Research Report

Characterising stimulus-specific adaptation using a multi-layer field model

Robert Mill\textsuperscript{a,c,*}, Martin Coath\textsuperscript{b,c}, Thomas Wennekers\textsuperscript{b,c}, Susan L. Denham\textsuperscript{a,c}

\textsuperscript{a}School of Psychology, University of Plymouth, UK
\textsuperscript{b}School of Computing and Mathematics, University of Plymouth, UK
\textsuperscript{c}Centre for Robotics and Neural Systems, University of Plymouth, A311 Portland Square, PL4 8AA, Plymouth, UK

ARTICLE INFO

Article history:
Accepted 26 August 2011

Keywords:
Stimulus-specific adaptation
Neural field model
Synaptic depression
Oddball sequence

ABSTRACT

The response of an auditory neuron to a tone is often affected by the context in which the tone appears. For example, when measuring the response to a random sequence of tones, frequencies that appear rarely elicit a greater number of spikes than those that appear often. This phenomenon is called stimulus-specific adaptation (SSA). This article presents a neural field model in which SSA arises through selective adaptation to the frequently-occurring inputs. Formulating the network as a field model allows one to obtain an analytical expression for the expected response of a simple two-layer model to tones in a random sequence. The sequences of stimuli used in SSA experiments contain hundreds—and sometimes thousands—of tones, and these experiments routinely measure the response to many such sequences. A conventional neural network model (e.g., integrate-and-fire) would require numerical integration over long time periods to obtain results. Consequently, a field model that offers an immediate, analytical solution for a given input sequence is helpful. Two routes to obtaining this solution are discussed. The first involves the convolution of two closed-form expressions; the second relies on a series of approximations involving Gaussian curves. The purpose of the paper is to describe the model, to develop the approximations that allow an analytical solution, and finally, to comment on the output of the model in light of the SSA results published in the physiology literature.

This article is part of a Special Issue entitled "Neural Coding".

© 2011 Published by Elsevier B.V.

1. Introduction

Regular sequences of tones have long been used to assess the response properties of neurons sensitive to sound. For example, often a tone is presented repetitively, and on each trial, some intrinsic feature of the tone is varied, such as its frequency or intensity. The spikes elicited by each tone are then counted, enabling the experimenter to produce a tuning curve for that feature (e.g., Dorrn et al., 2010; Sun et al., 2010).

In recent years, there has been a growing interest in how the responses of auditory neurons depend on the context in which tones appear. One context-dependent effect evident in neurons at multiple sites along the auditory pathway is stimulus-specific adaptation (SSA) (Anderson et al., 2009, ...
Antunes et al., 2010; Malmierca et al., 2009; Pérez-González et al., 2005; Ulanovsky et al., 2003, 2004; von der Behrens et al., 2009). A neuron that exhibits SSA responds more vigorously to a tone if it appears rarely in a sequence than if it appears often. As it is the same tone that is presented in each case, any statistical difference in the response must be due to the stimulus history, if external factors are controlled for. It has been suggested that SSA in single cells contributes to neural responses measured at other scales, for example, mismatch negativity (MMN) in the event-related potential (Nelken and Ulanovsky, 2007) and SSA in current source density recordings (Szymanski et al., 2009), and also that it might be implicated in auditory scene analysis (Winkler et al., 2009).

This article presents a neural field model of stimulus-specific adaptation. Neural field models aim to capture the dynamics that govern coarse-grained quantities in a patch of neural tissue, such as the firing rate. This contrasts with the approach of modelling discrete units, e.g., neurons and synapses, individually. (See Coombes (2005, 2010) for a review.) The earliest example of such work was that of Wilson and Cowan (1972, 1973) (but see also Beurle, 1956) who used mean field theory to describe the collective behaviour of a large population of neurons. (See Destexhe and Sejnowski, 2009, for a brief note on the work and its legacy.) Amari (1977) later used neural field models to identify patterns of electrical activity that form and propagate spontaneously in spatially-extended layers with homogeneous connectivity. The equations established by these earlier studies did not include the effects of spike-rate adaptation or synaptic depression. Pinto and Ermentrout (2001) introduced a linear negative feedback term to represent broadly “spike frequency adaptation, synaptic depression, or some other slow process that limits the excitation of the network.” More recently, Kilpatrick and Bressloff (2010) developed field equations that account for two separate sources of adaptation: a hyperpolarising current (related to spike-rate adaptation) and a synaptic scaling factor (related to synaptic resource availability). Synaptic depression operates in their equations along the same lines as the model proposed by Tsodyks et al. (1998); namely, the efficacies of the synapses originating at a neuron are multiplied by a term which depends on the recent activity of that neuron. Synaptic depression is treated similarly in this work (see Section 2.3 for more details). We do not model spike-rate adaptation.

In this model of SSA, repetition of a standard tone depletes the synaptic resources associated with a localised region of the input field, leading to a reduced response in the postsynaptic target. Conversely, the presentation of a deviant tone activates distinct synaptic resources, which are available, and this in turn results in a larger response. The suggestion that SSA arises from adaptation on the inputs to a neuron (due, for example, to synaptic depression) was advanced in one of the earliest papers to describe the phenomenon (Ulanovsky et al., 2004).

There are specific motivations in the present work for using field models rather than those that consist of discrete units, e.g., integrate-and-fire neurons. Firstly, the input stimuli submitted to the model are thousands of seconds long, and evaluation of the model for several conditions runs into hours of simulated time. Secondly, the model of SSA we propose does not possess any inhomogeneities that would hinder a neural field approach or lead us to prefer a discretised approach. Thirdly, and most importantly, formulation of the model in field terms allows us to search for an analytical solution, thus obviating the need to integrate the equations of the model in discrete time steps.

We present an approximate analytical solution to a particular kind of model (two layers connected via depressing synapses), in response to a particular kind of input (stimuli presented with regular timing and at randomly-chosen locations). Many of the approximations made exploit the fact that quantities in the model vary smoothly along the spatial axis, consisting of bumps and notches that can be approximated adequately using linear combinations of Gaussian curves. Gaussian curves are widely used in neural field modelling on account of their convenient analytical properties, in particular, closure under multiplication and convolution, and are frequently used to represent localised inputs or coupling kernels (e.g., Bressloff et al., 2003; Wennekers, 2002). Much of the analysis can proceed without making recourse to Gaussian approximations (Sections 3.1–3.2), but the final stage, which requires spatially integrating the activity across an entire layer must be carried out either numerically or by means of analytical approximations (Sections 3.3–3.5).

The remainder of the paper is structured as follows. Section 2 describes the format of the input sequence and presents the equations that govern the field model. Notational conventions for operations on Gaussian structures are introduced. Section 3 contains an analysis of a particular network, in which there are two layers connected by feed-forward depressing synapses. Analytical solutions for the steady-state of the synapses are given (as they have been previously: Tsodyks et al., 1998), and these are approximated using Gaussians. Section 4 presents the results obtained using the model and evaluates them from two perspectives: (i) whether the results are in keeping with those known from physiology, and (ii) whether the approximations using Gaussians are successful. Section 5 concludes the paper.

2. Methods

2.1. Oddball sequences

The extent to which a neuron exhibits SSA can be assessed using oddball sequences, which are sequences containing two frequencies, $f_1$ and $f_2$, one of which is common (the standard), and the other rare (the deviant). Typically, the two frequencies are evenly spaced on either side of the neuron’s best frequency (Ulanovsky et al., 2004) or chosen so that they elicit the same number of spikes when presented in isolation (Malmierca et al., 2009). Oddball sequences are typically parameterised by four values: $\Delta f$, the tone frequency separation, in octaves; $P_{\text{dev}}$, the probability of the deviant; $T_d$, the duration of a tone; and $T_s$, the length of the silent interval between the offset of one tone and the onset of the next. The interval between the onsets of consecutive tones is given by $T = T_1 + T_2$.

In order to control for the possibility that the neuron being studied is naturally tuned to prefer either $f_1$ or $f_2$, experimenters present the same sequence twice, with the standard and deviant swapped in the second instance. The degree of
stimulus specific adaptation is then quantified by the SSA index (SI), defined as (Ulansky et al., 2003):

\[
SI = \frac{d(f_1) - s(f_1) + d(f_2) - s(f_2)}{d(f_1) + s(f_1) + d(f_2) + s(f_2)}
\]  
(1)

where \(s(f_i)\) and \(d(f_i)\) denote the spike count elicited in response to tone \(k\), when it is presented as the standard or the deviant, respectively. A significantly positive SI values indicates SSA. SSA is absent when SI = 0. Fig. 1 illustrates two eight-tone oddball sequences. The sequences used in SSA experiments are several hundreds of tones long in practice.

2.2. Network structure

The systems of equations we propose in this paper assume that neural tissue is organised into \(n\) one-dimensional layers. The net synaptic input to layer \(j\) at spatial location \(s\) and time \(t\), denoted \(u_j(s,t)\), is a function of both direct input and activity in other layers, and evolves according to the equation

\[
\tau_x \frac{du_j(s,t)}{dt} = -u_j(s,t) + f_j(s,t) + \sum_{i=1}^{n} w_{ij}(r,s,t) u_i(r,t) dr.
\]  
(2)

Here \(f_j(s,t)\) is an external input to layer \(j\), and \(u_j(r,s,t)\) is the synaptic weight connecting location \(r\) in layer \(i\) to location \(s\) in layer \(j\), at time \(t\). \(\tau_x\) is a time constant, and \(v_j(r,t) = \Phi(u_j(r,t))\) is an instantaneous firing rate in layer \(j\) obtained by applying a static nonlinearity \(\Phi[\cdot]\). In this work, we assume that neurons are operating within a linear regime so that \(\Phi[u]=u\).

2.3. Synaptic dynamics

It should be noted that the coupling kernels depend on time. The efficacy of the synapses connecting locations \(r\) and \(s\) is taken to derive from a product of two factors:

\[
w_{ij}(r,s,t) = Z_{ij}(|r-s|) \cdot xi(r,t).
\]  
(3)

The first factor, \(Z_{ij}\), is static, and relates to a fixed density of synapses projecting from \(r\) to \(s\). This density is in turn taken to be a function of the distance between \(r\) and \(s\). Consequently, the coupling profile is symmetric and spatially-invariant, and the integrals in Eq. (2) become convolutions (i.e., \(Z_{ij} \cdot \Phi[v(r,t)]\)). In this work, we assume that the \(Z_{ij}\)s are Gaussians with height \(Z_0\) and width \(a_0\), i.e.,

\[
z_{ij}(|r-s|) = Z_0 \exp \left[ \frac{(r-s)^2}{-2a_0^2} \right].
\]  
(4)

The second factor, \(x(r,t)\), is dynamic, and it relates to the effectiveness of transmission across the synapse that originate at location \(r\) at time \(t\). The dynamics governing the effectiveness of a synapse are presumed to stem from the use, re-uptake and loss of transmitter substance, along similar lines to a number of other models that have been proposed (Meddis, 1986; Sumner et al., 2003; Tsodyks and Markram, 1997).

In this model, a unit quantity of resource at a synapse is partitioned into two states: effective, \(x(r,t)\); and ineffective, \(y(r,t)\). Resources are consumed at a rate proportional to the pre-synaptic firing rate at \(r\), and undergo continuous recovery, according to the following pair of equations:

\[
\dot{x}_i(r,t) = b_i(1-x_i(r))-a_i u_i(r,t) x_i(r)
\]  
(5)

\[
y_i(r,t) = 1-x_i(r,t).
\]  
(6)

Here, \(a_i\) and \(b_i\) are kinetic parameters, which govern the rate of depression and recovery, respectively. Both \(a\) and \(b\) are stated in units of \(s^{-1}\). The model of primary auditory cortex developed by Loebel et al. (2007) uses this set of dynamical equations to govern the synaptic depression on the feed-forward and lateral connections in an array of cortical columns. The model exhibits forward masking, but given the recurrent connectivity within layers, it is unclear whether SSA would be observed.

Synaptic dynamics similar to these have been analysed by Tsodyks et al. (1998) in the context of mean-field population dynamics. The synapse specified in Eqs. (5) and (6) is a simplified version of the one referred to there: it neglects the faster component associated with the post-synaptic potential, and a facilitating factor that, in the context of this formulation, effectively leads to a short-term usage-dependent increase in \(a_i\). Neither of these omitted features is essential to our analysis of SSA, which is chiefly concerned with depression.

The analysis of synaptic depression offered here differs to that of Tsodyks et al. (1998), in that the latter were concerned to model the dynamics of interconnected populations. In this work, we intend to model the effects of synaptic depression upon activity distributed along a spatial axis. In addition, our enquiry is restricted to the response of a specific network to oddball random processes in particular.

2.4. Motivation for Gaussian approximations

A notable feature of the response of simple field models to point- or bump-like inputs is that, under many circumstances, the spatial profiles of the dynamical variables \((u, v, x, y, \text{etc.})\) also assume a unimodal form. Similarly, multimodal inputs tend to generate extrema in the spatial profiles that correspond to peaks in the input and nonlinear interactions between them.
For simple nonlinear systems, expressions for the functional forms underlying these bump-like shapes may be forthcoming, even in a closed form. For instance, when the output function, \( \Phi \), is the unit step function, bumps can be constructed in closed form for any choice of synaptic kernel (e.g., Amari, 1977). However, in attempting to find the time-invariant responses of the kind of multi-layer network outlined above, spatial integrals arise which do not admit a closed form (see below).

Wennekers (2001, 2002), observed that a bell-shaped firing rate profile could be approximated by a Gaussian with identical zero-th and second order derivatives around its mean. A profile of infinite extent is thus summarised in just three parameters, the spread of activation in the input layer relates in some way to the width parameter could be approximated by a Gaussian with identical zero-th and second order derivatives around its mean. A profile of infinite extent is thus summarised in just three parameters, the spread of activation in the input layer relates in some way to the width parameter.

Wennekers (2002) used a field model of the kind introduced in Section 2.2 to trace the dynamics of a single bump in different layers. In order to shed light on SSA, this work has to differ in at least two respects. First, whereas in the previous work, the weight profiles were static, in this work, they depend on time and are themselves governed by a dynamic system (Eqs. (5), (6)). Second, the previous work considered the response to a single point-spread function, centred on \( s=0 \). In order to investigate SSA, the field model must capture the response at two, distinct input locations (at least).

These extra considerations demand a sharp departure from Wennekers’ original formulation; but the original insight of characterising curves using Gaussians is retained.

2.5. Operations on Gaussian combinations

A single Gaussian component \( g \) is parameterised by the triple \( (\lambda; \mu; \sigma) \) and, when supplied with an argument, \( s \), is evaluated according to

\[
g(s) = A \exp \left[ \frac{(s-\mu)^2}{-2\sigma^2} \right]
\]

Let the set of all possible triples be designated \( \mathcal{G} \). The power set, \( \mathcal{G}^2 \), then contains a description for every possible linear combination of Gaussians. If \( \mathcal{G} = \mathcal{G}^2 \), then

\[
G(s) = \sum_{g \in \mathcal{G}} g(s).
\]

Linear combinations of Gaussians are often applied in function approximation and, notably, in radial basis function networks and Gaussian mixture models for density estimation (Powell, 1987; Bishop, 1995, ch. 5).

The most convenient aspect of the set of linear combinations of Gaussians, from our current perspective, is its closure under several binary operations, including \( \forall G_1, G_2 \in \mathcal{G}^2 \): addition \( (G_1 + G_2) \in \mathcal{G}^2 \), subtraction \( (G_1 - G_2) \in \mathcal{G}^2 \), multiplication \( (G_1 \cdot G_2) \in \mathcal{G}^2 \), convolution \( (G_1 * G_2) \in \mathcal{G}^2 \). Furthermore, \( \mathcal{G}^2 \) is also closed under (i) unary operations such as negation, scalar multiplication and raising to a positive integer power, and (ii) linear transforms, such as translations \( (G(s-\alpha)) \in \mathcal{G}^2, \alpha \in \mathbb{R} \) and dilations \( (G(\sigma s)) \in \mathcal{G}^2, \sigma \in \mathbb{R} \).

We also note that: (i) \( \int G(s) ds \) has a closed-form solution, (ii) \( \mathcal{G}^2 \) contains real-valued constant functions, \(^1\) (iii) generally, closure under division is not satisfied (\( \frac{G_1}{G_2} \in \mathcal{G}^2 \)), (iv) nor is there closure under raising to non-integer powers.

These properties allow one to write the entire set of field equations outlined earlier (Eqs. (2), (4), (5), (6)) in the compact manner below, following a convention of moving the quantities being represented into superscripts on \( G \):

\[
\begin{align*}
\tau_{G_i} & = -G_i + \sum_{j=1}^{N} G_j \ast G_i \\
G_{ij} & = G_i \ast G_j \\
\dot{G}_i & = b_i (1-G_i) - a_i G_j \ast G_i \\
G_i & = 1 - G_i.
\end{align*}
\]

This Gaussian notation is used in Section 3 to simplify the analysis: intractable expressions are substituted for Gaussian equivalents that can be readily combined using operators, as in the example above. We note in passing that, in principle, it is possible to solve this system numerically whilst retaining the Gaussian notation, as finite differencing is a special case of addition and scaling, i.e., \( G(t + \delta t) = G + G \delta t \). However, this method is impractical for this set of equations, as cross-terms proliferate when Gaussians with different means are multiplied. At any rate, our concern is not the temporal integration of dynamic equations, but rather spatial integration of the time-invariant solution to those dynamic equations—as will become apparent.

3. Analysis of a two-layer field model

In this section, we develop and analyse the response of a field model to oddball sequences. We consider a two-layer model, in which the second layer spatially integrates the activity in the first layer, this integration being mediated by the convergence of depressing synapses. For simplicity, the time constants \( \gamma_i \) are set to zero, so that temporal integration is instantaneous. The static weight profile connecting the first layer to the second has a standard deviation of \( \sigma_{1,2} = 0.4 \).

The input to the model—an oddball sequence—is a random process. Consequently, the state variables of the model are best characterised as random processes (or random variables, for a given \( s \) and \( t \)). Thus our concern here is the expected response to a process, not the deterministic response to a realisation of the process.

3.1. Input patterns

We wish to characterise the response of the model to random sequences of inputs, presented isochronously with period \( T \). If the stimuli in question are pure tones, and the spatial dimension \( s \) (or \( r \)) is taken to be tonotopic, then frequencies \( f_1, \ldots, f_k \) map in an orderly way to spatial locations \( s_1, \ldots, s_k \), and the spread of activation in the input layer relates in some way to peripheral auditory tuning.

\(^1\) This is permitted provided that \( \sigma \) is notionally allowed to tend to \( +\infty \). This is reasonable, given that the width parameter could have been specified in the numerator of the exponential argument.
The time-varying input to layer one consists of a sequence of Gaussian-shaped pulses:
\[ I_1(s, t) = \begin{cases} A_{in} \exp\left(\frac{(s-s_0)^2}{2\sigma^2}\right) & \text{tone } f_0 \text{ at time } t \\ 0 & \text{otherwise.} \end{cases} \tag{7} \]
where \( A_{in} = 1 \) and \( \sigma = 0.2 \), if it is assumed that the dimension \( s \) has octave units, and the half-peak bandwidths of peripheral auditory filters is \( \approx \frac{1}{2} \) octaves. We set \( u_1(s) = u_1(s) = I_1(s) \), so that the firing rate reflects the input directly, and use \( u_1^{[k]} \) to denote the output of the first layer when tone \( k \) is presented. Fig. 2A shows an input profile for four input tones.

The input to the model at a given point, \( s \), consists of a series of pulses of differing height. Each pulse has duration \( T \). The time-varying input to layer one consists of a sequence of pulses of differing height. Each pulse has duration \( T \), and the period of inactivity following it has duration \( T \). The height of a pulse depends on how the tone frequency is resolved at \( s \). For example, if the input profile in Fig. 2A monitored at location \( s = 0.1 \) corresponds to the series of square pulses in plotted in Fig. 2B.

### 3.2. Equilibrium synaptic profiles

Let \( x_1^s(t) \) denote the expected fraction of effective resource at the onset of a pulse for the synapses that originate at \( s \). The synaptic dynamics at the point \( s \) depend only on the net input at \( s \), and so we can concentrate our analysis on a single point. Taking the probability of pulse heights into account, along with the synaptic dynamics described in Eqs. (5) and (6), it can be shown that
\[ x_1^s(t) = 1 + \sum_{k=1}^{K} \frac{b}{c_k(s)} + e^{-c_k(s)T_1} \left( x_1^s(t) - \frac{b}{c_k(s)} \right) - 1 \] \[ e^{-c_kT_1} \tag{8} \]
where \( c_k(s) = au_1^{[k]}(s) + b \). \tag{9}

Some equations that support the derivation of Eq. (8) are supplied in Appendix A. For brevity, we now suppress the dependency on \( s \), which appears only in the term \( c_k \). Eq. (8) can be readily solved for \( x_1^s \) by rearranging:
\[ x_1^s = \frac{1 + \sum_{k=1}^{K} \frac{b}{c_k} - \frac{b}{c_k} e^{-c_kT_1} - 1}{1 - \sum_{k=1}^{K} \frac{b}{c_k} e^{-c_kT_1}}. \tag{10} \]

Once the steady state of the effective resource variable at the onset time is known, the mean time course for a given tone frequency at time \( t \) after the onset can also be computed:
\[ x_1^s(t; t_1) = \begin{cases} \frac{b}{c_k} + \left( x_1^s(t_1) - \frac{b}{c_k} \right) e^{-c_kT_1} & 0 \leq t \leq T_1 \\ 1 + \left( x_1^s(t_1) - 1 \right) e^{-c_k(T-t_1)} & T_1 < t \leq T. \tag{11} \end{cases} \]

In order to compute the SI value for certain stimulus and model parameters, we require the mean response to the two stimuli, when presented as standard and deviant. The average firing rate at position \( s \) in the second layer, measured over a presentation of stimulus \( k \), is given by
\[ R_k(s) = \int_{t_0}^{t_1} \sum_{s} \left( u_1^{[k]}(s) \right) \left[ \frac{1}{T_1} \int_0^{T_1} x_1^s(t) \, dt \right] \, ds. \tag{12} \]
where the inner integral has the closed-form solution
\[ \frac{b}{c_k} + \frac{1}{T_1} \left( x_1^s(t_1) - \frac{b}{c_k} \right) \left( 1 - e^{-c_kT_1} \right). \tag{13} \]

Replacing Eq. (13) inside Eq. (12), one sees that the mean response is the convolution of two closed-form expressions. From here one may perform the convolution numerically to obtain a firing rate for different tones. In what follows, we perform convolutions or integrations numerically, by substituting the integrands for approximations in terms of sums of Gaussians.

Fig. 3 shows the mean effective transmitter for two tone inputs at \( s_1 = 0.25 \) and \( s_2 = 0.25 \), given that \( p_1 = 0.9 \). The location of the depression shifts depending on where the tone is positioned. Fig. 3A depicts the change in resources if the standard is presented. At the onset of a tone, the resources around \( s_1 \)
tend to be used up, and the presentation of another standard simply extends this groove: it deepens initially, as the stimulus is applied, then recovers when it is released. Fig. 3B demonstrates the effect of presenting a deviant. The impression now shifts away from $s_1$, and towards $s_2$.

The surfaces that characterise the average spatio-temporal course of $x_i$, in response to tones, though described exactly by unwieldy expressions (10) and (12), qualitatively speaking, are just two impressions that evolve in depth and width in a way that depends on whether $k=1$ or $k=2$. Subtracting Gaussians from a constant can generate similar indentations in a flat surface to those in Fig. 3. It is to this approximation that we now turn.

3.3. Gaussian approximation for equilibrium state

Integration of $x_i$ with respect to $s$ is obstructed by the fact that the ratio contains terms that depend on $s$ in the numerator and denominator. These terms are themselves nonlinear in $s$, and include reciprocals of exponentials ($1/C_k$) and exponentials of exponentials ($e^{-sT}$), where $C_k=1/C_k(s)$ is defined according to Eq. (9). Furthermore, summations over $k$ appear in both halves of the ratio, making it difficult to tackle each stimulus-dependent component separately. We now describe a procedure for dealing with these difficulties, under three section headings below.

3.3.1. Decouple terms that depend on $k$
First note that $x_i$ can be approximated as a sum over terms in $k$ if the interactions between stimuli are ignored. Effectively, the equilibrium synaptic state is computed for a train of stimuli of type $k$, where a pulse or omission arises with probability $p_k$ and $1-p_k$, respectively. The components are then combined linearly using Eq. (14) below.

\[
\xi_k = \frac{1 + p_k \left( \frac{b}{C_k^0} \right) (1-e^{-\alpha_k T}) - 1 - p_k e^{-\alpha_k T}}{1 - p_k e^{-\alpha_k T} e^{bT} - (1-p_k) e^{bT}}
\]

\[
x_i = \frac{1}{\xi_k} \left( 1 + \frac{b}{C_k^0} (1-e^{-\alpha_k T}) - 1 - p_k e^{-\alpha_k T} e^{bT} \right)
\]

Here $\xi_k$ and $1-\xi_k$ denote the expected fraction of effective and ineffective substance, respectively, if all the stimuli except $k$ are omitted from the oddball sequence. Fig. 4A compares the exact solution for $x_i$ (solid) with the approximation described above (dashed) for an exemplary parameter set.

3.3.2. Remove dependency of denominator on $s$
The next step is to find an expression for each $\xi_k$ in terms of a Gaussian, or a linear combination of Gaussians. The terms that depend on $s$ are those that refer to $C_k$, namely, $b/c_k$, which appears on the top of the ratio, and $e^{-\alpha_k T}$, which is on the top and bottom. These two functions have a bowl-like shape with respect to $s$, with a global minimum $m>0$, and horizontal asymptotes that approach some value $c\to\infty$, as $s\to\pm\infty$. This shape suggests that each might suitably be approximated by a function of the form $1-G$, where $G=2^g$. However, even replacing these terms will at best result in a ratio of sums of Gaussians—as one $e^{-\alpha_k T}$ term remains in the denominator.

A surprisingly effective strategy is to disregard the variation in the $e^{-\alpha_k T}$ terms altogether by replacing them with a suitably chosen constant. We designate this approximated version $\tilde{\xi}_k$. Substituting zero is only appropriate if $aA_{in}T_1$ is quite large. In that case, the shape of the $\xi_k$ depends predominantly on $b/c_k$ while $s=s_k$; then as $|s-s_k|$ grows, $b/c_k \to 1$, from which it can be seen that $\tilde{\xi}_k \to 1$, regardless of how $e^{-\alpha_k T}$ behaves. A better approach involves substituting the nested exponential term for its value at $s=s_k$. This ensures both that $\tilde{\xi}_k(s_k)=\xi_k(s_k)$ and $\tilde{\xi}_k(s)\to 1$ as $s\to\pm\infty$. Replacing $e^{-\alpha_k T}$ with the constant

\[
q=\exp \{-aA_{in} + b\}T_1
\]

as prescribed, means that $s$ now appears only in the numerator (via $C_k$), and one can rewrite $\tilde{\xi}_k(s)$ as

\[
\tilde{\xi}_k(s) = \alpha_k \left( 1 - \frac{b}{C_k(s)} \right) - 1
\]

\[
\alpha_k = \frac{p_k(1-q)e^{bT}}{p_kqe^{bT} + (1-p_k)e^{bT} - 1}
\]

where $\alpha_k$ is a constant. Figs. 4B and C compares an original curve $\tilde{\xi}_k$ (solid) with its approximation $\tilde{\xi}_k$ (dashed) for two values of $T_1$.

3.3.3. Substitute bowl-shaped terms for Gaussians
Thus far, we have obtained an approximation for $x_i$ as a weighted sum of bowl-shaped functions ($b/c_k$) and constant terms. Let $U(b/c_k)$ denote an approximation of $b/c_k$ in terms of a linear combination of Gaussians (i.e., an element of $2^v$). The final issue concerns the form that $U(b/c_k)$ should take. Here a basic approach is discussed (but see Appendix B).

First, note that $b/c_k \to 1$ as $s\to\pm\infty$. A simple bowl shape can thus be formed by subtracting a single Gaussian $g=\psi$ from unity:

\[
U \left( \frac{b}{C_k(s)} \right) = 1 - g(s)
\]
The mean of \( g \) clearly corresponds to \( s_k \), and its peak height corresponds to the depth of the bowl, namely, \( \frac{\alpha_k}{\alpha_{\infty}}. \) The width (standard deviation) of \( g \) can be approximated by ensuring that the curve \( g(s) \) passes halfway between the height and depth of the bowl at the same \( s \) locations as the curve \( b/c_\infty \) does. This results in a width proportional to that of the input profile:

\[
\alpha_k = \sqrt{\frac{\ln(\frac{s_k^2}{s^2} + 2)}{\ln 2}}
\]

The approximation that results from this final substitution we designate \( \xi_k \). Figs. 4B and C compare the three curves, \( \xi_k \), \( \xi_k' \) and \( \xi_k'' \), which are based on an exemplary set of parameters.

The grand approximation of \( X_k \) is found by summing over the components,

\[
X_k \approx X_k' = 1 - \sum_{k=1}^{K} (1-\xi_k).
\]

Here we use a tilde to denote the combined approximation. An example curve for \( X_k' \) is plotted in Fig. 4A as a dotted line.

\( X_k' \) is as well. Consequently, it can be approximated by ensuring that \( g \) is a linear combination of Gaussians, \( X_\text{approx} \), as well. Consequently, it can hereafter be expressed as \( G_k^≈ \in \mathbb{R}^2 \) and participate in all the operations on Gaussian combinations described in Section 2.5.

### 3.4. Reintroducing the time dimension

The previous section was concerned with the mean fraction of effective synaptic resource at the stimulus onset, \( X_k \), and how to approximate this function using Gaussians. We consider next how the surface \( X_k^\tau(t; X_k') \) varies following the onset of a given stimulus, \( k \). An exact expression for this quantity appears in Eq. (11), and the aim, once again, is to generate an approximation in terms of Gaussians, with a view to facilitating integration with respect to \( s \).

Fortunately, many of the terms appearing in Eq. (11) have already been treated in the previous section. These include the terms \( k' = -G_k^\tau(t; X_k') \) and \( b/c_\infty \). Once these terms are substituted, one factor that depends on \( s \) (via \( c_\infty \)) remains, namely, \( e^{-\alpha s} \). This function, like \( b/c_\infty \), is bowl-shaped in \( s \); and, as before, we can seek to devise an operator \( \mathcal{U}(e^{-\alpha s}) \) that finds an appropriate Gaussian combination to replace it.

In this case, we adopt a similar method to replace \( e^{-\alpha s} \) as was used for \( b/c_\infty \). A Gaussian is subtracted from a constant, yielding a bowl shape, and its parameters are chosen so that the curve matches certain strategic points of \( e^{-\alpha s} \), i.e., its base and halfway points. The expression that results from this procedure is written below.

\[
\mathcal{U}(e^{-\alpha s}) = e^{-\beta s}(1-g)
\]

\[
\mathcal{G}^{\tau}(t; X_k') = \left(1 - e^{-\beta s}; s_k; \sigma_\infty \right)
\]

\[
\mathcal{G}^{\tau}(t; X_k') = \ln \beta - \ln \left(\frac{1}{2} e^\beta + \frac{1}{2}\right)
\]

where \( \beta = -A \omega, at \).

The elements of the triple in Eq. (21) specify the height, mean and standard deviation of the Gaussian, respectively. Appendix B discusses an alternative method for choosing this Gaussian.

Once \( \mathcal{U}(e^{-\alpha s}) \) has been substituted for \( e^{-\alpha s} \), we arrive at a counterpart expression for Eq. (11):

\[
\mathcal{G}^{\tau}(t; X_k') = e^{-\beta s}(1-g)
\]

\[
\mathcal{G}^{\tau}(t; X_k') = \left(1 - e^{-\beta s}; s_k; \sigma_\infty \right)
\]

\[
\mathcal{G}^{\tau}(t; X_k') = \ln \beta - \ln \left(\frac{1}{2} e^\beta + \frac{1}{2}\right)
\]

This expression approximates the surface corresponding to the fraction of transmitter substance that is effective in the synapse located at \( s \) at time \( t \). Fig. 5 provides an example of such a surface, in both its exact (A) and approximate (B) forms.

### 3.5. Application to stimulus-specific adaptation

The sections above dealt exclusively with activity in the first layer and its effect on mean synaptic resources. The second layer receives activity in the first layer via depressing synapses and spatially integrates it according to a static weight profile \( z_{12} \). The response in layer two is therefore both stimulus- and history-dependent. For example, if stimulus \( k = 1 \) is presented often, the resources around \( s_1 \) will, on average, be depleted, and activity at \( s_1 \) will contribute little to the response in the second layer.

In order to approximate the SI value, we require the mean response to the two stimuli, when presented as standard and deviant. The response to stimulus \( k \), measured at position \( s = 0 \) in the second layer, averaged over the course of its presentation, taken from Eq. (12), is

\[
R_k = \int_{-\infty}^{\infty} z_{12}(\cdot s^0) \frac{1}{T_1} \int_{-\infty}^{\infty} X_k(t; s, t) dt \ ds
\]

\[ \int_{-\infty}^{\infty} G_{\omega}^x(t) dt \approx \int_{-\infty}^{\infty} G_{\omega}^{x|k}(t) dt \]  

(24)

Note that the spatial integrations in Eq. (23) cannot be achieved by any methods other than numerical (e.g., quadrature) integration. In contrast Eq. (24), although resting on a series of approximations, has an exact solution.

Up to this point, we have addressed every aspect of how Eq. (24) is evaluated, except for the temporal integration, which applies only to \( G_{\omega}^{x|k} \). From Eq. (13), we recall that the inner integral should approximate

\[ b \frac{1}{c_k} + \frac{1}{T_1} \left( x_1 - b + b \left( 1 - e^{-\omega T_1} \right) \right). \]  

(25)

All the factors and terms that comprise Eq. (25) are familiar from earlier discussions, and an equivalent expression in terms of a combination of Gaussians can be readily obtained:

\[ \int_{-\infty}^{\infty} G_{\omega}^{x|k}(t) dt \approx \int_{-\infty}^{\infty} G_{\omega}^{x'(k)}(t) dt + \int_{-\infty}^{\infty} G_{\omega}^{x''(k)}(t) dt \]  

(26)

The equation for the SI (3) refers to four quantities, viz., the mean responses to \( f_1 \) and \( f_2 \), when either \( f_1 \) or \( f_2 \) is deviant. Each quantity can now be computed using Eq. (24), substituting Eq. (26).

4. Results and discussion

SI values were computed using the approximations described in the previous section for various input and model parameters. Because the parameter space is quite large, we present results around a base point in the parameter space, varying a single parameter in one of six dimensions: deviant probability (\( p_{\text{dev}} \)), frequency separation (\( \Delta f \)), tone duration (\( T_1 \)), inter-stimulus interval (\( T \)), and depletion and recovery time constants (\( 1/a \) and \( 1/b \), seconds). The results are plotted in Fig. 6 as solid lines. The solid markers show the results obtained by performing the integration in Eq. (23) numerically. A cross on the abscissa marks each base parameter. The stimulus parameters are chosen to match those used in von der Behrens et al. (2009). The two model parameters, \( 1/a \) and \( 1/b \), cannot be matched directly in this way. Ulanovsky et al. (2003), referring to their own results, noted that depressing synapses with a recovery time constant of \( \sim 1.5 \) s could explain the extent to which one tone has an effect on the next, and this temporal window appears to be similar in other studies of cortical SSA (e.g., von der Behrens et al., 2009). Depressing synapses that recover on time scales with this order of magnitude are known to exist (Markram et al., 1998; Tsydyk and Markram, 1997). Consequently, we use \( 1/b = 1.5 \) s. We do not modify \( A \) or \( \sigma_{\text{SI}} \), as equivalent models can be obtained by rescaling \( a \) or \( \Delta f \), respectively.

There are two aspects to the results presented here. The first concerns how well the patterns in the results agree with those obtained in physiology. The second pertains to the adequacy of the Gaussian approximations being used. We comment upon each of these issues separately below.

4.1. Comparison with physiology

Fig. 6A shows that SSA decreases as \( p_{\text{dev}} \) increases. The same pattern was observed in all the SSA studies where \( p_{\text{dev}} \) was varied: the SI decreased as \( p_{\text{dev}} \) was increased towards 0.5 (Antunes et al., 2010; Malmierca et al., 2009; Ulanovsky et al., 2003, 2004; von der Behrens et al., 2009).

Fig. 6B shows that SSA increases as \( \Delta f \) increases, and it is equal to zero when the inputs are identical. This accords with the results in the physiological literature (Ulanovsky et al., 2003, and citations immediately above). SSA is smaller when the stimuli are closer together, because there is a greater overlap in the synaptic resources they consume; or, stated another way, the resources associated with the deviant undergo less recovery during runs of standards. The curve reaches a plateau as \( \Delta f \) grows and the inputs cease to interact.

Figs. 6C and D relate to timing: the duration of a tone (\( C \)) and the time between the onsets of tones (\( D \)). A few studies have manipulated the inter-onset interval (\( T \)), and those that do tend to report that SSA decreases as the interval between successive presentations grows longer (Antunes et al., 2010; Ulanovsky et al., 2003), although in some studies the pattern is weaker (Anderson et al., 2009; Malmierca et al., 2009). We are not aware of any studies that explicitly investigate the effect of tone duration upon SSA, and the duration used varies from study to study (e.g., 75, 200, 236 ms). In this model, the effect of tone duration and onset separation can both be understood in terms of the time that is available for synaptic resources to recover.

Figs. 6E and F concern parameters of the model rather than the stimulus. Consequently, they cannot be assigned the...
values used in original experiments in the same way. A justification for the base values of \( a \) and \( b \) used was given above. However, their effect upon SSA in this model is quite clear. Increasing the time constant \( 1/a \) means that effective resources are used up more slowly over the course of the tone, the response to the standard does not decay so rapidly following repetition, and SI decreases. As \( 1/a \rightarrow \infty \), resources are not consumed at all, and there is no SSA. The reverse principle applies in the case of the time constant \( 1/b \). As \( 1/b \) grows, ineffective resources are returned to the effective state more slowly, which causes SSA to increase. As \( 1/b \rightarrow 0 \), resources are recovered very rapidly, with the consequence that, at the onset of a tone, the system has no memory of previous tones, and there is no SSA.

4.2. Quality of the Gaussian approximations

In this final section we discuss the Gaussian approximations and their impact on the final estimate of the SI values. The ground truth against which the approximations are compared are SI values obtained by spatial numerical integration of Eq. (23). We shall address this issue through briefly replying to a series of questions: In what parameter regimes are the errors least, or most, severe? What stage in the approximation is this conclusion; yet it remains far from clear what stage in the Gaussian approximations cause these overestimations in the first place. Increased output could be attributed to the synapses being insufficiently depressed in the approximated system. Consideration of the mean synaptic depression at the onset of a tone (\( \tilde{x} \)) suggests the opposite. Fig. 7A shows that the approximation results in a slightly over-depressed mean synaptic state.

However, it must be recalled that the output depends not only on the synaptic state at the onset, but also on the synaptic state as it changes over the course of the input. Eqs. (23) and (24) show that the mean output (whether exact or approximate, respectively) depends on a time-integral of the synaptic state over the tone duration. Fig. 7B plots the fraction of local effective synapses averaged over the course of a tone. A comparison of the two profiles reveals no wholesale under- or overestimation of synaptic efficacy; rather, errors occur in different directions in different regions. Furthermore, synaptic depression is underestimated in the region around \( s = 0 \), where the response is measured. Thus, in summary, underestimates of synaptic depression lead to overestimated firing rates, which in turn lead to underestimated SI values.

4.3. Inhibition and nonlinear activation

In this work, we have assumed that the output function, \( \Phi \), which maps a mean sub-threshold potential to a mean firing rate, is a simply identity: \( \Phi(u) = u \). The trivial generalisation \( \Phi(u) = au \) does not impede the method, as \( a \) can at each stage be absorbed into the factors on the Gaussian components (A). Moreover, because there is only excitation in the model, \( u \geq 0 \); consequently, setting \( \Phi(u) = [u]^+ \) leads to the same behaviour as setting \( \Phi(u) = au \).

The technique encounters difficulties, however, when \( u \) can assume negative values (due to inhibition, for example), or \( \Phi \) is nonlinear for \( u < 0 \). There is experimental evidence that both inhibition and thresholding contribute to SSA. The application of the inhibitory blocker gabazine reduces SSA in...
the rat IC (Malmierca et al., 2010), and bicuculline abolishes SSA-like responses in ex vivo networks (Eytan et al., 2003). It is not clear how the current analytical approach would withstand the introduction of inhibitory connections. At the very least, the equation for the equilibrium synaptic profile (8), requires that \( v'(s) \) be non-negative, and a nonlinear output function (e.g., sigmoidal or a linear threshold) is needed to ensure this.

One means to circumvent these difficulties is to postulate that the inhibitory antagonist has the effect of broadening the width of the inputs \( (\sigma w) \); inhibition would then effectively be accounted for prior to the model. SSA in the neural field model is reduced as the inputs become less sharply tuned, and this is consistent with theories that hold that frequency tuning depends on a balance of excitation and inhibition. For a discussion, see de la Rocha et al. (2008); Wehr and Zador (2003).

5. Conclusion

This article analyses a neural field model of stimulus-specific adaptation. Neural field models allow quantities such as the rate of firing and resource depletion to be expressed as functions of continuous spatial variables. This model consists of an input layer and an output layer, which are connected by depressing synapses. We have derived an equation that characterises the expected synaptic effectiveness as it varies over the course of a tone presented in an oddball sequence, given the statistics of the sequence. The response at a point in the output layer is a weighted sum over activity in the input layer, where the input is modulated by a synaptic effectiveness profile. Consequently, SSA is generated in the output layer as the result of the synapses associated with localised inputs undergoing depression due to repeated presentations of a stimulus.

The expressions for the mean synaptic states assume a form that is difficult to integrate. However, the qualitative forms of the expressions are bumps and notches that can be replaced by Gaussian approximations with similar characteristics (i.e., heights, widths and centres). Making these adjustments leads to a set of expressions that can straightforwardly take part in sums, products and convolutions. With these approximations in place, the average response to a tone presented in an oddball sequence, given the statistics of the sequence, can be computed, and from there, an SI value is calculated. The solutions obtained when Gaussian approximations are used are close to those obtained by numerical integration, if the inputs are weak, or synaptic resources become ineffective slowly. When synapses depress rapidly, the effectiveness profile assumes a square-like shape, which is poorly approximated by a single Gaussian curve. However, the method is open to using several Gaussians to better approximate this kind of shape.

Acknowledgments

This research was supported by the European Community’s Seventh Framework Programme, Grant no. 231168–SCANDLE: “acoustic SCene ANalysis for Detecting Living Entities” (http://www.scandle.eu). RM wishes to thank Abul Azad for some helpful discussions, and three anonymous reviewers for their insightful remarks.

Appendix A. Solution to effective transmitter equation

The differential equation \( \dot{x} = b_1(1-x) - aux \), with positive constants \( a, b \) and \( v \), has the general solution

\[
x(t) = \frac{b}{av + b} + e^{(-av+b)t} \left( x(0) - \frac{b}{av+b} \right).
\]

If \( v \) is a time-varying pulse, such that \( v(t) = v^k \) on the interval \( [0,T_1] \), and \( v(t) = 0 \) on the interval \( (T_1,T_2) \) then

\[
x(T_1) = \frac{b}{av^k + b} + e^{(-av^k+b)T_1} \left( x(0) - \frac{b}{av^k+b} \right).
\]

\[
x(T_2) = 1 + e^{-bT_2} (x(T_1)-1).
\]

Appendix B. Substitution of bowl-shaped functions

Bowl-shaped functions arise naturally when point-like inputs are applied to a field of depressing synapses. One such function that appears in this paper is \( b/c_o \), which, in its expanded form, is written

\[
b/c_o = \frac{b}{Aa + b} + \frac{1}{Aa + b} = 1 - Y.
\]

For our purposes, it is more convenient to consider \( b/c_k \) as a bump-shaped profile, \( Y \), subtracted from a constant.

It is readily verified that the following Gaussian has the same zero-th, first and second derivatives at \( s = 0 \) as \( Y \) does:

\[
g_Y(s) = \left[ \frac{Aa}{Aa + b} \right] \exp \left( \frac{s^2}{-2u^2(Aa + b)} \right).
\]

However, as \( Y \) is flat-topped, the curvature at the peak fails to represent the curve as a whole, and the Gaussian substitutions are in general too broad. Instead, we opt to match the peak, as above, but use the half-peak-value intersections either side of centre as a guide to width (see Section 3.3.3). The curves that result from these procedures are shown in Fig. B.8A for an exemplary parameter set.

Similarly, for the nested exponential expression \( e^{-at} \), one can make the Gaussian approximation, such that the derivatives in the base of the bowl match:

\[
e^{-at} = \exp(At) \exp \left( \frac{s^2}{-2u^2(At + b)} \right) (1 - g_Y) e^{-bt}
\]

where

\[
g_Y(s) = \left[ 1 - e^{-at} \right] \exp \left( \frac{s^2}{-2u^2(1/b - a)} \right).
\]

The result of using the derivative-matching procedure is shown in Fig. B.8B.
REFERENCES


