysis

## Classical receptive fields

A classical receptive field or a receptive field is defined as the locus of the skin on which a stimulus triggers a neural response, and describes as well the impact of the stimulus to the firing rate of a neuron. A cRF is a two-dimensional surface, which slope indicates the impact of the stimulus and its location points to the skin surface which is represented by the corresponding neuron. We computed the cRF of each neuron using a grid of $p \times p$ regularly spaced stimuli over the normalized skin patch. For each stimulus and after the model has converged, the activity of each neuron of the model has been recorded and aggregated into a $p \times p$ matrix that was identified as the cRF of the neuron. cRFs have been further charaterized by determining their center of mass and size. The center of mass $\mathcal{C}$ is computed according to the following equation:

$$\mathcal{C} = \frac{\sum_{i=0}^{p^2} \mathcal{V}_i \mathbf{s_i}}{\sum_{i=0}^{p^2} \mathcal{V}_i} \tag{7}$$

where $\mathbf{s}_i$ is the position of a stimulus ($i$ in $[1, p^2]$) and $\mathcal{V}_i$ is the related activity of the neuron (activity of the neuron when stimulus $\mathbf{s}_i$ is presented). The cRF area is computed using a normalized sum of the elements of the cRF that are grater than a threshold value ($< 0.05$). Using these two informations, we can build plots of cRFs lying at their respective position on the skin patch with high precision. We can therefore qualify the topographic organization of the cRF and, when relevant, we can define the migration as well as size modification. The density of cRF over the skin patch was evaluated using a Gaussian filter with standard deviation of 0.5 in each dimension.

## Non-classical Receptive fields

We define non-classical receptive fields as the weights $\mathbf{b}$ that minimize the error between the predicted firing rates of neurons and the observed ones. Such non-classical receptive fields reveal what is the optimal stimulus and how stimulus affect the firing rates of neurons. We computed non-classical receptive fields (ncRFs as opposed to cRFs) using the protocol and method (analogue to reverse correlation) defined in **DiCarlo et al.** (1998). Therefore, the predicted impulse rate $r_p$ of a neuron in response to the $n$th stimulus is defined to be the sum of the effects of each skin subregion to the neuron (see drum protocol for more details). Hence,

$$\mathbf{r}_p = b_0 + \sum_{i=1}^{p} b_i x_i(n) \tag{8}$$

where $n$ is the number of the stimuli, $p$ is the number of the different skin subregions, $x_i$ is the stimulus relief, $b_0$ is the background firing rate and $b_i$ is the strength of the effect of a dot (since the stimulus each time is a random dot pattern). Rewriting equation (8) in vector form we have,

$$\mathbf{r}_p = \mathbf{X}\mathbf{b}$$

where $\mathbf{r}$ is a $n \times 1$ vector of the firing rates, $\mathbf{X}$ is a $n \times p$ matrix with values of ones in the first column and the stimuli in the remaining columns. Finally $\mathbf{b}$ is a $p \times 1$ vector, which contains the weights of the effect of a stimulus to the neuron. This term is actually the non-classical receptive field of a neuron, since its values indicate the way that a stimulus affect the firing rate properties of a neuron. In addition, $\mathbf{b}$ can be used in order to investigate further the optimality of each stimulus, in terms of neural responses. Besides the predicted firing rates, we had also a vector of observed discharge rates for every neuron, $\mathbf{r}_o$.

In order to compute the vector $\mathbf{b}$, which minimizes the mean-squared error between $\mathbf{r}_p$ and $\mathbf{r}_o$, we solved the linear normal equations $\mathbf{X}^T\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{r}_o$ by inverting the stimulus autocorrelation matrix,

$$\mathbf{b} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{r}_o \tag{9}$$

This method allows to compute the non-classical receptive fields, $\mathbf{b}$ for all the neurons at one and indicates skin area that contribute positively to the firing rate (excitatory part) or negatively (inhibitory part). Throughout of this work anytime we refer to the non-classical receptive fields we mean the optimal vectors $\mathbf{b}$. For more details about this method refer to **DiCarlo et al.** (1998).

## Signal-to-noise ratio (SNR) and Noise Index (NI)

When we applied the threshold method described in the main text, we also measured the SNR and the Noise Index. From signal processing definition, noise is defined to be the residual of the subtraction between an original signal $s$ and the filtered one $\bar{s}$ (see main text for more details about filtering and thresholding method). In our case, the original signal is the receptive field and the filtered signal is the receptive field after the application of a Gaussian filter with zero mean and variance $1,7$. SNR is given by $\text{SNR}=10\log_{10}\left(\frac{S}{N}\right)$, where $S$ is the power of the original signal $s$ and $N$ is the noise power. We measured the SNR for each RF and we found that the mean SNR was 8 (S.D.=1.5).

In addition, and in accordance with **DiCarlo et al.** (1998), we computed the Noise Index (NI) according to equation $NI = 100\frac{\text{Var}[noise]}{\max\{|s|\}}$. $NI$ is given in terms of the absolute peak of the original signal $s$. We measured the Noise Index for each of the RF, finding a mean value of 1.5% (S.D.=0.01). Conclusively, both measures $SNR$ and Noise Index pinpoint that the noise was eliminated.

## Relative histogram of cRFS

Histograms of receptive field sizes have been made relatively to the control case in order to underline the various changes. Considering the mean size $M$ (over the $32 \times 32$ RFs) of the control and a set $\{X_i\}_n$ of observations, histogram were built using a set $\{Y_i\}_n$ such that $Y_i = 100(X_i - m)/m$ with a fixed size bins of width 8%. Any bin with negative abscissa indicated a shrinking in RF size while positive abscissa indicated an expansion.

## Cortical representation

Cortical representation is defined as the area of the cortical surface that respond when a stimulus is present within the RoI. Using the RoI protocol, we recorded the two-dimensional activation of the whole cortical model if, and only if, the stimulus center was located within the RoI (roughly half of the 25,000 stimuli). We summed these activations and normalized the result to get values in the [0,1] interval.

## Spatial event plot

During the execution of the drum protocol, the alignment of the stimulus and the responses of a specific neuron was done by applying the spatial event plot (SEP) method described in **Phillips et al.** (1988) and used by **DiCarlo et al.** (1998). The latter used a variation of SEP method in order to achieve better resolution. Since we deal with a computational model, the first version of the method was used because it is easier to achieve the alignment of the stimulus and the neural response. We isolated the receptive field of a specific neuron and applied the drum protocol (see main text for more details), recording all the responses of that neuron. The stimulus and the responses of the neuron were recorded for later processing. We processed the array of the responses by assigning a one to the element of array which contained an evoked response and a zero otherwise. Because the events array is a one-dimensional, we had to transform it into a two-dimensional one in order to fit the drum (input space). Therefore, we started from the lower right

corner (where the first stimulus was applied according to the drum protocol) and we continuously skimmed the array. Every time the drum swept was completed, we moved the next line of the array above the current one. At the end of this process, we obtained a spatial event plot, as it is depicted in figure 9B. In addition to the SEP, we also convolved the stimuli (dot patterns) with the non-classical receptive field of a neuron such as to obtain a prediction of the responses of the neuron to each stimulus. This method leads to the results shown in figure 9A for a specific neuron.

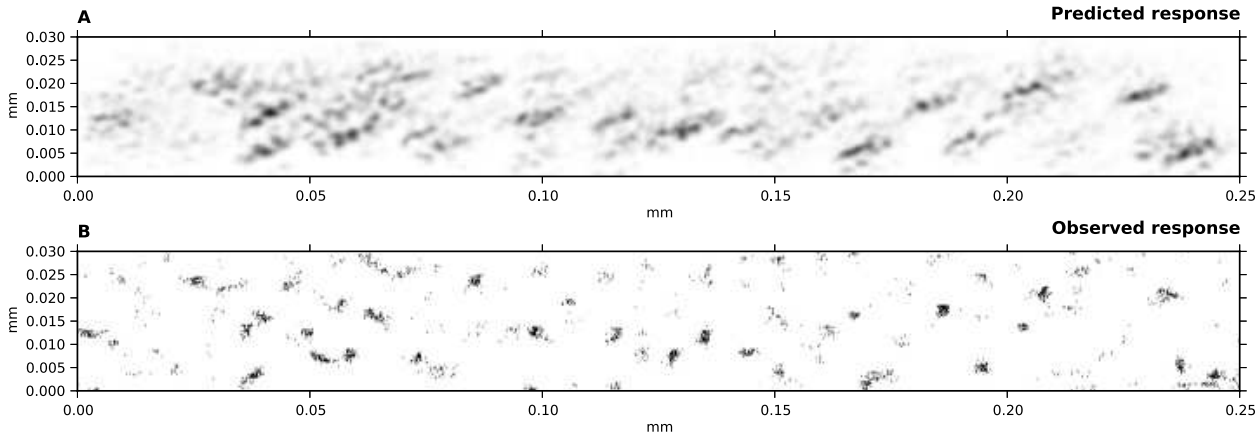


Figure 9: **Observed and predicted responses.** A. Neural impulse rates of the neuron $(20, 16)$. In order to obtain this plot, we convolved the ncRF of the neuron with the random dot stimulus pattern. B. Spatial event plot of neuron $(20, 16)$. Each dot represents the position where a stimulus triggers a response for the neuron. The axis $x$ and $y$ represent the length and the width of the drum (drum protocol, see the main text) in $mm$, respectively.