

Processing of sounds by population spikes in a model of primary auditory cortex

Bomin Kim¹, Yanpin Luo¹, Kamyab Mirhosseini¹, Laszlo Morgan¹, Paria Naderi Nasab¹
¹University of Nottingham, School of Psychology, Nottingham, United Kingdom

The primary auditory cortex plays a crucial role in localising and processing complex sounds. The tonotopic organisation of A1 is mapped such that neurons with similar frequencies are oriented in separate iso-frequency bands (i.e., tonotopic axis). In this study, we replicate a model whereby each iso-frequency column is described by a recurrent neural network with short-term synaptic depression. In particular, this network generates population spikes, in which a group of neurons synchronously fire for a short time period. To mimic the tonotopic organisation of A1, frequency bands are inter-connected where population spikes will disseminate from one column to neighbouring columns depending on the input's characteristics. Consequently, incoming sounds are processed through sequences of population spikes that contain distinct information about simple and complex tones. Furthermore, we demonstrate the effect of noisy sensory input to the propagation and response of population spikes. We replicate frequency tuning curves of iso-frequeuncy columns, which rely on the strength of intra-cortical inhibitory and excitatory connections. We also show that consecutive two-tone stimuli exhibit forward masking which are contingent on spatial and temporal properties of the inputs, alongside PS-based encoding scheme for processing complex sounds (i.e., different amplitude and frequencies).

Keywords: Primary auditory cortex (A1), recurrent neural networks, synaptic depression

INTRODUCTION

Decomposition of simple and complex sounds into perceptually distinct frequency components is a fundamental component of hearing. There is an ongoing debate on the degree to which the primary auditory cortex (A1) responses are non-linear (Machens et al., 2004). At the cortical level, sound processing can be well modelled using linear approximation as neurons are well attuned to specific aspects of spectro-temporal features of sound. Conversely, by definition, linear models seem to not provide a complete representation of the process but rather an adequate approximation. Several lines of evidence show that A1 neurons display non-linearities when responding to simple and complex sounds. Specifically, previous studies have found that linear models, such as the linear spectro-temporal receptive field, are insufficient in fully capturing the underlying mechanism and the neural computation in sensory systems (Linden et al., 2003). However, it is important to note that existing simplified non-linear mechanisms (e.g., adapting to sound intensity and rectification) also fail to explain these complex behaviours in A1 neurons (Machens et al., 2004) as not all response reductions to stimulus repetition are related to adaptation (Pérez-González & Malmierca, 2014). Shechter and Depireux (2010) have characterised that most A1 neurons in an awake ferret respond non-linearly to the time-course and frequency range of complex sounds; non-linearity is not an exclusive result of response rate saturation and half-wave rectification. Additionally, the suppression of A1 responses - which can persist for hundreds of milliseconds after the tone onset - cannot be explained solely by synaptic inhibition (i.e., the postsynaptic GABAergic inhibition provided by the inhibitory synapses (Calford & Semple, 1995)), as the duration of inhibitory activity lasts maximally for 100ms (Wehr & Zador, 2005). Meanwhile, synaptic depression acting on the thalamocortical or intra-cortical synapses explains

the forward suppression experienced by consecutive tone (Chung et al., 2002). Therefore, *higher* mechanisms, such as short-term synaptic depression, can be attributed to explaining the non-linear interactions between sound frequencies and time-dependent properties of the neural encoder (Las et al., 2005).

Most existing models considering synaptic depression in A1 are feed-forward (Elhilali et al., 2004), largely ignoring the role of intra-cortical circuitry and the synaptic inter-connectivity of A1 neurons (Atzori et al., 2001). Past studies have shown that recurrent synaptic excitation-inhibition loops induce fast, rhythmic activity modulated by the inhibitory decay time constant (Augustin et al., 2013). Examining the effects of synaptic depression on a single cortical column, Tsodyks et al. (2000) proposed a recurrent network to demonstrate the sensitivity of PS response to both stimuli and spontaneous background activity. They found that when the synapses are intra-cortical connected, the recurrent network emits synchronous population neuron activity, named Population Spike (PS), which is then terminated via synaptic depression. With strong recurrent connections, a periodic sequence of spontaneous PSs are formed, whereas, with weak recurrent connections, PSs are evoked via external inputs. In other words, when given a persistent sub-threshold excitatory input, integrated with a single onset PS, the network will subsequently exhibit increasing asynchronous activity. asynchronous activity refers to neurons exhibiting dissimilar frequencies (see Fig.1); synchronous activity is exhibited during a PS response, in which neurons simultaneously fire at a similar frequency. With a stronger input amplitude, an earlier onset of PS and higher firing peak are observed (Yarden & Nelken, 2017). This temporal characteristic is in line with the response of A1 neurons to stimuli, distinguished by a transient locked onset response followed by a continuous increase to firing rate (Nelken et al., 1999; Ulanovsky et

al., 2003). Moreover, as the input frequency increases, the network's ability to follow a periodic input with a PS response diminishes (Kilgard & Merzenich, 1999).

Following past research findings, Loebel et al. (2007) hypothesised that the synchronous timing of the neural firing of A1 in response to sound stimulus represents the concurrent activity of neuron populations sharing similar best frequencies (i.e., preferred frequency range for a given iso-frequency column). The temporal coherence in the neuronal population's activity originates from intrinsic properties of the intracortical connections (i.e., short-term synaptic depression). Consequently, similar responses of A1 neurons in their respective cortical column are predicted from the same collection of neurons (Chen & Jen, 2000; Kilgard & Merzenich, 1999; Wallace et al., 2005). They also predicted that PSs are crucial in sound processing in A1, where PSs laterally spread along the tonotopic map, carrying precise signals between cortical regions with high temporal accuracy and speed. Additionally, several recent studies utilising intracellular (Deweese & Zador, 2006; Las et al., 2005) and multi-unit recordings (Luczak et al., 2013; Harris et al., 2005) have suggested that synchronised synaptic inputs received by A1 neurons correspond to PSs, further validating the existence of PSs in A1. They further hypothesised that recurrent connections with short-term synaptic depression can explain various responses of neurons in A1, including forward masking of simple and complex sounds.

Aim and overview of the current study

The aim of this study is to reproduce a neural network model of A1 (Loebel et al., 2007) with recurrent intra- and inter-columnar connections, incorporating synaptic depression. In other words, the activity of a neuron is not only dependent on itself but also the weighted sum activity of its neighboring neurons and columns. This type of network provides further insight into the sensitivity of the PS response to the sensory input, as well as to the spontaneous background activity from other cortical regions. Also, it allows for exploration into the mechanisms behind the forward masking phenomena to subsequent tones, observed in A1 (Wehr & Zador, 2005). Overall, this model provides an insight into the role of spontaneous activity in the auditory cortex, dynamic changes in response to various frequencies, and the thalamocortical connection patterns in A1.

MATERIALS AND METHODS

Modelling a Single Column

A single iso-frequency column is modelled by a recurrent network of A1, consisting of excitatory and inhibitory neurons whose rate dynamics are based on the Wilson-Cowan

model (1972). A neuron is expressed by its firing rate:

$$\begin{aligned} \tau_E \frac{dE_i}{dt} &= -E_i + (1 - \tau_{ref}^E E_i) \left[\frac{J_{EE}^0}{N_E} \sum_{j=1}^{N_E} Ux_j E_j + \frac{J_{EI}^0}{N_I} \sum_{j=1}^{N_I} Uy_j I_j + e_i^E + s_i \right]^+ \\ \tau_I \frac{dI_l}{dt} &= -I_l + (1 - \tau_{ref}^I I_l) \left[\frac{J_{IE}^0}{N_E} \sum_{j=1}^{N_E} E_j + \frac{J_{II}^0}{N_I} \sum_{j=1}^{N_I} I_j + e_l^I \right]^+ \end{aligned} \quad (1)$$

The number of excitatory and inhibitory neurons in a column is given by N_E and N_I , where the rate variables are E_i and I_l (i and l represent the indices of a neuron per column). τ_E (τ_I) denotes the time constant; the neurons' refractory period is given by τ_{ref} . The total input a neuron receives is dependent on the synaptic connection, represented by synaptic efficacies, J^0 terms (see Fig.2). e_i, e_l , the background synaptic inputs, are the effects of projections from other cortical regions. These values were randomly generated and selected from an uniform distribution with higher neuronal indices corresponding to bigger background inputs (i.e., $e_1^E < e_2^E < e_3^E$). It is important to note that only excitatory neurons received sensory inputs from the medial-geniculate, that is, the inhibitory neurons did not receive any sensory inputs nor experienced synaptic depression (see Eq.1).

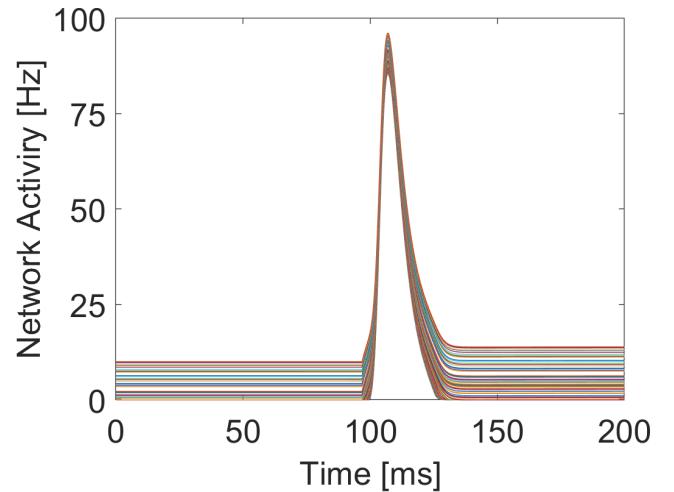


Figure 1. Asynchronous and synchronous activities of 100 excitatory neurons. Preceding and following a PS, neurons exhibit different frequencies (asynchronous), but during a PS response, neurons fire at similar frequencies (synchronous).

Synaptic depression was not introduced to inhibitory neurons as Tsodyks et al.'s (2000) experiment found that they do not significantly affect the PS response. Synaptic depression for excitatory and inhibitory neurons is described by the following equations:

$$\begin{aligned} \frac{dx_i}{dt} &= \frac{1 - x_i}{\tau_{rec}} - Ux_i E_i \\ \frac{dy_l}{dt} &= \frac{1 - y_l}{\tau_{rec}} - Uy_l I_l \end{aligned} \quad (2)$$

The parameters x and y are the maximum resources available for an action potential from excitatory and in-

hibitory neurons, of which the fraction of the amount available is determined by Ux and Uy , respectively. The resources recovered with a time constant τ_{rec} . These values come from somatosensory cortex experimentation data of Tsodyks & Markram (1997). τ_{ref} are approximated from synaptic conductance dynamics (Loebel et al., 2007). Loebel & Tsodyks (2002) found that PS responses could be evoked spontaneously if the excitatory connection strengths are high enough. Therefore, values of the J's are chosen to be below those values, so that only external inputs could evoke a PS response.

A threshold-linear gain function was applied for the inputs to the neurons (i.e., $[z]^+ = \max(z, 0)$). e and $s(t)$'s units were given in Hz and the synaptic efficacies, J 's, had no units.

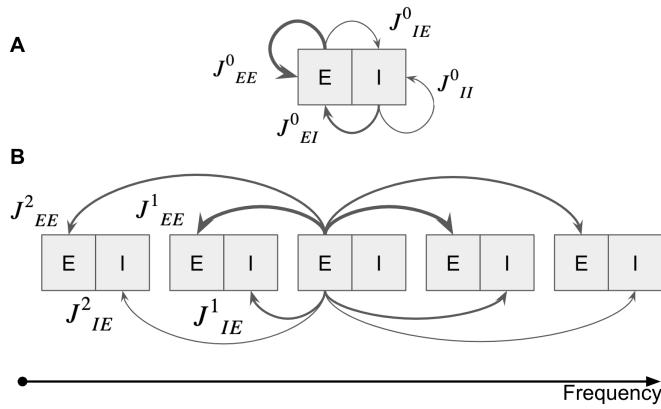


Figure 2. A schematic illustration of the iso-frequency column and A1 model. Adapted from Loebel et al. (2007). **A.** Illustrates one column in the network where excitatory and inhibitory neurons are connected. The neuronal connection strength is described by the width of the arrows. **B.** A representation of the tonotopic axis in the auditory cortex. Inter-columnar connections originate from excitatory neurons and target its two nearest neighboring columns.

Modelling the Primary Auditory Cortex (A1)

A1 is arranged tonotopically, with each iso-frequency column increasing in the preferred auditory frequency rate (Figure 2B). A neuron within column Q has intra-cortical connections J_0 to all other neurons in column Q , as well as inter-columnar connections to all neurons of column $Q + 1$ and $Q + 2$ (J_1, J_2). The strengths of the connections vary depending on the type of connection and the distance of the column, which take the form of:

$$\begin{aligned} \frac{dE_i^Q}{dt} &= -E_i^Q + (1 - \tau_{ref}^E E_i^Q) \left[\sum_{R=-2}^2 \frac{J_{EE}^{[R]}}{N_E} \sum_{j=1}^{N_E} Ux_j^{Q+R} E_j^{Q+R} + \frac{J_{EI}}{N_I} \sum_{j=1}^{N_I} Uy_j^Q I_j^Q + e_i^{E,Q} + \sum_{M=1}^P s_i^{Q,M} \right]^+ \\ \tau_I \frac{dI_i^Q}{dt} &= -I_i^Q + (1 - \tau_{ref}^I I_i^Q) \left[\sum_{R=-2}^2 \frac{J_{IE}^{[R]}}{N_E} \sum_{j=1}^{N_E} E_j^{Q+R} + \frac{J_{II}}{N_I} \sum_{j=1}^{N_I} I_j + e_i^{I,Q} \right]^+ \end{aligned} \quad (3)$$

The sensory inputs to A1 consist of a temporal and spatial component (see Eq. 4). For simplicity, sensory input is only given to the excitatory neurons as input given to the inhibitory neurons had no effect on the simulations (Loebel et al., 2007). The contribution to the neuron is taken as a sum of all of the inputs to the column, which propagate to

other columns via the following function:

$$s_i^{Q,M}(t) = A_M \zeta^M(t) e^{-\frac{|Q-M|}{\lambda_S(A)}}, \quad (4)$$

where $S_i^{Q,M}$ is the sensory input received by the i th neuron in column Q with best frequency (BF) of column M . $\zeta^M(t)$ is the temporal component which represents the onset and offset of the input. $h_i^{Q,M}$ represents the spatial component. A_M is the input amplitude for correspondent column (i.e., Besides column M, all columns are set to 0; see Fig.3). $\lambda_{s(a)}$ determines the localisation effect of the input. $\lambda_{s(A)}$ is dependent on the amplitude of A :

$$\lambda_S(A) = \begin{cases} \lambda_C & A \leq \alpha \\ \lambda_C + \frac{(A-\alpha)}{\delta} & A > \alpha \end{cases} \quad (5)$$

α is a threshold determining the sensory input's efficacy on PS's propagation to the neighbouring columns. When the amplitude is below α , the extent of the sensory input disseminating to the neighbouring columns is fixed. As amplitude increases above α , the decay of the signal become slower and propagates further. Additionally, δ determines the extent of localization of the sensory input at sound levels above α .

P is the number of columns in the recurrent neural network. Only neurons that displayed spontaneous activity were given sensory inputs because the thalamo-cortical inputs prefer to target neurons with non-zero spontaneous activity (Loebel et al., 2007). To do this, we ran the simulation from a negative time-point to zero seconds and observed which neurons exhibited non-spontaneous activity values (see results).

Table I. Parameter values used throughout the model.

Parameter	Value	Parameter	Value
N_E, N_I	100	τ_E, τ_I	1×10^{-3}
$\tau_{ref}^E, \tau_{ref}^I$	$3 \times 10^{-3}s$	τ_{rec}	0.8s
U	0.5	J_{EE}^0	6
J_{EI}^0	-4	J_{IE}^0	0.5
J_{II}^0	-0.5	e_1^E, e_1^I	-10Hz
e_{NE}^E, e_{NI}^I	10Hz	P	15
J_{EE}^1	4.5×10^{-2}	J_{EE}^2	1.5×10^{-2}
J_{IE}^1	3.5×10^{-3}	J_{IE}^2	1.5×10^{-3}
λ_C	0.25	α	2
$\delta_{left,right}$	5		

We randomly initialise state variables ($range : [0, 1]$) for both excitatory and inhibitory neurons. All simulations run using custom-written scripts written in MATLAB 2019a and an ODE solver from MATLAB was used to solve the rate equations at each time-step for excitatory and inhibitory neurons.

RESULTS

Basic properties of the model

Precedent to a PS response, excitatory and inhibitory neurons exhibit asynchronous activity at a steady state which is a non-zero value of a few Hz. The total receiving inputs (e.g., intra and inter-columnar, and background input from other brain areas) are relatively weak and do not reach the threshold required for the neurons to synchronously spike together. Yet, if the sensory input is strong enough for a given column, then it will emit a PS and spread out to neighbouring columns. PS spread depends on the spatial component of the sensory input and the inter-columnar connections (see Fig.4). Succeeding a PS response, the neurons reach a new steady state with increased asynchronous activity, usually slightly higher than the initial steady state before PS (see Fig.1).

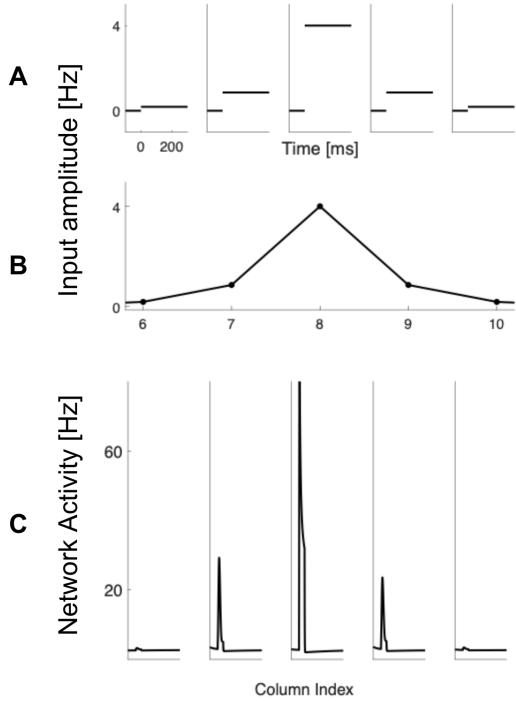


Figure 3. The model response to a continuous pure tone. The sensory input is given at time 0, with the BF of column 8. **A.** The temporal dependence of sensory input across auditory columns. **B.** The sensory input (product of temporal and spatial component; see Eq. 4). **C.** Mean activity of excitatory neurons in columns 6-10. The population spike response originates at column 8 and propagates to other columns via inter-columnar cortical connections.

Past studies have observed a temporal coupling to the PS response of excitatory and inhibitory synaptic conductance. The delay of the inhibitory population is independent to the spatial component of the input, so the intensity of the input is inconsequential (Las et al., 2005; Wehr & Zador, 2003; Zhang et al., 2003). The same findings were observed in the model, where the PS response firstly occurred in the excitatory populations, before recruiting the inhibitory population. Although the latency to the excitatory PS depends on both the temporal and spatial component of the sensory stimulus, the delay between the excitatory and inhibitory response is solely due to an intrinsic dynamic of the network (see Fig.5).

Frequency tuning curves

A frequency tuning curve (FTC) illustrates how the threshold to an input pure tone is proportional to the tone frequency. Evidently, it shows the minimum amplitude required to evoke a PS response to different tone frequencies.

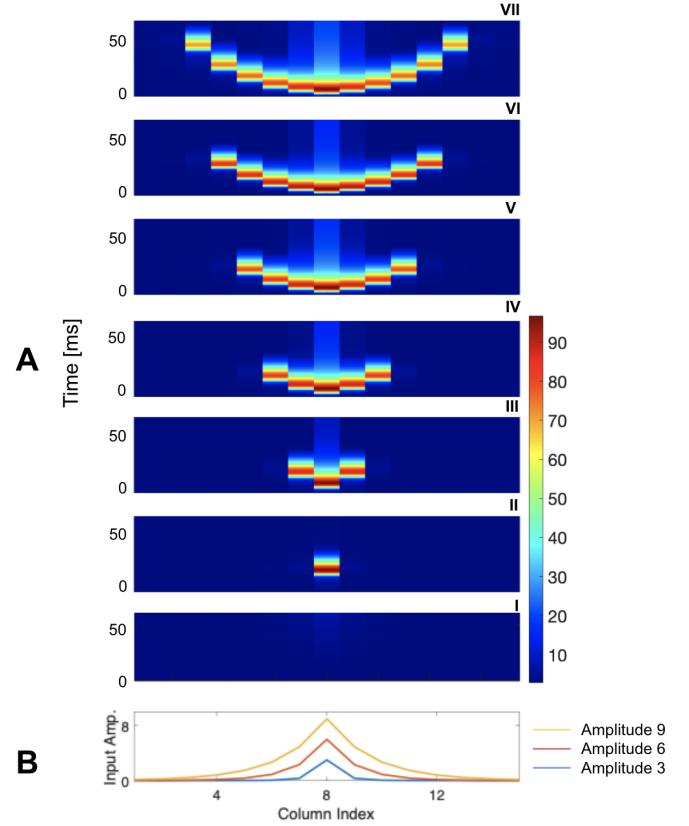


Figure 4. Spread of PS activity. **A.** A Colour Map of the PS activity with increasing input values. **B.** Spatial extent of three different stimuli corresponding to panels II, IV and VII.

The V-shape response of a FTC is due to the fact that columns whose BF response is furthest from the input amplitude, require a higher amplitude in order to respond, as spatial spread of the sensory input is gradually decreased by distance (see Fig.6).

We compare two different scenarios varying synaptic efficacies (J_{EE}^1 and J_{EE}^2), and observed that with the weak inter-columnar connections, the PS does not propagate as strongly to neighbouring columns, thereby higher inputs were required to evoke a PS response (see Fig.6A). This is because when the connections between the columns are weak, sensory input is the main stimuli to evoke the PS responses of neighbouring columns, as inputs received from the synapses of other neurons have been reduced. Furthermore, the symmetric effect of the tone input can be manipulated by modifying the parameter δ , which affects the localisation effect of the sensory input around the BF. δ_{left} and δ_{right} determine the spread of the input to the left or right columns. When these values are unequal, there is a non-symmetric effect of the tone input (see Fig.6B). Furthermore, the occurrence of the PS in the model is not only determined by the synaptic depression of the recurrent excitatory connections, but also by the recurrent inhibition. By decreasing the intensity of the recurrent

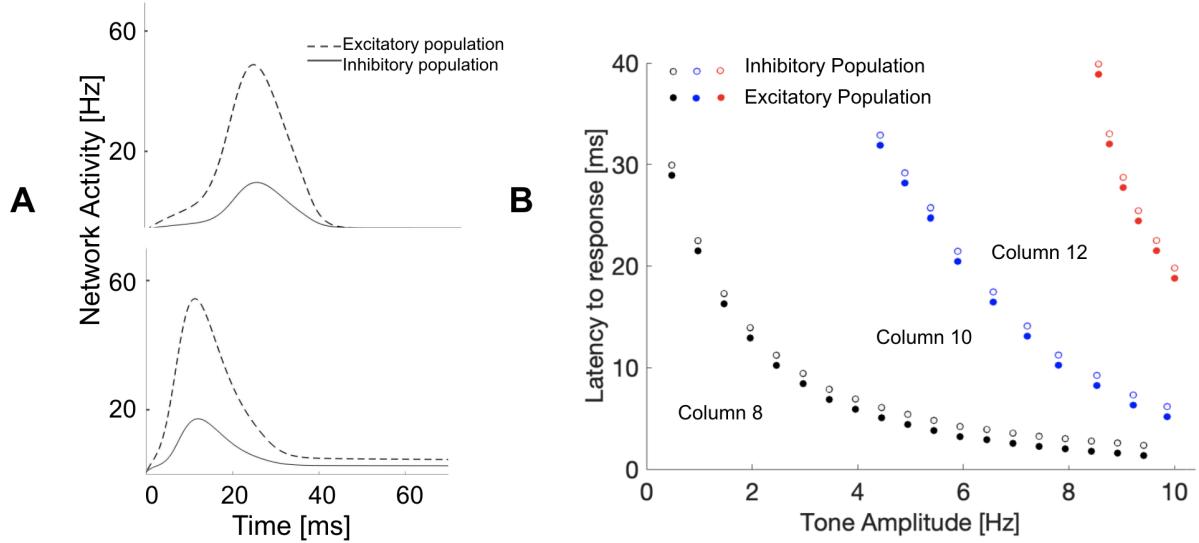


Figure 5. Temporal co-tuning of the excitatory and inhibitory population. **A.** Excitatory and inhibitory neuronal activity in the same column responding to a weak tone (top) and strong tone (bottom). The onset times for the excitatory PS depends on the strength of the input whereas the onset of the inhibitory PS just follows the onset of excitatory PS. **B.** The latency to the PS response for different tone frequencies in column 8 as a function of amplitude. As the BFs become more distant to column 8's, the PS's latency to threshold (20 Hz) increases. The inhibitory PS follows the excitatory PS with identical delay for each column and amplitude.

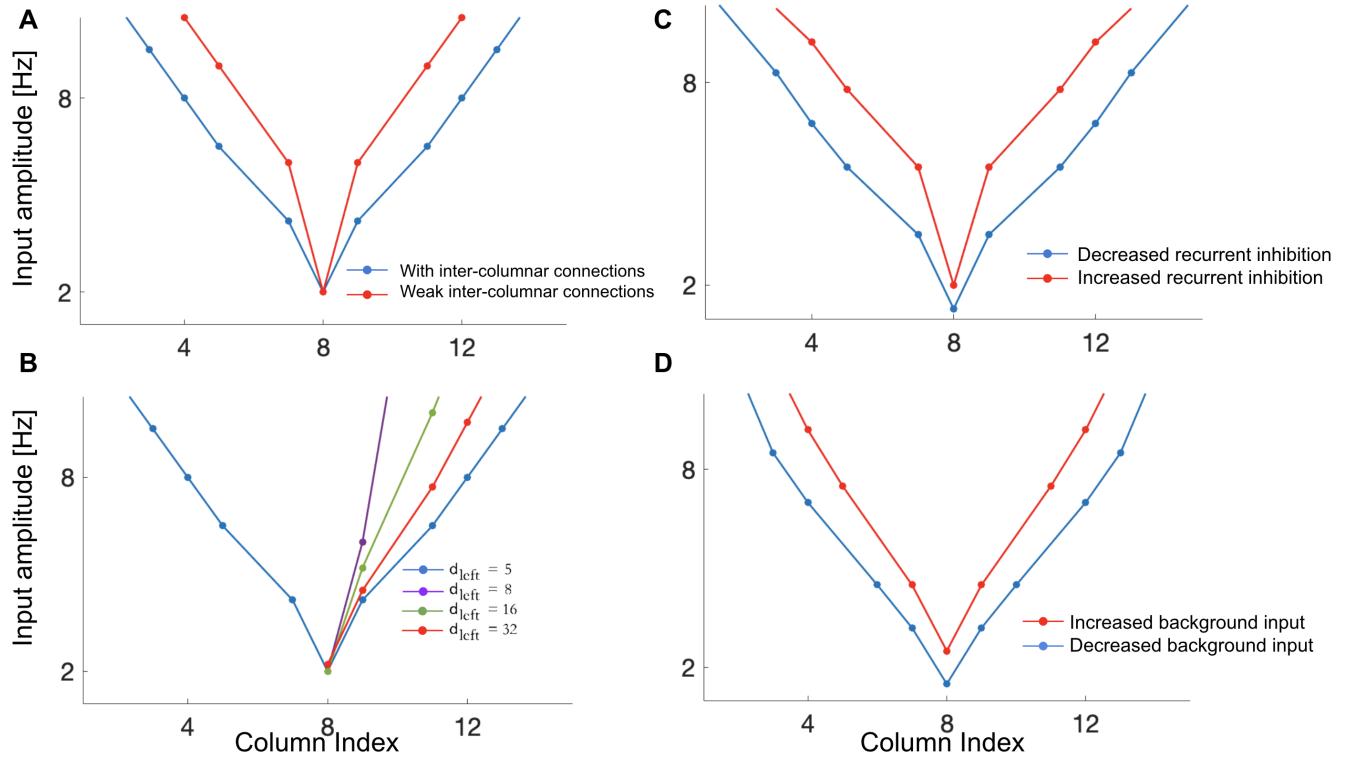


Figure 6. Frequency Tuning curves; **A.** FTC in networks with inter-columnar connections ($J_{EE}^1 = 4.5 \times 10^{-2}$, $J_{EE}^2 = 1.5 \times 10^{-2}$) and weak intercolumnar connections ($J_{EE}^1 = 4.5 \times 10^{-4}$, $J_{EE}^2 = 1.5 \times 10^{-4}$). **B.** Selecting different left values while keeping right constant at 5. **C.** Effects of varying recurrent inhibition by varying J_{EI}^0 , J_{II}^0 synaptic efficacies which are inputs from the inhibitory neurons (increased: $J_{EI}^0 = -1$, $J_{II}^0 = 2.5$; decreased: $J_{EI}^0 = -7$, $J_{II}^0 = -3.5$). **D.** Effects of background input modulation on the shape of FTC. Increased the distribution of e from -5 to 15 , and decreased it from -15 to 5 .

inhibition (J_{EI}^0 and J_{II}^0), the FTCs become wider, as the columns experienced less inhibition and exhibited PSs with lower amplitudes (see Fig. 6C). This is supported by studies where the addition of the antagonist GABA a led to a wider FTC of A1 neurons and a decreased threshold in response to different tones (Chen & Jen, 2000; Wang et al., 2002).

Rapid and reversible modulation of network responses by the changes in spontaneous activity
PS responses mainly depend on the excitatory connections within a column, therefore, a neuron which exhibits high

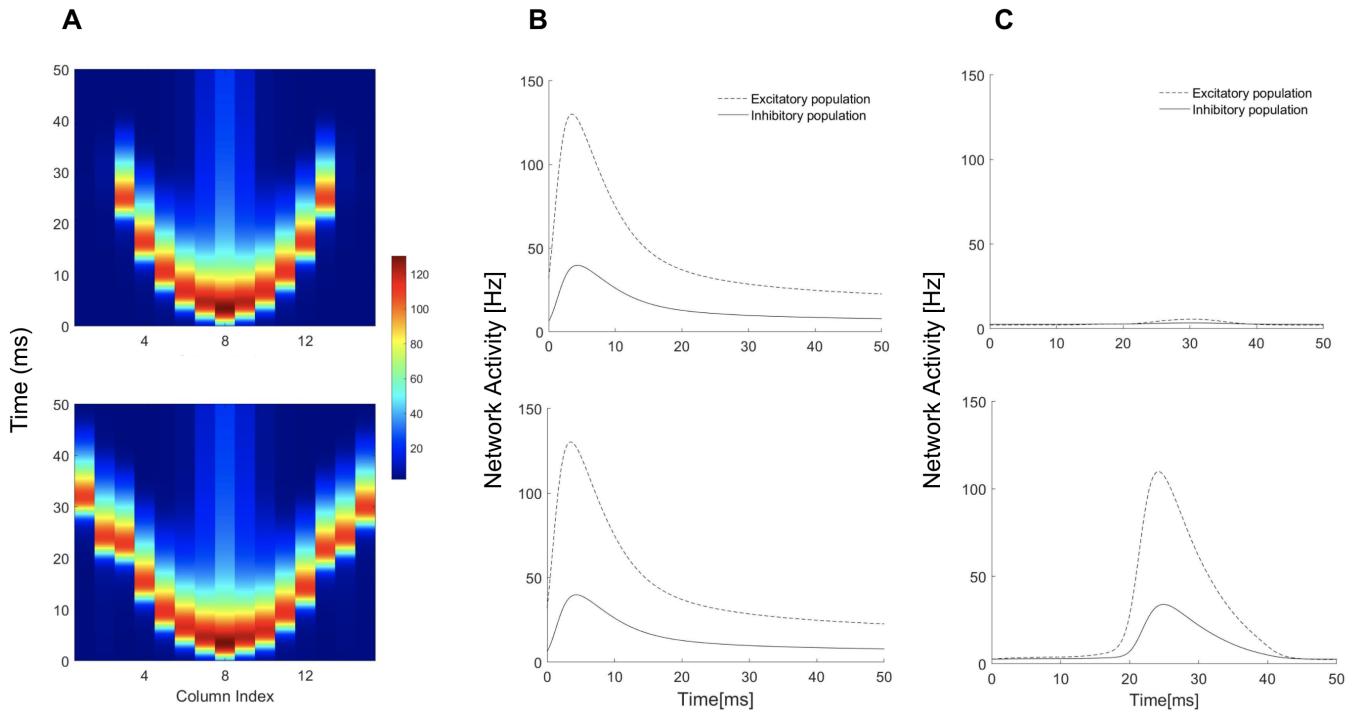


Figure 7. Introduction of background noise. **A.** A colour map comparing spread of PS activity in noisy (top) and non-noisy (bottom) sensory input. **B.** The activity of excitatory and inhibitory neurons to the BF of column 8 with noisy and non-noisy input. **C.** The activity of excitatory and inhibitory populations in column 2 to with noisy (top) and non-noisy (bottom) sensory input.

spontaneous activity will experience a weakened synaptic connection as synapses are in a depressed state, subsequently a high level of spontaneous activity. Furthermore, varying the distribution of the spontaneous activity in the excitatory neurons in a given column would result in dissimilar states of the synaptic connections. This in turn results in different sensory input responses. To test this, the distribution of the background inputs $e_i^{(E,Q)}$ and $e_l^{(I,Q)}$ were reduced by -5 . Figure 6D demonstrates that the FTC is wider, hence the threshold for an input amplitude is reduced, meaning that the PS can propagate further to other columns. If the background inputs return to the default values, then the spontaneous activity of the network will be return to its initial state along with the FTC; this may be a method for appropriately adjusting the background activity for rapid and reversible modulation of the response of the network to spontaneous activity.

Introducing noisy tone: further simulations

Noisy tone is introduced to mimic a more ecologically valid scenario, where a low-level noise is ubiquitous in forms of wind, tress, computer noise, etc. Introducing a noisy tone is achieved via setting each column, besides column 8, to a small random background amplitude (range: $0 - 0.5\text{Hz}$). Precisely, unlike typical pure tone input, the noisy tone has low-level noise in all columns except column 8. The noise is not time-dependent, and is present and constant for each simulation. This background sensory noise acts as an additional excitatory input to all columns, allowing even distant columns to emit PSs, as opposed to no noise. Figure 7 illustrates that when an input is presented to column 8, this input has the ability to spread to further columns when background noise is implemented, and

PSs appear in columns where there was minimal activity before. The noise changes the response of the model to an input frequency, with columns containing a high initial noise requiring less input to produce a population spike.

Two-tone interactions: forward masking

In A1, consecutive presentation of two tones elicits a forward suppression where the first sound stimulus (a masker) suppresses the signal of the second stimulus (a probe) for hundreds of milliseconds, a process termed forward masking (Wehr & Zador, 2005). Wehr & Zador (2005) showed that inhibitory conductance evoked by the first stimuli only last around $50 - 100\text{ms}$, whereas the effects of forward masking last around a few hundred milliseconds. This suggests that postsynaptic inhibition is not the only mechanism contributing to forward masking, but other processes such as synaptic depression could also be contributing. The excitatory synapses deplete more than 50% of resources after the presentation of the first stimulus and are strongly suppressed even until 500ms (see Fig.8B). Therefore, the forward suppression from the masking is due to the inhibitory conductance and can also be attributed to mechanisms such as synaptic depression.

The forward masking was simulated by presenting two identical tones to column 8, whereby both tones evoked a PS response when introduced individually. However, when the tones were presented in succession with a small interstimulus interval (ISI) (a unit measurement of τ_{rec}), a forward masking effect was observed, in which the second tone elicited substantially weaker PS response (see Fig.8A). A reason for this is that when there is high level of activity after a PS response, the synapses experience depression and require time to recover (which follows an exponential

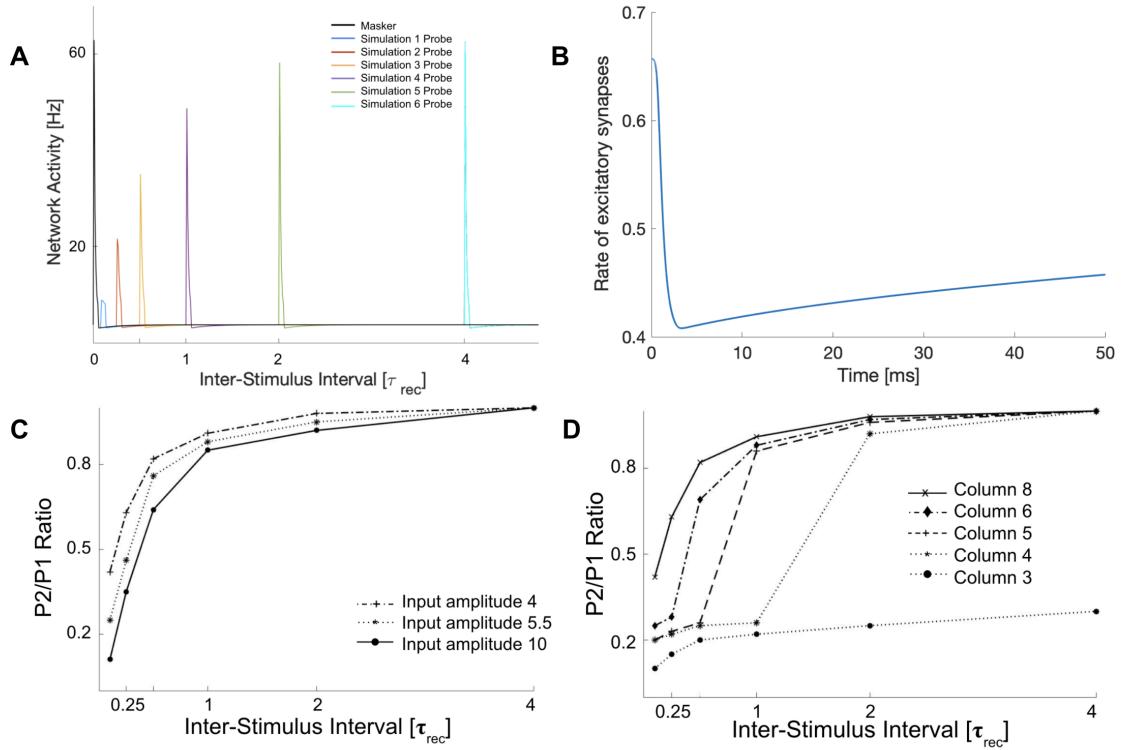


Figure 8. Forward masking. **A.** The activity of the network to tones with the BF of column 8, presented for a duration of 50ms, recorded from column 8. The first tone (masker) was always given at time 0, and the second tone (probe) was presented at different ISIs. **B.** The rate of excitatory synapses after an a sensory input was given to column 8 for 50ms. **C.** The recovery dynamics for different input amplitudes given as the ratio between the columnar responses to the first and second tone. **D.** The recovery dynamics at column 8 using input pairs with the BF of different columns.

shape (see Fig.8C)). This is in accordance with the claim that synaptic depression is a mechanism underlying forward suppression. The recovery dynamics for different input amplitudes follow a similar pattern, mainly because once a PS occurs in response to a masker, it is independent to the stimulus (Loebel & Tsodyks, 2002). Hence, the synaptic connections have similar amounts of depression with similar recovery dynamics across different amplitudes.

On the contrary, the recovery dynamic of the forward masking is dependent on stimulus tone frequency, that is, distant columns have a slower recovery dynamic (see Fig.8D). For instance, in Figure 8D, the second stimulus is presented at column 3, indicating that lower level of column 3 PS response propagates to column 8. This is both due to the depressed intra-columnar connections within column 3 and also the gradually decreasing PS level at the intra-cortical pathway. Additionally, this pattern is depicted when a masker is presented to column 4 followed by a probe in the same column.

The forward-masking modulates the level of response in column 8 because the intermediate columns experience suppression, and therefore, PS is restrained (see Fig.11).

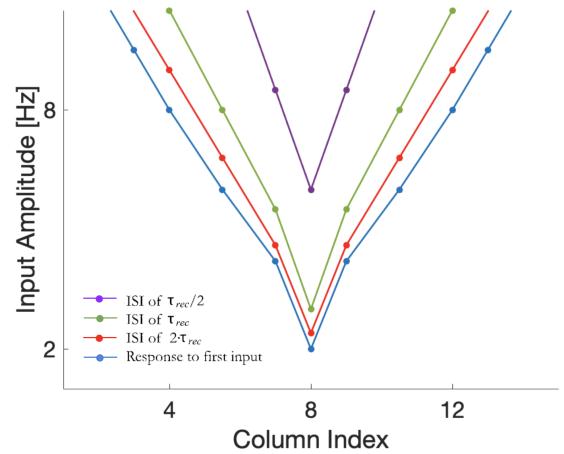


Figure 9. Receptive field response to different ISIs. Identical pure tones of varying amplitudes were presented with different ISIs to column 8. The minimum amplitude required for the second tone to evoke a PS response was recorded as the threshold.

Therefore, recovery from masking in A1 neurons do not follow similar patterns to each others'. Each A1 neuron recover has different recovery dynamics, where the neurons receiving the BF will recover first. Subsequently, all other columns' receptive fields will recover to its initial state. The FTC of the short ISI is much narrower because a higher amplitude is required to generate a second PS response (see Fig.9). If the second tone is followed with a small ISI, the synapses have not fully recovered, thus, requiring a higher amplitude to generate a second PS response. Specifically,

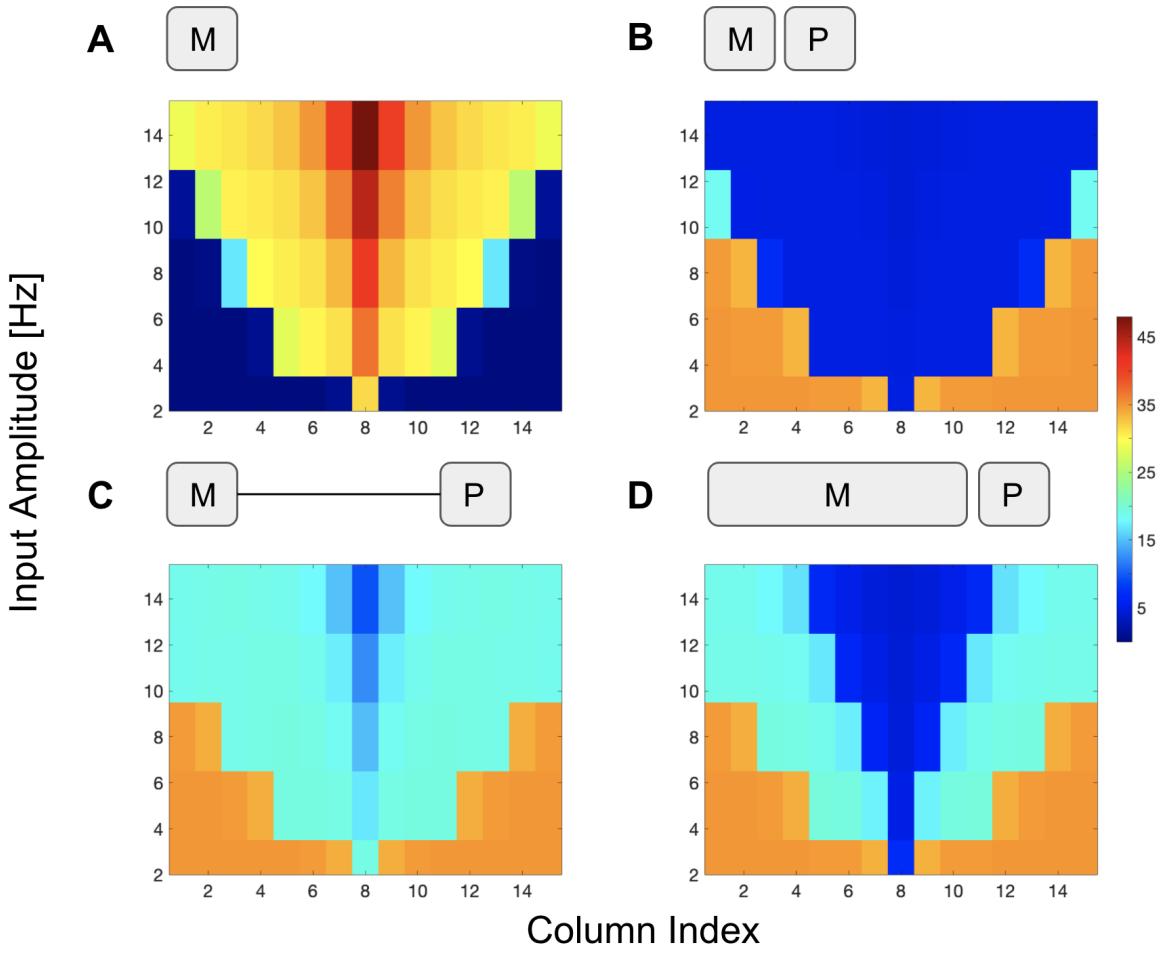


Figure 10. Forward masking with varied maskers; **A.** The responses of column 8 to a short masker (50ms) with varying frequencies and amplitude. The colour shows the average response of column 8 during the input as a function of frequency. **B.** The responses of column 8 to a probe (50ms) using the BF of column 8, when presented with a delay of 10ms. **C.** The responses of column 8 to a probe (50ms) using the BF of column 8, when presented with a delay of 400ms. **D.** Response to the probe (50ms), with the masker applied for 300ms with a 10ms delay. Stimulus configurations are shown above each panel.

longer the ISI, the closer the FTC resembles the response to just a single tone input. Therefore, it is demonstrated that the receptive field depends on the interval between the inputs, prohibiting response to a wide range of frequencies when the ISI is too small.

Forward suppression is also dependent on the duration of the masker and the interval between the masker and the probe. It is important to note that we present the masker in different columns but the probe is always presented to column 8. When the masker is presented for a duration of 50ms, distant columns require higher amplitudes to evoke a PS response in column 8 (see Fig.10A). However, when a probe immediately follows the short mask (presented after 10ms for a duration of 50ms), substantial suppression of the probe response is observed (see Fig.10B). With the presence of a longer delay (400ms) between the masker and the probe, column 8 is not suppressed as much from the mask in all columns (see Fig.10C) because the depressed synapses partially recover. When the masker is presented for a longer period (300ms) and then is immediately followed by a probe (with 10ms delay), suppression is much higher than the case in which a probe is presented after a long delay. The long masker results in increased asynchronous activity following a PS response (Loebel & Tsodyks, 2002), in which the synapses are

depressed and prevent a PS response to a probe. Maskers with BFs further away from the probe (i.e., column 8),

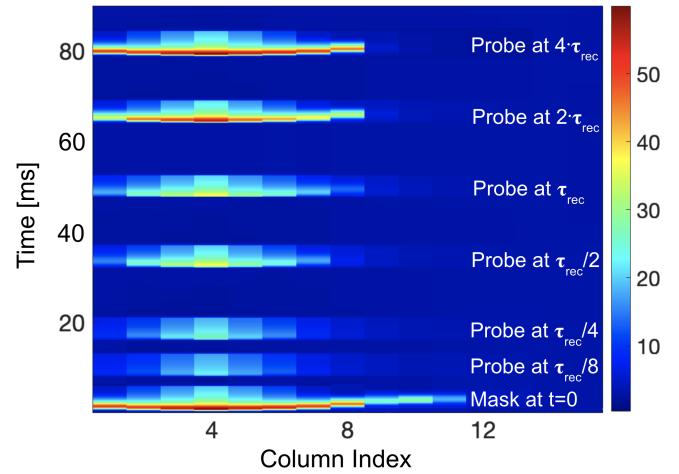


Figure 11. The dynamics of the recovery after forward masking depends on the intracortical pathway. The responses of the probes presented at column 4 with different ISI. The masker is also presented at column 4.

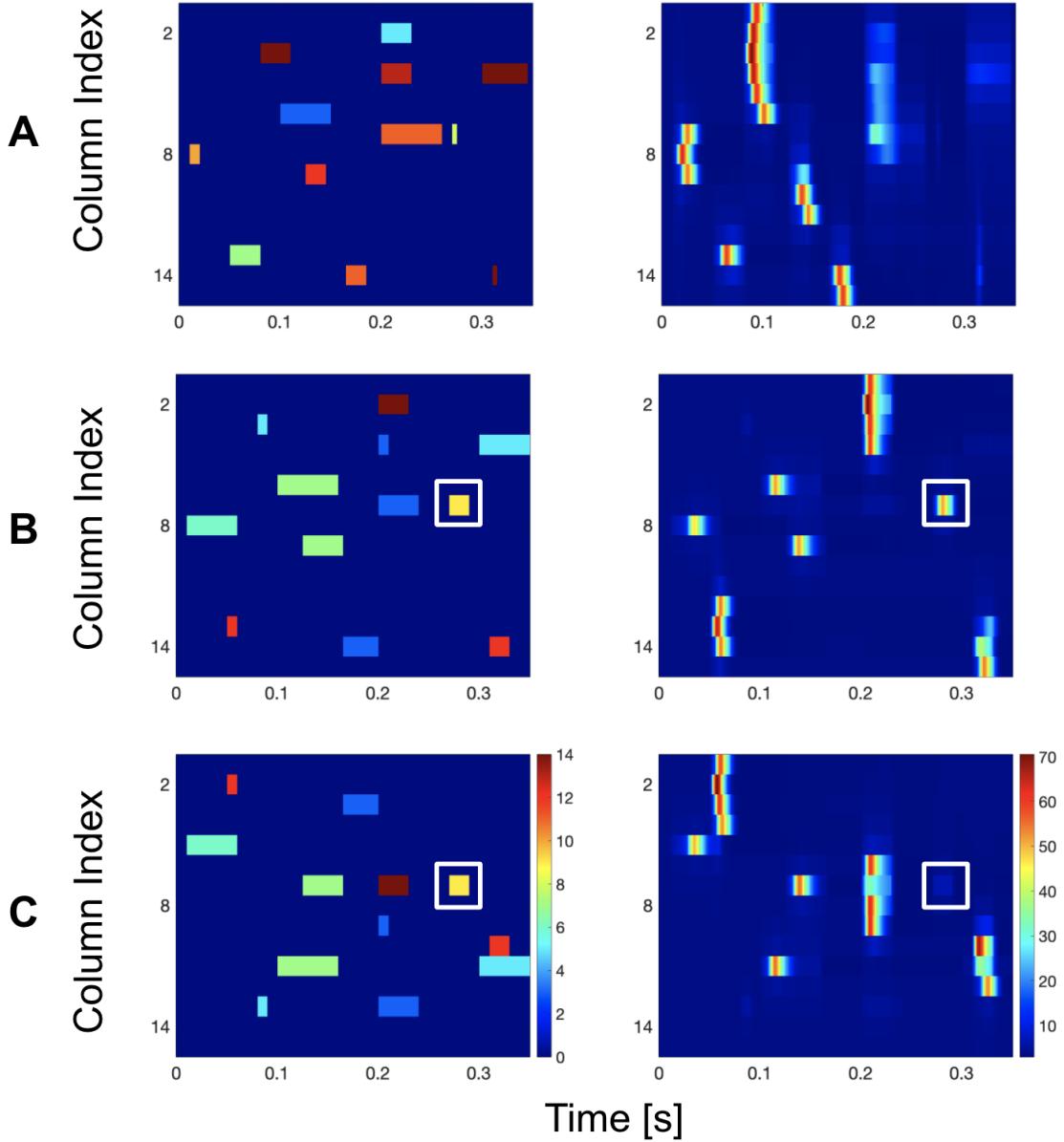


Figure 12. Illustration of varying PS encoding to complex sounds. Figures on the left represent the spectrograms of three different sounds; figures on the right illustrate the corresponding PS activity pattern output. In **A** and **B**, the onset times and frequencies of the sound were kept constant, but their duration and amplitudes were varied. In **B** and **C**, the onset times, duration and amplitudes of the sounds were kept constant, but the frequencies were varied. The white square demonstrates a tone pip which was identical in **B** and **C** but only evoked a strong PS response in **B**.

exhibit high enough asynchronous activity to maintain the depressed state of the synapses, limiting subsequent PS (see Fig.10D). Furthermore, if the response to the probe solely depends on the delay between the masker and the probe, we would expect Figure 10D to resemble Figure 10B, however, the suppression is moderately lower in Figure 10D. This could be explained by the recovery of the synapses with longer maskers. To test this, we plotted the dynamics of excitatory synapses and found that for 50ms masker, the excitatory synapses did not deplete as much as the 300ms. We also observed that with the 300ms masker, the excitatory synapses recovered to almost its initial value after 2 seconds (figures not show). Therefore, a longer masker recovers faster, enabling the neurons to respond more to a probe. Evidently, forward masking depends on both the duration of the mask but

also the interval between the two tones.

PS-based encoding of complex sounds

Given the behaviour of the PS propagation along the tonotopic map of A1 and its sensitivity to the spatial and temporal components of the input stimuli, Loebel et al. (2007) suggested a PS-coding scheme of complex sounds. That is, this encoding configuration precisely transforms complex input patterns into temporal A1 responses. We formulated three complex sounds with varying range of frequencies, amplitudes, onset time, alongside duration of the time. Each sound produced a unique response along the tonotopic map of A1 exhibiting PS-based encoding scheme, such as two-tone masking. It is important to note that the response to the complex sound cannot be inferred from just the response of the column to the component (Bar-Yosef

et al., 2002) because a component may or may not evoke the exact same response based on the inputs preceding it. Evidently, the tone pip (demonstrated by a white box), induces a higher PS response in Figure 12B than in Fig 12C, due to the suppression it experiences from the tone preceding it. If the onset times, duration and the amplitude are identical, but frequency is different (e.g., blue and maroon boxes in Fig. 12B and 12C), masking can occur (white rectangle).

DISCUSSION

In the current study, we qualitatively replicate the findings of Loebel et al. (2007) by simulating a fully-connected recurrent neural network model of the primary auditory cortex, with inclusion of depressing intracortical synaptic connections. Firstly, we investigated the response of the columns to different input amplitudes and showed that the threshold for the PS response to different sensory inputs ($s_i^{Q,M}(t) = A_M \zeta^M(t) e^{-\frac{|Q-M|}{\lambda_S(A)}}$) depended on the strength of the inter-columnar connections (see Fig.6A), the recurrent inhibition (see Fig.6B, and the localisation of the sensory input (see Fig.6C) and background input from other brain regions (see Fig.6D). In detail, modulation of background input could be a mechanism which allows for the regulation of rapid and reversible cortical processing. Secondly, we show that with longer intervals between two consecutive tones, the forward masking effect diminished as more synapses had recovered from the earlier synaptic depression (see Fig.10C). The time course of recovery is independent to the input amplitude, but strongly dependent on the frequency of the input. Additionally, forward masking varied as a function of the duration of the masker and the delay between the masker and the probe (see Fig.10). Finally, expanding from a simple sound context, we observed that by presenting different complex sounds to the A1, the model generated unique PS activity patterns along the tonotopic map (see Fig.12). These findings elucidate the intricate interplay between asynchronous neural activity and the varying conditions (e.g., intracolumnar connections, τ_{rec} , input amplitude) to generate population spikes, which is also dependent on synaptic depression. In conclusion, we show that a recurrent model is more sensitive in capturing the processing of sound in A1 than simple feed-forward network.

Differences from Loebel et al. (2007)

Our results predominantly replicate Loebel et al. (2007), yet there is one key consideration to note. In the current study, many of our parameters are set randomly, such as the initial states of excitatory and inhibitory neurons, and the generation of background input. The lack of explicit detail on the original simulation (e.g., initial states, and input amplitude (Hz)) is a limitation in replicating the exact findings. That is, the results in our paper may not be precisely consistent and quantitatively identical, but we successfully demonstrate the qualitative pattern and intrinsic properties, such as different PS propagation as input amplitude changes. In order to mitigate randomness within different simulations, we kept track of random seed generators, yet some were run with different seeds. However, despite employing different random seeds, we confirmed that similar patterns were generated by running the same simulation multiple times. In the future, the

suggestion is to keep the random seed consistent throughout the study when randomness is introduced. Secondly, another key difference between our study and Loebel et al. (2007) is the arbitrarily threshold set to detect PS (in the current study, 20Hz). There are multiple different thresholds to identify synchronous activity, such as the peak of the PS, the onset of the PS, or even an arbitrary value that induces population spikes (i.e., 20 Hz or 40 Hz). That is, we decided to set the threshold of detecting PS to 20 Hz because lower frequencies correspond to asynchronous activity (see Fig.1).

Limitations and future directions

There are several limitations of the present study that should be noted. First, this model is a basic recurrent neural network, focusing on the intrinsic dynamics of A1, and does not take into account the entire auditory cortex (Wrigley & Brown, 2004; Loebel et al., 2007). For instance, expansion of the model to the entire auditory cortex, incorporation of A2 and A3 would be an interesting next step to gain a deeper insight into presumably more complex and interactive coupling of the synaptic depression. Second, the parameters for excitatory and inhibitory neurons are largely identical (e.g., $\tau_E = \tau_i, e_1^E = e_1^I$), making excitatory and inhibitory neurons to have the same dynamics. In particular, the connection among neurons per column and the inter-columnar are fixed, not fully depicting the biological features in in-vivo neurons. Additionally, the number of excitatory and inhibitory neurons are identical so that all columns have the same number of corresponding neurons. Again, it is unrealistic for biological organisms to have the exact number of excitatory and inhibitory neurons in tonotopic gradients in A1. Therefore, an interesting future direction would be to allow some flexibility in the parameters so the model is representative of the plastic and volatile nature of the brain.

Additionally, the model is simplified, simulating in the most rudimentary dynamics and connections, not considering any external variations such as the development stage of the cortex, specifically in the A1. Interestingly, similar to other high-level cognition processes like language, the auditory process exhibits experience-dependency in which there is a postnatal period of sensory input-driven plasticity (Villers-Sidani et al., 2007). “Hebbian rule”, a principle underlying neuronal plasticity states that connected neurons that “fire together, wire together” (Hebb, 1949). That is, to strengthen the mutual connections between neurons, one neuron persistently (consistency) partakes in firing (causality) the other, subsequently yielding brain plasticity (Gustafsson et al., 1987). A recent study examining young infant mice has found relatively weak ability for distinct frequency selectivity to the onset (ON) and offset (OFF) of sounds (Sollini et al., 2018), presumably due to the premature cortical development of the auditory cortex. They found that with the addition of a novel implementation of Hebbian plasticity model, it is sufficient in the development of such ON and OFF frequency tuning. That is, this simple learning rule can provide further detail in the development of topographic organisation in the auditory cortex. In this model, neighbouring columns connected in which long-distance connectivity (i.e, distant columns have weaker connectivity) are suppressed due to the intrinsic properties of the topographic configuration of the A1. Yet, it is unclear why some tones are suppressed due to re-

peated exposure of similar frequency, insinuating a latent learning mechanism may be involved in the reinforcement process of Hebbian theory. It would be interesting to examine the interplay between frequency tuning selectivity, different stages of A1 development, alongside PS dynamics (e.g., forward-masking) to gain further insight into the intricate process of transforming auditory sounds into neuronal activities.

Neuronal adaptation to specific stimuli is ubiquitous in the neuronal system. That is, with more repeated exposure to the same stimulus, there is a decrease in response (Yarden & Nelken, 2017; Augustin et al., 2013). Ulanovsky et al. (2003) presented a sequence of repeating sounds which was then interrupted with a novel “unexpected” sound (varied in both frequency and amplitude). They found that stimulus-specific adaptation occurs, with adapted neurons exhibiting a lower activity response for the repeating sound than the deviant sound, suggesting that in addition to processing acoustic features, A1 is implicated in novelty detection and sensory memory. Connecting to our current study, the forward-masking effect we observed could be a factor in the reduced cortical activity in response to the successive repeated sound. The “unexpected” noises could be an interesting future direction to observe the subsequent effect of tone habituation varying frequency and amplitude on the onset and propagation of PS.

Indeed, we examine the effect of noisy data on the onset and propagation of PS. Yet, our current simulation is not time-dependent, that is, the noise is constant with identical pure tone, lacking fluctuation throughout time (i.e., complexity). It still remains unclear the intricate interaction between dynamic noisy tone (i.e., time-dependent) and PS responses. In other words, noisy data can be presented simultaneously or randomly alongside both pure tones and other noisy tones, determining the synaptic depression, subsequently the PS response. Particularly, with the sudden presence of noisy data at one time-point, we speculate the noisy tone to have a stronger masking effect on the subsequent tone as the noisy input has a stronger propagation of PS.

REFERENCES

- Augustin, N. H., Cummins, R. P., & French, D. D. (2001). Exploring spatial vegetation dynamics using logistic regression and a multinomial logit model. *Journal of Applied Ecology*, 38(5), 991–1006. <https://doi.org/10.1046/j.1365-2664.2001.00653.x>
- Bertalmío, M., Calatroni, L., Franceschi, V., Franceschiello, B., Gomez Villa, A., & Prandi, D. (2020). Visual illusions via neural dynamics: Wilson-cowan-type models and the efficient representation principle. *Journal of neurophysiology*, 123(5), 1606–1618. <https://doi.org/10.1152/jn.00488.2019>
- Bressloff, P. C., Cowan, J. D., Golubitsky, M., Thomas, P. J., & Wiener, M. C. (2002). What geometric visual hallucinations tell us about the visual cortex. *Neural computation*, 14(3), 473–491. <https://doi.org/10.1162/089976602317250861>
- Buxhoeveden, D. P., Semendeferi, K., Buckwalter, J., Schenker, N., Switzer, R., & Courchesne, E. (2006). Reduced minicolumns in the frontal cortex of patients with autism. *Neuropathology and applied neurobiology*, 32(5), 483–491. <https://doi.org/10.1111/j.1365-2990.2006.00745.x>
- Calford, M. B., & Semple, M. N. (1995). Monaural inhibition in cat auditory cortex. *J Neurophysiol*, 73(5), 1876–1891.
- Chambers, J. D., Elgueda, D., Fritz, J. B., Shamma, S. A., Burkitt, A. N., & Grayden, D. B. (2019). Computational neural modeling of auditory cortical receptive fields. *Frontiers in computational neuroscience*, 13, 28–28. <https://doi.org/10.3389/fncom.2019.00028>
- Charles, C. L., & Sherman, S. M. (2008). Synaptic properties of thalamic and intracortical inputs to layer 4 of the first- and higher-order cortical areas in the auditory and somatosensory systems. *Journal of Neurophysiology*. <https://doi.org/10.1152/jn.90391.2008>
- Chen, Q. C., & Jen, P. H. S. (2000). Bicuculline application affects discharge patterns, rate-intensity functions, and frequency tuning characteristics of bat auditory cortical neurons. *Hearing research*, 150(1), 161–174. [https://doi.org/10.1016/S0378-5955\(00\)00197-0](https://doi.org/10.1016/S0378-5955(00)00197-0)
- DeWeese, M. R., & Zador, A. M. (2006). Non-gaussian membrane potential dynamics imply sparse, synchronous activity in auditory cortex. *The Journal of neuroscience*, 26(47), 12206–12218. <https://doi.org/10.1523/JNEUROSCI.2813-06.2006>
- Elhilali, M., Fritz, J. B., Klein, D. J., Simon, J. Z., & Shamma, S. A. (2004). Dynamics of precise spike timing in primary auditory cortex. *The Journal of neuroscience*, 24(5), 1159–1172. <https://doi.org/10.1523/JNEUROSCI.3825-03.2004>
- Ermentrout, G. B., & Cowan, J. D. (1979). A mathematical theory of visual hallucination patterns. *Biol Cybern*, 34(3), 137–50. <https://doi.org/10.1007/BF00336965>
- Fakhraei, L., Gautam, S. H., & Shew, W. L. (2017). State-dependent intrinsic predictability of cortical network dynamics. *PloS one*, 12(5), e0173658–e0173658. <https://doi.org/10.1371/journal.pone.0173658>
- Frédéric, P., & Massimo, S. (2001). Enforcement of temporal fidelity in pyramidal cells by somatic feed-forward inhibition. *Science (American Association for the Advancement of Science)*, 293(5532), 1159–1163. <https://doi.org/10.1126/science.1060342>
- Fritz, J., Elhilali, M., & Shamma, S. (2005). Active listening: Task-dependent plasticity of spectrotemporal receptive fields in primary auditory cortex. *Hearing research*, 206(1), 159–176. <https://doi.org/10.1016/j.heares.2005.01.015>
- Gabriel, S., Nancy, K., & Kamal, S. (2006). Network architecture, receptive fields, and neuromodulation: Computational and functional implications of cholinergic modulation in primary auditory cortex. *Journal of Neurophysiology*, 96(6), 2972–2983. <https://doi.org/10.1152/jn.00459.2006>
- Gustafsson, B., Wigström, H., Abraham, W. C., & Huang, Y. Y. (1987). Long-term potentiation in the hippocampus using depolarizing current pulses as the conditioning stimulus to single volley synaptic potentials. *J Neurosci*, 7(3), 774–780.
- Happel, M. F. K., Jeschke, M., & Ohl, F. W. (2010). Spectral integration in primary auditory cortex attributable to temporally precise convergence of thalamocortical and intracortical input. *The Journal of neuroscience*, 30(33), 11114–11127. <https://doi.org/10.1523/JNEUROSCI.0689-10.2010>
- Jennifer, F. L., Robert, C. L., Maneesh, S., Christoph, E. S., & Michael, M. M. (2003). Spectrotemporal structure of receptive fields in areas ai and aaf of mouse auditory cortex. *Journal of Neurophysiology*, 90(4), 2660–2675. <https://doi.org/10.1152/jn.00751.2002>
- Kamal, B., Holman, C., & de Villers-Sidani, E. (2013). Shaping the aging brain: role of auditory input patterns in the emergence of auditory cortical impairments. *Front Syst Neurosci*, 7, 52.

- Kang, H., Auksztulewicz, R., An, H., Abi Chacra, N., Sutter, M. L., & Schnupp, J. W. H. (2021). Neural correlates of auditory pattern learning in the auditory cortex. *Frontiers in neuroscience*, 15, 610978. <https://doi.org/10.3389/fnins.2021.610978>
- Khalighinejad, B., Herrero, J. L., Mehta, A. D., & Mesgarani, N. (2019). Adaptation of the human auditory cortex to changing background noise. *Nat Commun*, 10(1), 2509.
- Kilgard, M. P., & Merzenich, M. M. (1999). Distributed representation of spectral and temporal information in rat primary auditory cortex. *Hearing research*, 134(1), 16–28. [https://doi.org/10.1016/S0378-5955\(99\)00061-1](https://doi.org/10.1016/S0378-5955(99)00061-1)
- Kuo, S. P., & Trussell, L. O. (2011). Spontaneous spiking and synaptic depression underlie noradrenergic control of feed-forward inhibition. *Neuron (Cambridge, Mass.)*, 71(2), 306–318. <https://doi.org/10.1016/j.neuron.2011.05.039>
- Las, L., Stern, E. A., & Nelken, I. (2005). Representation of tone in fluctuating maskers in the ascending auditory system. *The Journal of neuroscience*, 25(6), 1503–1513. <https://doi.org/10.1523/JNEUROSCI.4007-04.2005>
- Liang, L., Wang, X., & Lu, T. (2001). Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nature neuroscience*, 4(11), 1131–1138. <https://doi.org/10.1038/nn737>
- Loebel, A., Nelken, I., & Tsodyks, M. (2007). Processing of sounds by population spikes in a model of primary auditory cortex. *Frontiers in neuroscience*, 1(1), 197–209. <https://doi.org/10.3389/neuro.01.1.1.015.2007>
- Loebel, A., & Tsodyks, M. (2002). Computation by ensemble synchronization in recurrent networks with synaptic depression. *Journal of computational neuroscience*, 13(2), 111–124. <https://doi.org/10.1023/A:1020110223441>
- Luczak, A., Bartho, P., & Harris, K. D. (2013). Gating of sensory input by spontaneous cortical activity. *The Journal of neuroscience*, 33(4), 1684–1695. <https://doi.org/10.1523/JNEUROSCI.2928-12.2013>
- Machens, C. K., Wehr, M. S., & Zador, A. M. (2004). Linearity of cortical receptive fields measured with natural sounds. *The Journal of neuroscience*, 24(5), 1089–1100. <https://doi.org/10.1523/JNEUROSCI.4445-03.2004>
- Merzenich, M. M., Schreiner, C. E., Zhang, L. I., & Tan, A. Y. Y. (2003). Topography and synaptic shaping of direction selectivity in primary auditory cortex. *Nature (London)*, 424(6945), 201–205. <https://doi.org/10.1038/nature01796>
- Nelken, I. (2004). Processing of complex stimuli and natural scenes in the auditory cortex. *Current opinion in neurobiology*, 14(4), 474–480. <https://doi.org/10.1016/j.conb.2004.06.005>
- Pérez-González, D., & Malmierca, M. S. (2014). Adaptation in the auditory system: An overview. *Frontiers in integrative neuroscience*, 8, 19–19. <https://doi.org/10.3389/fnint.2014.00019>
- Rankin, J., Sussman, E., & Rinzel, J. (2015). Neuromechanistic model of auditory bistability. *PLoS computational biology*, 11(11), e1004555–e1004555. <https://doi.org/10.1371/journal.pcbi.1004555>
- Schreiner, C. E., Read, H. L., & Sutter, M. L. (2000). Modular organization of frequency integration in primary auditory cortex. *Annual review of neuroscience*, 23(1), 501–529. <https://doi.org/10.1146/annurev.neuro.23.1.501>
- Sedlacek, M., & Brenowitz, S. D. (2014). Cell-type specific short-term plasticity at auditory nerve synapses controls feed-forward inhibition in the dorsal cochlear nucleus. *Frontiers in neural circuits*, 8(JULY), 78–78. <https://doi.org/10.3389/fncir.2014.00078>
- Shechter, B., & Depireux, D. A. (2010). Nonlinearity of coding in primary auditory cortex of the awake ferret. *Neuroscience*, 165(2), 612–620. <https://doi.org/10.1016/j.neuroscience.2009.10.034>
- Simranjit, K., Ronit, L., & Raju, M. (2004). Intracortical pathways determine breadth of subthreshold frequency receptive fields in primary auditory cortex. *Journal of Neurophysiology*, 91(6), 2551–2567. <https://doi.org/10.1152/jn.01121.2003>
- Tsodyks, M., Uziel, A., & Markram, H. (2000). Synchrony generation in recurrent networks with frequency-dependent synapses. *J Neurosci*, 20(1), RC50. <https://www.ncbi.nlm.nih.gov/pubmed/10627627>
- Ulanovsky, N., Nelken, I., & Las, L. (2003). Processing of low-probability sounds by cortical neurons. *Nature neuroscience*, 6(4), 391–398. <https://doi.org/10.1038/nn1032>
- Vincent, B., Frédéric, C., Larry, G., & Yves, F. (1999). Horizontal propagation of visual activity in the synaptic integration field of area 17 neurons. *Science (American Association for the Advancement of Science)*, 283(5402), 695–699. <https://doi.org/10.1126/science.283.5402.695>
- Wallace, M. N., Shackleton, T. M., Anderson, L. A., & Palmer, A. R. (2005). Representation of the purr call in the guinea pig primary auditory cortex. *Hearing research*, 204(1), 115–126. <https://doi.org/10.1016/j.heares.2005.01.007>
- Wang, J., McFadden, S. L., Caspary, D., & Salvi, R. (2002). Gamma-aminobutyric acid circuits shape response properties of auditory cortex neurons. *Brain research*, 944(1), 219–231. [https://doi.org/10.1016/S0006-8993\(02\)02926-8](https://doi.org/10.1016/S0006-8993(02)02926-8)
- Wehr, M., & Zador, A. M. (2005). Synaptic mechanisms of forward suppression in rat auditory cortex. *Neuron*, 47(3), 437–445.
- Wilson, H. R., & Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical journal*, 12(1), 1–24. [https://doi.org/10.1016/S0006-3495\(72\)86068-5](https://doi.org/10.1016/S0006-3495(72)86068-5)
- Wolfgang, M., Ursula, K., & Michael, H. (2005). Feed-forward inhibition shapes the spike output of cerebellar purkinje cells. *The Journal of physiology*, 563(2), 369–378. <https://doi.org/10.1113/jphysiol.2004.075028>
- Wrigley, S. N., & Brown, G. J. (2004). A computational model of auditory selective attention. *IEEE transactions on neural networks*, 15(5), 1151–1163. <https://doi.org/10.1109/TNN.2004.832710>
- Yarden, T. S., & Nelken, I. (2017). Stimulus-specific adaptation in a recurrent network model of primary auditory cortex. *PLoS computational biology*, 13(3), e1005437–e1005437. <https://doi.org/10.1371/journal.pcbi.1005437>
- Zador, A. M., & Wehr, M. (2003). Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature*, 426(6965), 442–446. <https://doi.org/10.1038/nature02116>