

Processing of Sounds by Population Spikes in a Model of A1

[AUTHOR REDACTED]^{1,*}, [AUTHOR REDACTED]^{1,*}, [AUTHOR REDACTED]^{1,*} and [AUTHOR REDACTED]^{1,*}

¹School of Mathematical Sciences, University of Nottingham, Nottingham, NG7 2RD, UK

*These authors contributed equally to this work

Abstract

Understanding the computational principles underlying the processing of sounds in the primary auditory cortex (A1) requires accounting for its pronounced non-linear response properties. Loebel et al. (2007) explores a model of A1 in which iso-frequency columns are realised as recurrent neural networks incorporating short-term synaptic depression. These dynamics generate transient, synchronous bursts of increased firing rates termed Population Spikes (PS), which propagate along the tonotopic map of A1. We successfully replicate the results presented by Loebel et al. (2007), and further build on their work, by probing three additional aspects of the model: (i) stochastic fluctuations in background inputs modulate the propagation of PS, and (ii) how variations in the signal-to-noise ratio (SNR) influence network activity and frequency response area (iii) the dynamics of stimulus-specific adaptation (SSA) in the model. Our results reveal that PS generation and propagation is highly sensitive to minor variations in background input, and the model mostly captures the key experimental trends observed in SNR studies.

1. Introduction

The primary auditory cortex (A1) is pivotal in transforming auditory stimuli into meaningful neural representations. Yet early linear models, particularly those relying on spectro-temporal receptive fields (Bar-Yosef et al. 2002; Machens et al. 2004), which are the response of neurons to sound frequencies and time, struggle to predict the non-linear responses elicited by natural stimuli. Loebel et al. (2007) provide a path to understanding these complex responses by modelling the tonotopic organisation of A1 with a rate-based paradigm (Figure 1). This model was able to accurately predict a range of responses to forward masking, hypersensitive locking suppression, frequency modulated sweeps, and complex sounds. In this report, we will replicate the key results of Loebel et al. (2007), providing comparisons to experimental data, as well as diagnosing problems with reproducibility, exploring the effect of altered signal-to-noise ratios and the dynamics of stimulus-specific adaptation (SSA).

1.1. Primary Auditory Cortex (A1)

Non-linear behaviours in (A1) include pronounced phasic responses, in which neurons generate a brief, high-amplitude population spike when a sound begins (Calford et al. 1995; Phillips et al. 1991; Bendor et al. 2005). Another salient example is hypersensitive locking suppression, where even weak pure tones can disrupt a neuron's ability to lock to the amplitude envelope of fluctuating noise (Las et al. 2005; Nelken et al. 1999). Such observations underscore the shortcomings of purely linear or strictly feed-forward models. Furthermore, synaptic inhibition, which lasts for roughly 100 ms, cannot fully account for the more enduring suppression reported in A1 (Brosch et al. 1997; Wehr et al. 2005).

Consequently, attention shifted to intracortical mechanisms, particularly short-term synaptic depression (STD) (Thomson and Deuchars, 1994; Fishbach et al., 2003; Elhilali et al., 2004) (Thomson et al. 1994; Fish-

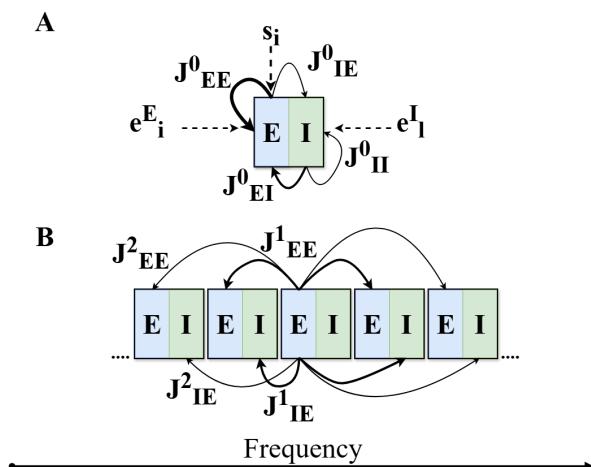


Figure 1. Schematic of recurrent neural network model of A1. (A) Each iso-frequency column consists of interconnected excitatory (E) and inhibitory (I) neuronal populations, with background inputs e_i^E , e_i^I , and sensory inputs s_i . Arrow thickness indicates relative connection strength. (B) Columns are arranged along the tonotopic axis, with inter-columnar connections originating from excitatory neurons and targeting neighbouring columns. Adapted from Loebel et al. (2007).

bach et al. 2003; Elhilali et al. 2004). Building on this, Loebel et al. (2007) proposed a recurrent neural network model, incorporating STD, to capture the nonlinearities observed in A1. In their framework, each iso-frequency column is a network of excitatory and inhibitory neurones connected by recurrent loops (Figure 1A). Because excitatory synapses are subject to STD, their transmission strength decreases with repeated activation.

The interaction between recurrent excitation, inhibition, and synaptic depression can give rise to a population spike (PS): transient, synchronised bursts of increased firing rates across the neural assembly. We define PSs in terms of firing rate as each neurone is modelled based on the rate model of Wilson et al. (1972). In the Loebel-Nelken-Tsodyks model, multiple such columns are connected according to the tonotopic or-

$$\tau_E \frac{d E_i^Q}{dt} = -E_i^Q + (1 - \tau_{\text{ref}}^E E_i^Q) \left[\sum_{R=-2}^2 \frac{J_{EE}^{|R|}}{N_E} \sum_{j=1}^{N_E} U x_j^{Q+R} E_j^{Q+R} + \frac{J_{EI}}{N_I} \sum_{j=1}^{N_I} U y_j^Q I_j^Q + e_i^{E,Q} + \sum_{M=1}^P s_i^{Q,M} \right]^+ \quad (1)$$

$$\tau_I \frac{d I_l^Q}{dt} = -I_l^Q + (1 - \tau_{\text{ref}}^I I_l^Q) \left[\sum_{R=-2}^2 \frac{J_{IE}^{|R|}}{N_E} \sum_{j=1}^{N_E} E_j^{Q+R} + \frac{J_{II}}{N_I} \sum_{j=1}^{N_I} I_j + e_l^{I,Q} \right]^+ \quad (2)$$

ganisation of A1 (Figure 1B). This arrangement gives rise to characteristic A1 activity where PS events initiate in a specific column (reflecting the frequency of the input), weaken in time as STD accumulates, and spread laterally across neighbouring columns through excitatory connections.

By emphasising intracortical interactions, rather than focusing solely on thalamocortical inputs, this approach offers a mechanism for producing a range of non-linear response features. In particular, recurrent excitation combined with synaptic depression explains how robust onset responses or phasic activity can self-terminate.

2. Methods

We modelled each iso-frequency column of A1, following Loebel et al. (2007) as a fully connected recurrent network with N_E excitatory and N_I inhibitory neurones. Neuronal activity was represented by excitatory, E_i , and inhibitory, I_l , firing rate variables, whose dynamics follow the Wilson et al. (1972) model. Synaptic inputs comprise (i) local recurrent excitation and inhibition with baseline efficacies $J_{EE}^0, J_{EI}^0, J_{IE}^0, J_{II}^0$; (ii) background inputs e_i^E, e_l^I , drawn from a random uniform distribution; and (iii) sensory inputs s_i that affect excitatory neurons only (Figure 1A). We used a threshold-linear function, $[z]^+ \equiv \max(z, 0)$, to convert net input drive into firing rates.

To then capture the tonotopic organisation of A1, we followed Loebel et al. (2007) to arrange P such columns ($Q = 1, \dots, P$) along an increasing frequency axis. Each excitatory neurone in column Q projected to excitatory, $J_{EE}^{1,2}$, and inhibitory, $J_{IE}^{1,2}$, populations in columns $Q \pm 1$ and $Q \pm 2$ (Figure 1B).

With intercolumnar coupling, the excitatory and inhibitory dynamics at column Q , can be found in Equations 1 and 2.

Recurrent and intercolumnar connections onto excitatory neurons exhibit short-term depression governed by variables x_i (for E to I connections) and y_l (for I to E connections). These obey:

$$\frac{d x_i^Q}{dt} = \frac{1 - x_i^Q}{\tau_{\text{rec}}} - U x_i^Q E_i^Q, \quad (3)$$

$$\frac{d y_l^Q}{dt} = \frac{1 - y_l^Q}{\tau_{\text{rec}}} - U y_l^Q I_l^Q, \quad (4)$$

where τ_{rec} is the recovery time constant, and U is the fraction of available synaptic resources used per presynaptic spike. Intuitively, each excitatory synapse has a finite pool of resources (x_i, y_l); strong activity depletes the resource, and recovery occurs over τ_{rec} .

Following Loebel et al. (2007), we modelled sensory inputs as frequency-specific drive to excitatory neurons. A pure tone, at frequency matching column M , provides an input:

$$s_i^{Q,M}(t) = \zeta^M(t) h_i^{Q,M}, \quad (5)$$

where $\zeta^M(t)$ represents the temporal envelope of the stimulus, while $h_i^{Q,M}$ is a spatial factor controlling amplitude across the tonotopic axis:

$$h_i^{Q,M} = A e^{-\frac{|Q-M|}{\lambda_s(A)}}, \quad \lambda_s(A) = \begin{cases} \lambda_C & A \leq \alpha \\ \lambda_C + \frac{(A-\alpha)}{\delta} & A > \alpha \end{cases}, \quad (6)$$

where A is the maximum magnitude of s_i at column M and $\lambda_s(A)$ determines the localisation of an input's effect which, as a function of A , allows increasing amplitudes to broaden the cortical extent of activation, reflecting known increases in tuning width with stimulus intensity (Hudspeth 2000). λ_C is a baseline parameter that dictates cortical extent and is modified, at amplitudes greater than a threshold α , by an auxiliary parameter δ . This allows the introduction of asymmetrical input into the model by choosing different values of δ when $Q < M$ (δ_{left}) and when $Q > M$ (δ_{right}).

Notably, s_i was only applied to spontaneously active neurones following a period of equilibration. This number was highly dependent on the variability in $e^{E,I}$ as a function of the random seed, therefore a seed was chosen that matched well with the findings of Loebel et al. (2002) of 59 spontaneously active neurones.

Unless stated otherwise, the parameter choices used throughout this report are identical to those used in Loebel et al. (2007) with the addition of setting the random seed to 47.

3. Results

3.1. Basic Properties

In simulating the tonotopic map of A1 (Zhang et al. 2003), we followed the work of Loebel et al. (2007) to simulate interconnected iso-frequency columns. Each neurone was provided background input that was randomly distributed and normalised (between $e_1^{E,I}$ and $e_{N,E,I}^{E,I}$). This random distribution was obtained by selecting a random seed that matched the number of spontaneously active neurones obtained by Loebel et al. (2007). The variability of model responses to changing this seed, along with a stereotypical model response that aligns with Loebel et al. (2007) is shown in Figure 2. This variability will be further discussed in Section 4. Excitatory neurones, within each column, also received sensory input, representing auditory stimulus-induced signals from the thalamus.

Following a period of equilibration, the average activity of neurones is asynchronous and less than 5 Hz. However, when a stimulus, above a certain threshold, is introduced the firing rate increases significantly and synchronously, producing a population spike (PS). This PS is single-onset and does not repeat with sustained input but may spread to neighbouring columns if the stimulus is strong enough. This propagation is a prod-

uct of inter-column synaptic connections (e.g. J_{EE}^1) as well as the spatial extent of the sensory stimulus (see methods). These parameters were fixed to allow gradual propagation of PS across the cortical sheet (Figure 3). With a sustained input, both the level of asynchronous activity and the magnitude of the population spike increased as a function of stimulus amplitude.

Loebel et al. (2007) note that excitatory and inhibitory activity within A1 is co-tuned, with the onset of inhibitory activity fixed to excitatory activity onset with a small delay (Wehr et al. 2003). While the latency to onset of excitatory activity is modulated by the frequency and amplitude of the sensory stimulus, the delay to inhibitory activity onset remains the same (Wehr et al. 2003). Loebel et al. (2007) replicate these results in their model with excitatory populations being first recruited by the stimulus before triggering increased activity in connected inhibitory neurones of the same column. Therefore, the recruitment of inhibitory neurones in the model depends only on the strength of connections into them and not on input characteristics.

3.2. Frequency Tuning Curves

Following the methods of Loebel et al. (2007), we characterised frequency tuning curves (FTCs) in our model by determining the minimum amplitude required for a given frequency to elicit a population spike (PS) within a cortical column. Consistent with experimental find-

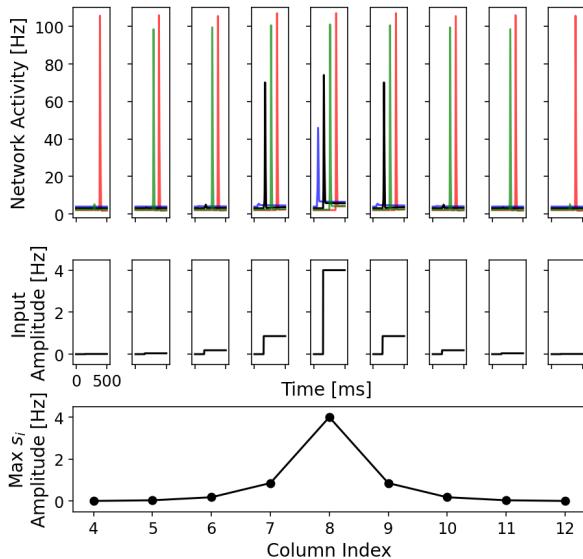


Figure 2. Model Responses to a pure-tone input. A sensory input is presented at 100 ms, for 400 ms with a best frequency of column 8 (black trace). Spatial spread of this input is shown in the bottom panel, with input amplitudes across the cortical sheet presented above. Mean activity over all neurones in a particular column is presented in the top panel. Black traces indicate simulations using a random seed of 47, whereas coloured traces (red, seed 2; blue, seed 28; green, seed 64) illustrate how changes in the random background input distribution (e^E) lead to varying population spike (PS) generation, with anywhere from 0 to 15 columns recruited. Note that latency to PS onset is unaffected, altered stimulus start times were chosen to separate the curves for clarity.

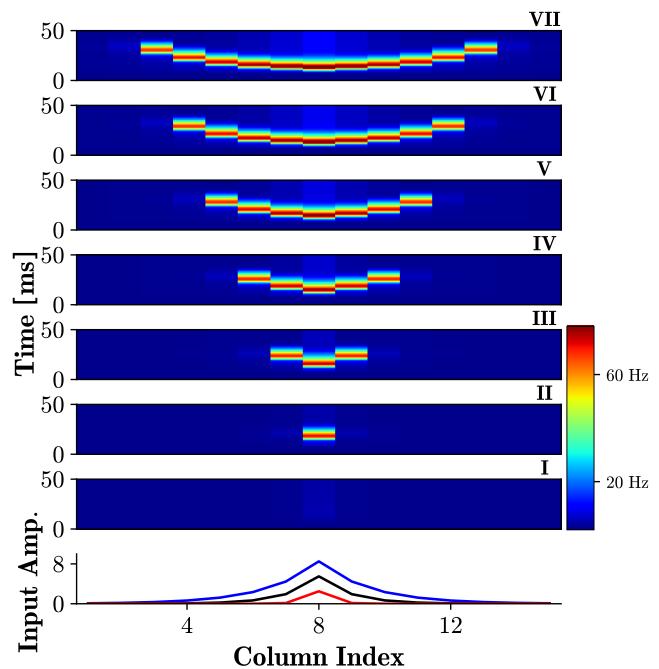


Figure 3. PS propagation through time. Bottom panel illustrates spatial extent of stimuli with three different peak amplitudes (1, 4.5, 8.5 Hz). Top panels show PS spread through time with sensory inputs of increasing amplitudes (1.0, 2.5, 4.0, 5.5, 7.0, 8.5, 10 Hz; Panels I-VII, respectively).

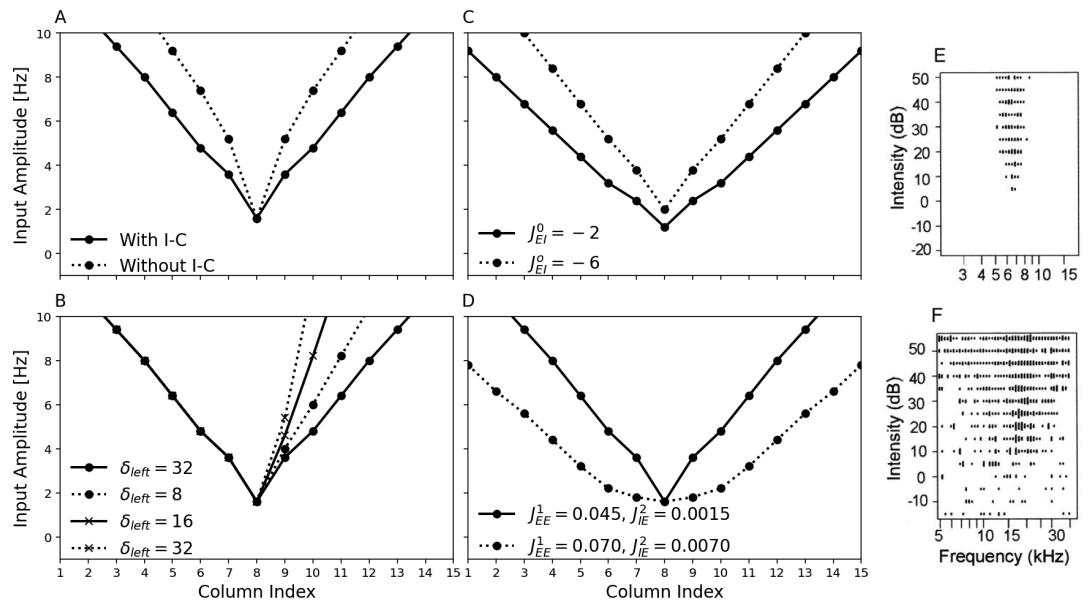


Figure 4. Frequency Tuning Curves (FTC). FTC were determined by applying a sensory input, with varying frequency, and measuring minimal amplitude required to evoke a PS at column 8. (A) FTC in networks with and without inter-columnar (I-C) connections. (B) Asymmetric FTC induced by asymmetry in the spatial spread of sensory input (altered δ_{left}). (C) FTC with increased and decreased recurrent inhibition. (D) FTC with increased primary neighbour excitatory connections and secondary neighbour inhibitory connections. These parameter modulations achieve a 'U'-shaped FTC observed in experimental data in A1 (Sutter 2000). (E, F) 'V' and 'U'-shaped FTC recorded from cat A1 (Sutter 2000).

ings (Schreiner et al. 2000; Ulanovsky et al. 2003), FTCs exhibit a characteristic 'V'-shaped profile, with a distinct best frequency (BF).

Experimental work suggests that thalamo-cortical and intracortical pathways differentially contribute to frequency selectivity (Kaur et al. 2004). In line with Loebel et al. (2007), we found that FTC bandwidth in our model is influenced not only by the spatial spread of sensory input but also by the ability of PSs to propagate between columns. By adjusting intercolumnar connection strengths ($J_{EE}^1, J_{EE}^2, J_{IE}^1, J_{IE}^2$), we observed that reducing lateral connectivity, or removing it altogether, restricts PS spread, narrowing the FTC (Figure 4A). However, PS threshold and onset latency at BF remained unchanged, as responses at this frequency were mediated directly by sensory input.

We also examined the impact of asymmetric sensory input ($\delta_{left}, \delta_{right}$) on FTC shape, replicating the asymmetry observed in experimental (Moshitch et al. 2006) and computational (Loebel et al. 2007) studies. Indeed, asymmetric sensory input resulted in a narrowing of one edge of the FTC (Figure 4B). Furthermore, following Loebel et al. (2007), we tested the role of recurrent inhibition (J_{EI}^0, J_{II}^0) and found that reducing inhibitory strength resulted in broader FTCs and lower response thresholds, with an opposite effect for increases inhibitory connection strengths (Figure 4C). Furthermore, Sutter (2000) highlight significant variability in the shape of the FTCs in A1 neurones. Figure 4E-F highlight their results, showing a stereotyp-

ical 'V'-shaped and 'U'-shaped FTC, respectively. To explore how our model might replicate these findings, we increased primary-neighbour excitatory connections (J_{EE}^1) and secondary-neighbour inhibitory connections (J_{IE}^2) which results in a 'U'-shape FTC (Figure 4D).

In contrast to the findings of Loebel et al. (2007), the impact of adjusting background input (e^E, e^I) was not immediately explainable. Changes brought about by altered background input were propagated through the model both in the number of spontaneously active neurones present and in the equilibrated firing rate of these populations. The impacts of these changes will be discussed further in Section 4.

3.3. Forward Masking

3.3.1. Dynamics of Recovery

Forward masking is a phenomenon where an initial stimulus (the "masker") suppresses neuronal responses to a following stimulus for hundreds of milliseconds. This suppression has been extensively studied (Brosch et al. 1997; Calford et al. 1995; Wehr et al. 2005) and was widely thought to be driven by postsynaptic GABAergic inhibition.

A subsequent study by Wehr et al. (2005), which involved *in vivo* cell recordings measuring synaptic conductances following the presentation of an identical pair of brief sounds, suggests that this inhibition plays a part in the only the first 50-100 ms of the hundreds of milliseconds long suppression. They suggest that

synaptic depression, however, may play a larger role in suppression and could be the more likely candidate following observations of similar dynamics in other parts

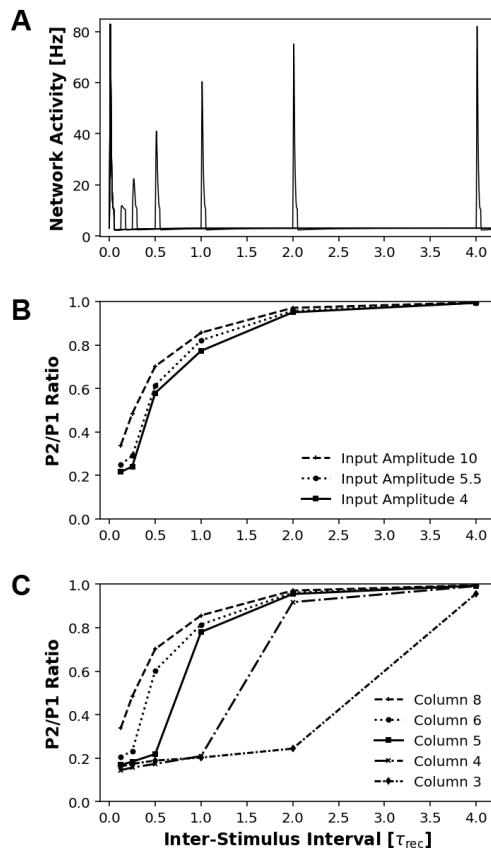


Figure 5. Recovery Dynamics of Forward Masking. (A) Network responses to pairs of identical pure tones at the best frequency (BF) of column 8, with varying inter stimulus intervals.(B) Recovery dynamics for different amplitudes, showing the ratio of responses relating the second tone to the first.(C)Recovery dynamics at column 8 for the pair of tones with frequencies corresponding to different columns, explaining frequency-dependent suppression.

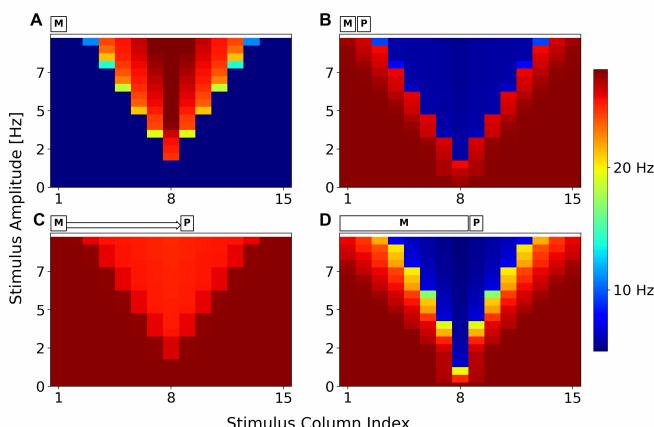


Figure 6. Temporal Dynamics of Forward Masking.(A) Network responses to brief maskers with varying Amplitudes and frequencies (B) Responses of probe at the best frequency (BF) of column 8 following a masker.(C)Responses of probe after a prolonged silent interval.(D)Responses of probe after extended durations of masker, showing partial suppression.

of the brain (Chung et al. 2002; Freeman et al. 2002).

In their model, Loebel et al. (2007) demonstrated that it exhibits forward masking with features similar to those observed in the study by Wehr et al. (2005). Two identical pure tones, with big enough amplitude to recruit a PS in the corresponding columns, were presented with varying inter-stimulus intervals (ISIs). For sufficiently short ISIs, the PS response to the masker tone effectively suppressed the probe tone, as expected. In the model, this suppression appears to result from a gradual increase in short-term synaptic depression of the connections between excitatory neurons in the column, following the PS response to the masker tone. This depression makes it more difficult for a second PS to be recruited by the probe tone shortly after. Further evidence for this came from the finding that masking follows an exponential recovery process which is governed by the time constant of synaptic depression recovery shown in Figure 5B.

The recovery rate seems largely unaffected by the amplitude of the stimulus that triggered the PS (Figure 5B). However, the same does not seem to hold true for the frequency of the tone stimuli. While recording the activity from a selected column, Loebel et al. (2007) found that the further the frequency of the stimuli was from the BF of the column, the longer the delay for the recovery from suppression was. Since the recorded column depends on signal propagation from the BF column through intracortical connections, the depression caused by the PS in response to the masker tone leads to a weaker signal propagating after the probe tone as well as smaller amplitude PS in intermediate columns.

In an effort to replicate these results, we presented the model with a pair of identical pure tones, each 50ms in duration and with an amplitude of 10 Hz (matching the conditions of the original study). We were able to reproduce nearly identical results, confirming the influence of both amplitude and frequency of the stimuli on the recovery duration. However, there was one discrepancy: in the paper, when column 8 was presented with the BF of column 3, no recovery occurred by ISI $\tau_{rec} = 4$, whereas our results showed full recovery by that point. This difference may be attributed to the stochastic elements of the model, which will be discussed further in Section 4 along with reproducibility.

3.3.2. Dependency on Masker Duration

Some biophysical studies suggest that in A1, the duration of the masker significantly influences the extent of forward masking (Brosch et al. 1997; Alves-Pinto et al. 2010). Brosch et al. (1997) examined neural responses to a probe following a masker with varying frequencies and intensities. Loebel et al. (2007) sought to replicate these findings using their computational model (Figure

9 in (Loebel et al. 2007)).

Their results demonstrated that when a masker was followed by a probe with a short offset—insufficient for synaptic depression recovery—the probe’s response was completely suppressed (Figure 6B). When the same masker was followed by a longer offset, a clear probe response was observed (Figure 6C). However, when a prolonged masker with the duration of the offset of Figure 6C was followed by a probe after an offset identical to that in Figure 6B, the resulting behaviour differed from both previous conditions, indicating that masker duration plays a crucial role (Figure 6D). We were able to reproduce these results accurately, this is shown in Figure 6.

The study suggests that prolonged maskers help sustain the depressed state, and that the further the masker’s frequency deviates from the BF of the measured column, the weaker this sustained suppression becomes, allowing for partial population spike generation in response to the probe.

3.4. Hypersensitive Locking Suppression

Hypersensitive locking suppression refers to the suppression of auditory neuronal locking to the amplitude envelope of slowly fluctuating noise when a low-level tonal input is introduced. This phenomenon, originally described in experimental studies (Las et al. 2005; Nelken et al. 1999), occurs even when the tone is presented at sub-threshold levels that do not evoke an onset spiking response in isolation. Loebel et al. (2007) demonstrated this effect following a setup similar to that of Las et al. (2005) (see Figure 7a).

To simulate the modulated noise, a train of excitatory clicks ($0.01 \tau_{rec}$ click duration) was applied simultaneously across all columns. Since Loebel et al. (2007) did not explicitly describe their simulation parameters, we based our implementation on the approach of Loebel et al. (2002), which initially suggested synaptic depression as a primary mechanism underlying locking suppression. Specifically, each column received excitatory clicks of equal amplitude $\Delta E = a\Delta E_{min}$, where ΔE_{min} represents the minimal stimulus amplitude required to elicit a PS response. This threshold was determined by presenting stimuli across a range of amplitude values and defining ΔE_{min} as the lowest amplitude at which network activity exceeded 10 Hz. In line with Loebel et al. (2002), we set $a = 1.5$.

A similar method was used to identify the maximum tone amplitude $\Delta E_{threshold}$, delivered at the BF of column 8, which did not evoke a PS response in isolation. This threshold value was used as a reference to assess whether suppression occurred at sub-threshold tone amplitudes, as observed experimentally (Las et al. 2005) and replicated in Loebel et al. (2007).

Loebel et al. (2007) noted that the tone amplitude required for suppression depends on the connectivity pattern of thalamic inputs within the network. When the tone was applied uniformly across all neurones, suppression occurred only at amplitudes significantly higher than $\Delta E_{threshold}$ (Figure 7b). This suggests an importance in network connectivity patterns: by focusing sensory input on spontaneously active neurones, the model facilitated an earlier or stronger engagement

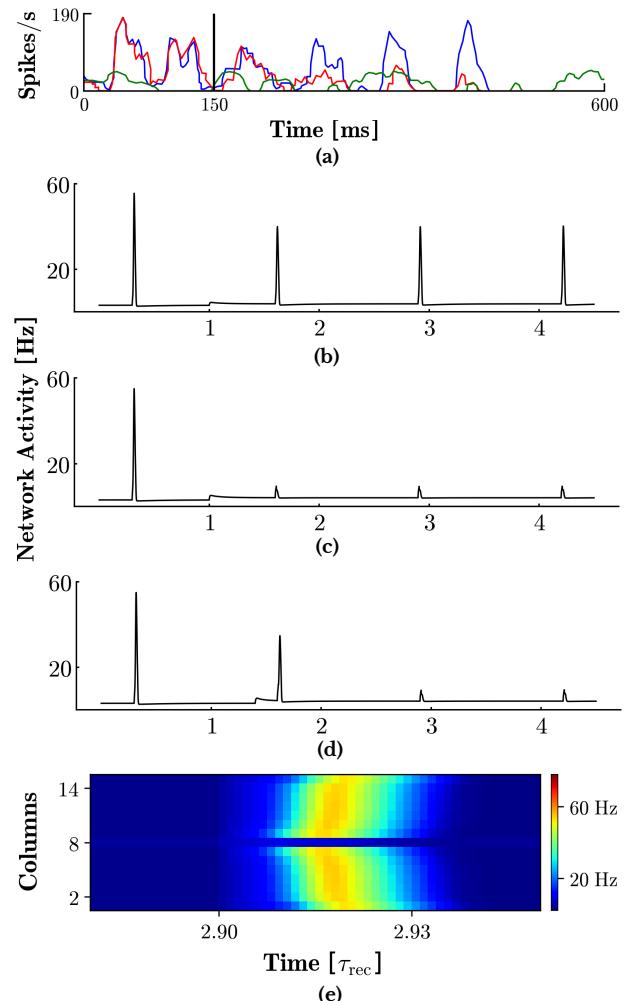


Figure 7. Hypersensitive Locking Suppression. (a) Experimental results from Las et al. (2005) showing the spike responses of a single neurone in A1 to tones and modulated noise. The plot presents average responses to a noise-plus-tone stimulus at the minimal tested tone level (red), the tone alone at the same level (green), and the noise alone (blue). The vertical bar at 150ms indicates tone onset. (b) Response of column 8 to a sustained pure tone at its BF with a sub-PS threshold amplitude, presented alongside the modulated noise to all neurones. (c) When both the modulated noise and the sustained pure tone were delivered only to spontaneously active neurones, complete suppression of PS responses to clicks following tone onset was observed. (d) Illustration of second NC suppression effects attained using the Loebel et al. (2007) model. When the tone onset is shifted closer to the first NC, full suppression emerges only by the second NC. (e) Average network activity per column during a suppressed spike in the simulation presented in (b) - suppression was localised to column 8.

of synaptic depression, therefore producing hypersensitive locking suppression at sub-threshold amplitudes (Figure 7c).

Consistent with experimental findings (Rupp et al. 2007), the suppressive effects of the tone were localised to the BF column, with minimal or no effect on the responses of other columns to the modulated noise (Figure 7e).

As noted by Loebel et al. (2007), the simulated results do not fully capture all aspects of the *in vivo* phenomenon. In the model, suppression emerges from the first noise cycle (NC) following tone onset, whereas *in vivo* data suggest a delay until the second NC (Las et al. 2005). However, it remains unclear whether this delayed suppression is time-dependent or tied to the NC structure. A similar delay can be reproduced in the model by adjusting the tone onset relative to the first NC, as shown in Figure 4d. In this case, the network does not have sufficient time to transition into a more depressed state before the first NC arrives, and thus full suppression is only observed by the second NC.

Furthermore, suppression in the model is highly sensitive to the duration of the noise signal. If the click duration is slightly prolonged (eg. $0.012 \tau_{rec}$ rather than $0.01 \tau_{rec}$), then hypersensitive locking suppression can no longer be achieved for sub-threshold tonal inputs, regardless of the connectivity pattern of thalamic inputs.

3.5. Response to Sweeps

A key aspect of this model is the ability to simulate responses to frequency modulated (FM) sweeps. PS mimic the neural responses to these sweeps by propagating along the tonotopic map. These neural activity patterns and the propagating nature of PS are influenced by the direction of the sweep and their velocity.

(Figure 4b shows the linear relationship between the sweep duration and the onset latency of PS responses in column 8 for downward FM sweeps. Figure 8 compares the effective tuning curve for FM sweeps with the pure tones. A1 neurones are sensitive to sweep direction and sweep velocity (Nelken et al. 2000; Zhang et al. 2003). By capturing the features, the model suggests the interaction between the dynamics of the stimulus and the intrinsic properties of the network. This model contrasts with purely feedforward models, such as Fishbach et al. (2003), which still accounts for the direction and velocity by fail to fully capture the role of intracortical circuitry. Heil et al. (1992) and Nelken et al. (2000) found this linear relationship indicating the response occurs at a specific “triggering frequency”, which is independent of the sweep velocity. This triggering frequency is above the best frequency (BF) for downward sweeps and below BF for upward sweeps.

The Loebel et al. (2007) model replicates this observation, suggesting that the front of activity propagation is synchronized with the FM sweep, regardless of its velocity. When the intercolumnar connections are removed, the tuning curve for FM sweep is similar to that of pure tones. This result arises due to inputs from the neighbouring columns are desynchronized by the FM sweeps. Isolated tones with frequencies far from BF needs synchronization, FM sweeps disrupt this by effectively disconnecting the intercolumnar connections. However, a PS in the target column occurs only when the feedforward input is strong enough to induce PS on its own, this leads to a triggering frequency independent of sweep duration.

The model also suggests that the triggering frequency of FM sweeps is primarily calculated by feedforward inputs from the auditory thalamus. This leads to the understanding that comparing tuning curves for FM sweeps and pure tones can outline the extent of thalamic inputs into A1 at a given cortical site. FM sweeps dissociate the contribution of thalamic inputs from lateral intracortical inputs, providing a method to probe the relative influence of these two input streams.

3.6. PS-Based Coding of Complex Sounds

The propagation of PSs along the model’s tonotopic map provides a structured encoding mechanism for complex sounds, where distinct patterns of PS activity emerge in response to different auditory inputs. Loebel et al. (2007) demonstrated that the occurrence of a PS response depends on the temporal context of preceding

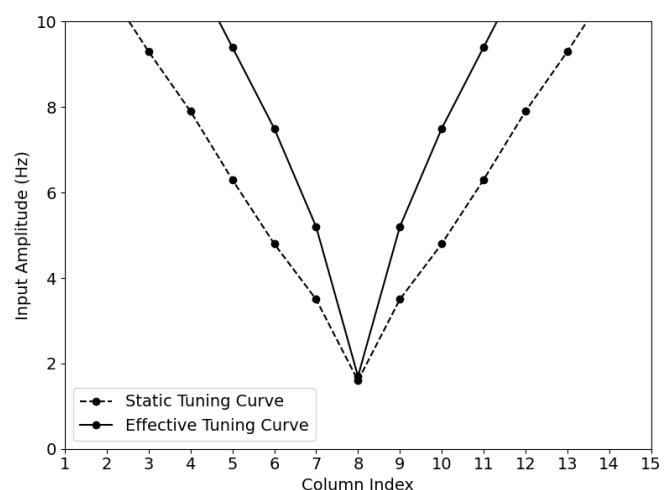


Figure 8. Comparison of Effective and Static Tuning Curves. The effective tuning curve (solid line) is derived from responses to frequency-modulated (FM) sweeps. They closely matches the static tuning curve (dashed line) obtained from pure tones without inter-columnar connections. This alignment demonstrates that the triggering frequency for population spikes (PS) during FM sweeps is determined by feed-forward thalamic inputs, which are independent of the sweep velocity.

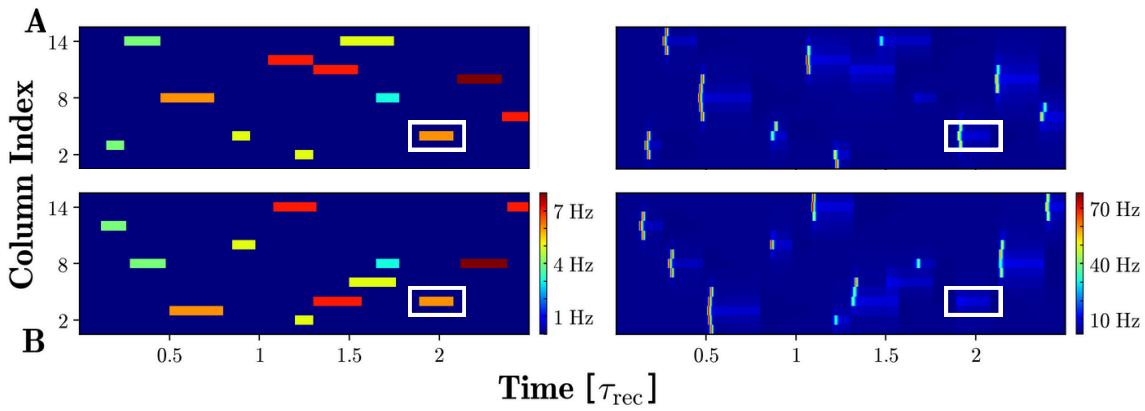


Figure 9. PS-based encoding of complex sounds. The left panels represent the spectrograms of two different complex sounds (the equivalent of **B** and **C** from Figure 13 in Loebel et al. (2007)). The right panels illustrate the corresponding network activity. The white squares highlight a particular tone pip that was identical in both conditions, but evoked a PS response only in **A**. This suggests that the response selectivity is influenced by changes in the spectral composition of the stimuli.

inputs, meaning that the representation of a complex sound cannot be directly inferred from the responses to its individual spectral components (Figure 9). This nonlinearity suggests a PS-based encoding scheme in A1, where synchronous, temporally precise PS events across neural populations transform complex input patterns into distinct neural representations.

More recent research has reinforced the role of PS-based encoding in auditory perception. Studies show that temporally precise population activity in A1 is critical for encoding fast sound modulations, with PS timing outperforming rate-based codes in representing amplitude-modulated stimuli (Downer et al. 2021; Ince et al. 2013). Similarly, perceptual discrimination of complex acoustic sequences, such as detecting subtle timing changes in natural sounds, is best predicted by spike-timing-based neural codes rather than firing rate alone (Walker et al. 2008). The sensitivity of PS responses to stimulus history also aligns with models of stimulus-specific adaptation (SSA) which we briefly explore in 6, where recurrent cortical circuits amplify responses to deviant sounds through population bursts (Yarden et al. 2017). This suggests that PS activity may contribute to novelty detection and the encoding of auditory expectations, potentially linking PS-based coding to predictive coding frameworks. Therefore, these studies identify PS activity as a key mechanism in the cortical representation of complex sounds, offering a temporally precise, context-dependent coding strategy that extends beyond traditional rate-based models.

4. Model Reproducibility

4.1. Impact of Random Seed on Model Response

As shown previously in Section 3.1, variations in the random seed that sets e^E and e^I distributions led to wide fluctuations in the number of columns recruited

to generate a PS in response to a sustained pure tone (amplitude 4 Hz). Some seeds produced little to no PS activity, while others yielded as many as 15 spiking columns (Figure 2). Figure 10 illustrates the distinct distributions of e^E in relation to a linear spacing from e_1^E to $e_{N_E}^E$ as a function of varying seed.

We noted that each random realisation of e^E and e^I influenced three key factors: (i) the number of spontaneously active neurones, (ii) the equilibrated activity of excitatory and inhibitory populations in each column, and (iii) the dynamics of the synaptic depression factors x and y . However, these factors proved to be insufficient to explain the changes in column recruitment. For example, contrary to expectation, even identical numbers of spontaneously active neurones yielded a diverse range of column recruitment (Figure 11A-C).

Because we found no correlation between changes

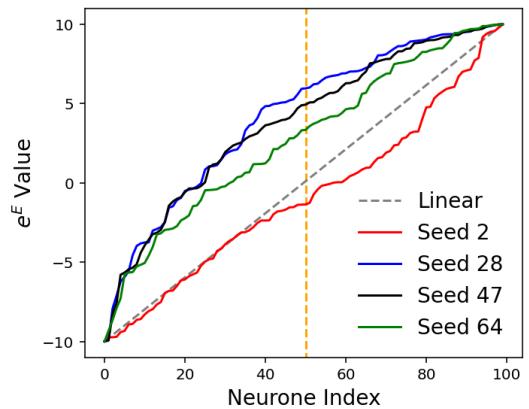


Figure 10. Distribution of e^E away from the linear spacing
Distributions of e^E vary significantly as a function of random seed. Grey, dashed line indicated the linear spacing from e_1^E to $e_{N_E}^E$. Solid, coloured lines illustrate distributions of e^E using different seeds. Orange, dashed line indicates the threshold between the negative and positive domains of the linear e^E distribution.

in any key variable in the model and the number of columns recruited, we next examined the intrinsic distributions of the background input. In examining deviations from linearly spaced e^E and e^I distributions, we observed that shifts towards more positive or negative values in distinct domains of e^E and e^I could alter the degree of column recruitment. Here, we take domain to mean the positive or negative portion of a linear $e^{E,I}$ distribution, as illustrated in Figure 10. We quantified these deviations by measuring the net shifts, in each domain, away from the linear distribution (Figure 11D-E). Correlating these shifts with column recruitment across many seeds did not show significant causation from e^I (Figure 11F-G). This was presumably because e^I exerts only an indirect effect on excitatory neurones through recurrent connections from inhibitory neurones. In contrast, e^E directly influences excitatory populations and therefore exerts a stronger effect on spontaneous neuronal activity and thus the extent of sensory input. This was supported by a strong correlation between e^E distribution and column recruitment (Figure 11D-E). Consequently, further analysis will focus on e^E , while e^I remains randomly distributed.

4.2. Quantifying Random Effects

Closer inspection of domain-specific shifts revealed that a positive sine shift (i.e. a positive shift in the negative domain and a negative shift in the positive domain) elicited more robust column recruitment and thus broadened the FTC. Conversely, a negative sine shift narrowed the FTC by limiting column recruit-

ment.

Taking this idea that sine shifts may be sufficient to explain the correlation between changes in Ee^E distribution and column recruitment, we applied a Fast Fourier Transform (FFT) to this random distribution. This allowed visualisation of the most significant components to the random distribution in terms of sine and cosine waves. Components with spatial frequencies above 30% of the highest spatial frequency (with units: cycles per neurone index) were preserved, while those with amplitudes below 7% of the maximum were removed. This approach ensured that the random distribution retained only its slower and most significant spatial fluctuations, making the connection between random variability and population response more transparent. In total, we retained 6 components with frequencies 0.01, 0.02, 0.03, 0.05, 0.09 & 0.14.

We found that this restricted distribution produced an identical FTC to the original random distribution. As such, specific positive and negative shifts in distinct regions of the random distribution of $e^{E,I}$ are responsible for the profound sensitivity of the model to random seed choice. Therefore, specifying the seed used, and thus background input distribution, is paramount to achieving model reproducibility. Further, this highlights the significant effect that the background input has on PS generation and propagation and so, alterations can bring about rapid modulations of network responses. Loebel et al. (2007) highlight this by shifting the entire distribution towards more negative values, however, we have shown that even slight, restricted

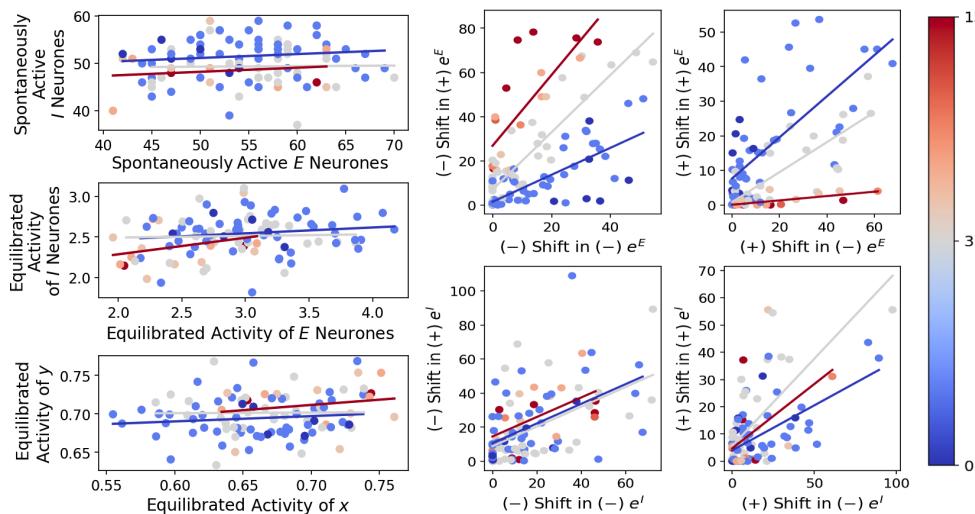


Figure 11. Variability in model response as a function of random seed choice All panels use random seed from 0 - 100 to assess correlation between column recruitment and changes in model variables (A - C) or parameters (D - G). Colour bars indicate number of columns recruited to generate a PS. Regression lines calculated from realisations with <3 columns recruited (blue), 3 columns recruited (grey), and >3 columns recruited (red). (A) Number of spontaneously active excitatory and inhibitory neurones, no correlation. (B) Equilibrated firing rate variables, no correlation. (C) Equilibrate synaptic depression factors, no correlation. (D) Negative shifts in positive and negative domains of e^E , strong correlation. (E) Positive shifts in positive and negative domains of e^E , strong correlation. (F) Negative shifts in positive and negative domains of e^I , no correlation. (G) Positive shifts in positive and negative domains of e^I , no correlation.

changes in values away from a linear spacing can result in profoundly different network responses.

5. Effects of Signal-to-Noise Ratio

The auditory system can process sounds over a very wide dynamic range, spanning many orders of magnitude in intensity. However, background noise effectively shrinks this usable range, reducing the ability to detect and classify sounds (Pollack et al. 1958). As the noise level increases relative to a given signal, the signal-to-noise ratio (SNR) drops, leading to progressively worse performance in speech and sound detection (French et al. 1947; Dubno et al. 2005). Notably, even when the SNR is held constant, making both the signal and noise louder can still degrade perception—speech at a high overall level in noise is less intelligible than speech at the same SNR but a lower level (Liang et al. 2014). This suggests that the absolute sound level, in conjunction with SNR, influences auditory processing outcomes, especially at extremes of

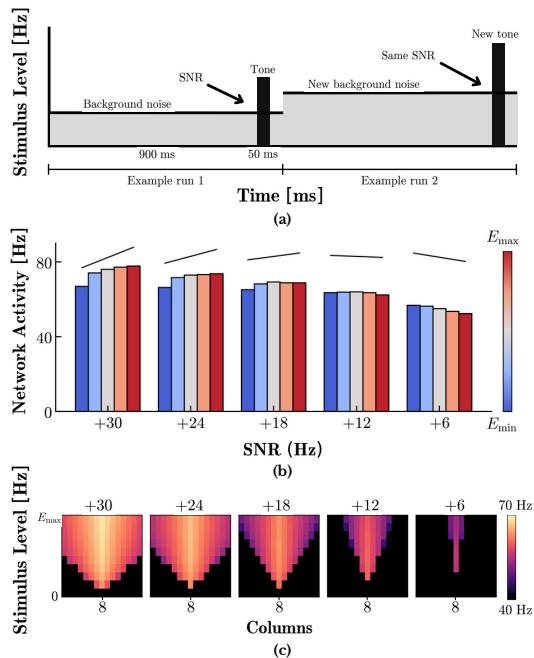


Figure 12. Effects of Signal-to-Noise Ratio on Network Activity and Frequency Response. (a) Schematic of the simulation design for tone-in-noise presentation, modelled after the experimental paradigm of Teschner et al. (2016). The amplitude of the background noise (grey) and the tone (black) is adjusted to maintain a fixed signal-to-noise ratio (SNR), with both components varying between trials. Each run lasted 1s. (b) Network activity across different SNR conditions, measured over a range of stimulus levels. The black lines above the bars indicate the average slope of activity change across increasing stimulus levels. Consistent with Teschner et al. (2016), higher SNRs exhibit a positive slope, whereas lower SNRs show a negative slope. (c) Proxy-FRAs obtained for different SNR conditions. At higher SNRs, the proxy-FRA bandwidth increases, in agreement with Teschner et al. (2016). However, in contrast to their results, the proxy-FRA bandwidth also tends to increase at lower SNRs.

intensity. One way to capture how neurons respond to varying frequency and intensity combinations under these different conditions is by measuring frequency response areas (FRAs), which track changes in neuronal firing rates across a range of frequencies and intensities to reveal which combinations most effectively drive the cell. Recently, Teschner et al. (2016) found that for high SNRs, a joint increase in the tone and noise level led to increased firing rates and expanded frequency response area bandwidths. However, at low SNRs, a joint increase led to reduced firing rates and narrowed FRA bandwidths.

To explore the Loebel et al. (2007) model responses under various SNR conditions, we followed a similar experimental paradigm to Teschner et al. (2016) (Figure 12a). Unlike many prior studies that varied either the tone level or the noise level while holding the other fixed (Shetake et al. 2011; Schneider et al. 2013), Teschner et al. (2016) maintained a constant SNR by jointly varying both signal and noise levels. Specifically for our simulations, a pure tone was applied at the BF of column 8 to serve as the foreground signal. This ensured that column 8 was the most strongly driven by the tone. To simulate white noise, a continuous uniform input was delivered at equal amplitude to all columns in the network. The amplitudes of the tone and noise inputs were calibrated such that the tone reliably evoked PSs in the BF column, while the noise amplitude remained below the PS threshold. As such, each SNR was tested starting from a slightly different minimum stimulus level E_{min} – due to higher thresholds at lower SNRs – up to the same maximum stimulus level E_{max} . We defined SNR on a linear scale as the power ratio of signal to noise

$$SNR = \frac{A_{signal}^2}{A_{noise}^2}.$$

By keeping this ratio fixed for a given simulation run, any increase in the signal amplitude was accompanied by a proportional increase in noise amplitude. Significantly low SNR values ($SNR < 5$) were not practical to test since the range of possible stimulus levels was too restricted to provide meaningful results, similar to how Teschner et al. (2016) excluded SNRs < 3 dB due to distortions.

5.1. Network Activity for Different SNRs

The simulations reveal a clear dependence of network activity on SNR that mirrors experimental findings (Figure 12b). At a relatively low SNR, increasing the overall sound level reduced the total spiking activity in the network. In other words, making both the tone and noise louder led to a paradoxical decrease in evoked responses when the noise was intense. By contrast, at high SNR,

increasing sound level produced the expected increase in network activity – louder sound evoked more firing across the network. This crossover in behaviour is qualitatively consistent with the observations of Teschner et al. (2016).

5.2. Effect on the Frequency Response Area

In addition to the observed firing rate changes, we also examined how SNR influences the spread of activity across the iso-frequency columns, used here as a proxy for FRA bandwidth (proxy-FRA). Under low SNR conditions, raising the overall sound level caused the BF column's response to extend into adjacent columns, thereby broadening the proxy-FRA. This outcome contrasts with Teschner et al. (2016), who reported that increasing the combined level under low SNR conditions narrowed rather than expanded the FRA. However, at high SNRs, our model results aligned with Teschner et al. (2016)'s findings: higher stimulus levels recruited more of the tonotopic map, mirroring the normal FRA expansion typically observed in quiet.

6. Stimulus-specific adaptation

Stimulus-specific adaptation (SSA) is the reduction in neuronal activity in response to a frequently repeated stimulus (standard), while sensitivity to a rarer stimulus (deviant) is maintained. This phenomenon has been

consistently observed in the primary auditory cortex (A1) across multiple studies involving rats (Taaseh et al. 2011; Yaron et al. 2012), cats (Ulanovsky et al. 2003; Ulanovsky 2004), monkeys (Fishman et al. 2012), and even the human auditory cortex (Briley et al. 2013). Pioneering this research, Ulanovsky et al. (2003) demonstrated that when exposed to an Oddball sequence, individual neurons responded more strongly to the deviant tone than to the standard one. The Oddball sequence is an experimental paradigm in which a frequently occurring standard tone is occasionally replaced by a deviant tone with a low probability. SSA is often regarded as nearly synonymous with "deviance detection," a critical survival mechanism that enables the detection of sudden, important sounds amid constant background noise (Nelken 2014).

Despite extensive research and experimental data, the mechanisms underlying SSA remain unclear. Recent computational models (Taaseh et al. 2011) suggest that feed-forward synaptic depression of thalamocortical input to A1 may play a key role in SSA. However, using a modified version of Rupp et al. (2007)'s model, Yarden et al. (2017) argue that previous models based on thalamocortical input depression do not fully align with experimental findings. Building on this idea, we tried to explore SSA using our implementation of the model introduced by Loebel et al. (2007).

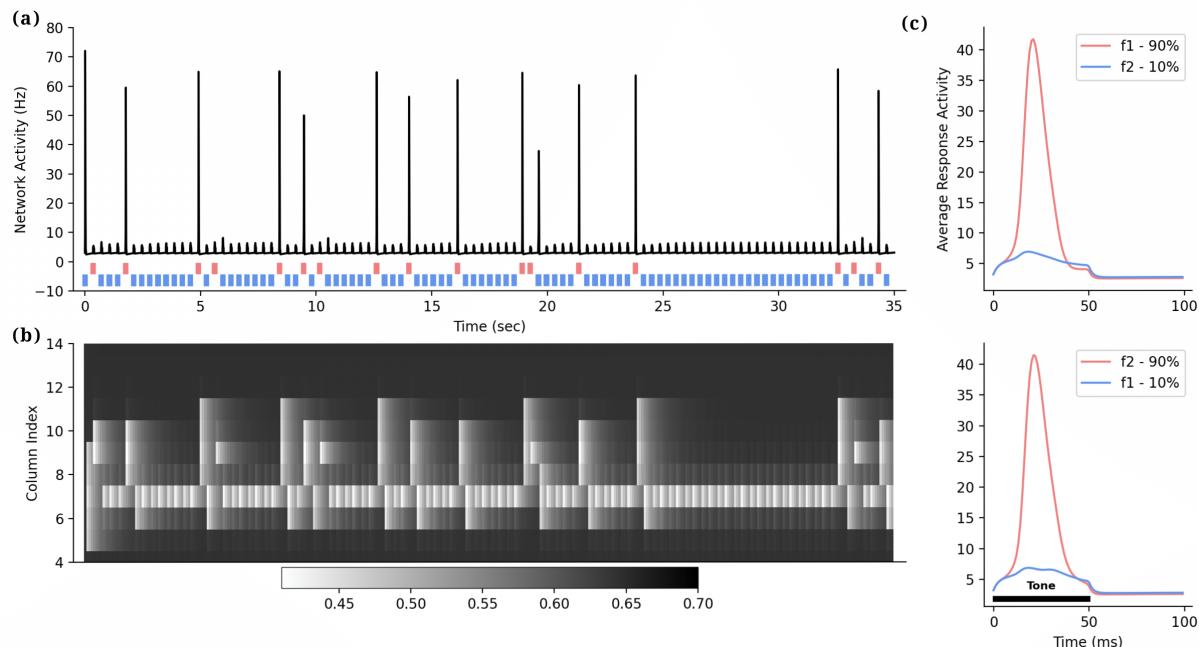


Figure 13. Stimulus-specific adaptation (SSA) dynamics in the model. (a) The mean network activity in column 9 "middle column" in response to the oddball sequence ("standard" = blue, "deviant" = red). f1 = BF of 8 and f2 = BF of 10. (b) The mean resources available, x , for each column along the time span of the oddball sequence in (a). (c) **Top:** The average of the responses for standard (blue) and deviant (red) the probabilities of f1 tones occurring are 90% and 10% for f2. **Bottom:** Same as top but f1: 10%, f: 90%. **In all simulations, input amplitude (A) = 5 and inter-stimulus interval (ISI) = 350 ms.**

6.1. Results

To investigate this, we presented the model with a regularly occurring "standard" pure tone at the best frequency (BF) of column 8, occasionally replaced by an identical "deviant" tone at the BF of column 10 over a 35-second period. Throughout this time, we measured the average network activity in the "middle column" (Figure 13A). Both standard and deviant tones elicited population spikes (PSs); however, deviant tones triggered them more frequently, as synaptic depression in the deviant column had more time to recover. PSs appeared to significantly deplete available resources, preventing subsequent PS generation (Figure 13B). Standard tones, on the other hand, produced only small spikes with each tone, gradually exhausting available resources. To confirm that these results were independent of tone frequency, experimental studies (Nelken 2014) measured average responses for both standard and deviant tones after swapping their frequencies. Following this approach, we generated similar plots and observed comparable results (Figure 13C).

7. Discussion

In this report, we have successfully replicated the key findings of Loebel et al. (2007): (i) reproducing model responses to sustained pure tone inputs, including gradual PS propagation through time; (ii) obtaining FTC matching Loebel's findings and extending their model to match FTC shapes found in cat A1; (iii) examining forward masking to highlight the effect of synaptic depression on subsequent stimulation, as well as highlighting the effect of masker duration on probe response; (iv) illustrating that lateral suppression is only achieved when the sensory input is supplied only to spontaneously active neurones; (v) detailing how PS-based encoding helps to explain how perception of complex stimuli arises.

Further, we examined the reproducibility of the model by probing the effect of random seed choice on model responses. We highlighted, using Fourier decomposition, the causal link between shifts in distinct domains of e^E and PS generation and propagation. Additionally, we examined how different SNRs influence network activity and proxy-FRAs in the model. Results show that increasing sound level increases firing rates at higher SNRs whilst decreasing activity at lower SNRs, consistent with prior findings (Teschner et al. 2016). However, at low SNRs, the model deviates from Teschner et al. (2016) by exhibiting a widening of the proxy-FRA instead of the reported narrowing. Finally, inspired by a newer version of the model (Yarden et al. 2017), we decided to explore SSAs and their dynamics in our implementation of the current model. We were

able to replicate experimental results and show that deviant tones that occur at a rate less than standard ones seemed to elicit most of the PS responses and showed how the synaptic resources available are depleted due to the tones and thus linking synaptic depression to the mechanisms behind SSAs and 'deviance detection'.

■ References

- Alves-Pinto, Ana et al. (2010). [\[link\]](#).
- Bar-Yosef, Omer, Yaron Rotman, and Israel Nelken (2002). [\[link\]](#).
- Bendor, Daniel and Xiaoqin Wang (2005). [\[link\]](#).
- Briley, Paul M. and Katrin Krumbholz (2013). [\[link\]](#).
- Brosch, Michael and Christoph E. Schreiner (1997). [\[link\]](#).
- Calford, Mike B and Malcolm N Semple (1995). [\[link\]](#).
- Chung, Sooyoung, Xiangrui Li, and Sacha B. Nelson (2002). [\[link\]](#).
- Downer, Joshua D. et al. (2021). [\[link\]](#).
- Dubno, Judy R., Amy R. Horwitz, and Jayne B. Ahlstrom (2005). [\[link\]](#).
- Elhilali, Mounya et al. (2004). [\[link\]](#).
- Fishbach, Alon, Yehezkel Yeshurun, and Israel Nelken (2003). [\[link\]](#).
- Fishman, Y. I. and M. Steinschneider (2012). [\[link\]](#).
- Freeman, Tobe C.B et al. (2002). [\[link\]](#).
- French, N. R. and J. C. Steinberg (1947). [\[link\]](#).
- Heil, Peter, R. Rajan, and Dexter R.F. Irvine (1992). [\[link\]](#).
- Hudspeth, AJ (2000). [\[link\]](#).
- Ince, Robin A. A., Stefano Panzeri, and Christoph Kayser (2013). [\[link\]](#).
- Kaur, Simranjit, Ronit Lazar, and Raju Metherate (2004). [\[link\]](#).
- Las, Liora, Edward A. Stern, and Israel Nelken (2005). [\[link\]](#).
- Liang, Feixue et al. (2014). [\[link\]](#).
- Loebel, Alexander and Michail Tsodyks (2002). [\[link\]](#).
- Loebel, Alexander et al. (2007). [\[link\]](#).
- Machens, Christian K., Michael S. Wehr, and Anthony M. Zador (2004). [\[link\]](#).
- Moshitch, Dina et al. (2006). [\[link\]](#).
- Nelken, Israel (2014). [\[link\]](#).
- Nelken, Israel, Yaron Rotman, and Omer Bar Yosef (1999). [\[link\]](#).
- Nelken, Israel and Huib Versnel (2000). [\[link\]](#).
- Phillips, Dennis P and Sarah A Sark (1991). [\[link\]](#).
- Pollack, Irwin and J. M. Pickett (1958). [\[link\]](#).
- Rupp, André, Liora Las, and Israel Nelken (2007). [\[link\]](#). Springer.
- Schneider, David M. and Sarah M. N. Woolley (2013). [\[link\]](#).
- Schreiner, Christoph E, Heather L Read, and Mitchell L Sutter (2000). [\[link\]](#).
- Shetake, Jai A. et al. (2011). [\[link\]](#).
- Sutter, Mitchell L (2000). [\[link\]](#).
- Taaseh, Nevo, Amit Yaron, and Israel Nelken (2011). Ed. by Izumi Sugihara. [\[link\]](#).
- Teschner, Magnus J. et al. (2016). [\[link\]](#).
- Thomson, Alex M and Jim Deuchars (1994). [\[link\]](#).
- Ulanovsky, N. (2004). [\[link\]](#).
- Ulanovsky, Nachum, Liora Las, and Israel Nelken (2003). [\[link\]](#).
- Walker, Kerry, Bashir Ahmed, and Jan Schnupp (2008). [\[link\]](#).
- Wehr, Michael and Anthony M Zador (2003). [\[link\]](#).
- Wehr, Michael and Anthony M Zador (2005). [\[link\]](#).
- Wilson, Hugh R and Jack D Cowan (1972). [\[link\]](#).
- Yarden, Tohar S. and Israel Nelken (2017). [\[link\]](#).
- Yaron, Amit, Itai Hershenhoren, and Israel Nelken (2012). [\[link\]](#).
- Zhang, Li I et al. (2003). [\[link\]](#).