

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30

Title: A Circuit Model of Auditory Cortex

Short Title: A Circuit Model of Auditory Cortex

Authors: Youngmin Park¹ and Maria N. Geffen^{1,2*}

¹Department of Otorhinolaryngology: HNS, University of Pennsylvania, Philadelphia, PA, 19104, USA

²Department of Neuroscience, Department of Neurology, University of Pennsylvania, Philadelphia, PA, 19104, USA

*Corresponding author *mgeffen@pennmedicine.upenn.edu*

Conflict of interest: The authors declare no competing financial interests

Acknowledgements: This work was supported by National Institutes of Health (Grant numbers NIH R03DC013660, NIH R01DC014779, NIH R01DC015527, NIH R01NS113241), Klingenstein Award in Neuroscience, Human Frontier in Science Foundation Young Investigator Award and the Pennsylvania Lions Club Hearing Research Fellowship to MGN. MNG is the recipient of the Burroughs Welcome Award at the Scientific Interface. The authors thank the members of the Geffen laboratory, the Hearing Research Center at the University of Pennsylvania, Jason Kim and Lia Papadopoulos for insightful comments and discussions.

31 **Abstract**

32 The mammalian sensory cortex is composed of multiple types of inhibitory and excitatory neurons,
33 which form sophisticated microcircuits for processing and transmitting sensory information. Despite rapid
34 progress in understanding the function of distinct neuronal populations, the parameters of connectivity that
35 are required for the function of these microcircuits remain unknown. Recent studies found that two most
36 common inhibitory interneurons, parvalbumin- (PV) and somatostatin-(SST) positive interneurons control
37 sound-evoked responses, temporal adaptation and network dynamics in the auditory cortex (AC). These
38 studies can inform our understanding of parameters for the connectivity of excitatory-inhibitory cortical circuits.
39 Specifically, we asked whether a common microcircuit can account for the disparate effects found in studies
40 by different groups. By starting with a cortical rate model, we find that a simple current-compensating
41 mechanism accounts for the experimental findings from multiple groups. They key mechanisms are two-fold.
42 First, PVs compensate for reduced SST activity when thalamic inputs are strong with less compensation when
43 thalamic inputs are weak. Second, SSTs are generally disinhibited by reduced PV activity regardless of
44 thalamic input strength. These roles are augmented by plastic synapses. These differential roles reproduce
45 the differential effects of PVs and SSTs in stimulus-specific adaptation, forward suppression and tuning-curve
46 adaptation, as well as the influence of PVs on feedforward functional connectivity in the circuit. This circuit
47 exhibits a balance of inhibitory and excitatory currents that persists on stimulation. This approach brings
48 together multiple findings from different laboratories and identifies a circuit that can be used in future studies
49 of upstream and downstream sensory processing.

50

51 **Significance Statement**

52 The mammalian auditory cortex is composed of multiple types of inhibitory and excitatory neurons,
53 which form sophisticated microcircuits for processing and transmitting sensory information. Distinct inhibitory
54 neuron subtypes play distinct functions in auditory processing, but it remains unknown what simple set of
55 underlying mechanisms is responsible for inhibitory cortical function. Here, we built minimal rate and spiking
56 models and identified a specific set of synaptic mechanisms that could best reproduce the broad set of
57 experimental results in the auditory cortex. The simplicity of our model provides an understanding of inhibitory
58 cortical processing at the circuit level, which explains results from different laboratories, and provides for a
59 novel computational framework for future studies of cortical function.

60

61

62 Introduction

63 Detecting sudden changes in the acoustic environment and extracting relevant acoustic features from
64 noise are important computations for auditory navigation and scene analysis. The mammalian auditory cortex
65 (AC) is a key region for processing temporally patterned sounds (Ulanovsky, Las, and Nelken 2003). Neurons
66 in AC exhibit adaptation to repeated tones, which may be selective for an overrepresented stimulus, such as
67 in stimulus-specific adaptation, or SSA (Ulanovsky, Las, and Nelken 2003; Natan et al. 2015). They
68 furthermore exhibit forward suppression, in which a preceding stimulus masker tone drives a decrease in
69 responses to the subsequent target tone (Phillips and Hasenstaub 2016; Loebel, Nelken, and Tsodyks 2007).
70 How these computations are carried out by cortical circuits has been subject of extensive research.

71 The AC is composed of tightly coupled networks of excitatory and inhibitory neurons. Recent studies
72 have identified the differential involvement of two distinct major classes of inhibitory neurons, parvalbumin-
73 positive (PV) and somatostatin-positive (SST) neurons in these temporal paradigms. These neurons differ
74 morphologically and physiologically (Fino, Packer, and Yuste 2013; Isaacson and Scanziani 2011), and recent
75 studies found that they play differential functions in auditory processing. Specifically, SSTs, but not PVs
76 facilitate stimulus-specific adaptation (Natan et al. 2015). PVs and SSTs play distinct roles in adaptation to
77 repeated tones along the frequency response function of the target neuron (Natan, Rao, and Geffen 2017).
78 SSTs and PVs drive bi-directional effects on forward suppression (Phillips, Schreiner, and Hasenstaub 2017).
79 In addition, PVs enhance feedforward connectivity in the auditory cortex (Hamilton et al. 2013). These
80 experimental results can be used to constrain currents and connections in an idealized auditory cortex model
81 consisting of PVs, SSTs and Exc. Here, we tested whether these results can be accounted for by the same
82 set of mechanisms.

83 In this paper, we build up from a simple dimensionless model consisting of one iso-frequency unit. We
84 transition to a three-unit rate model to understand the mechanisms of inhibitory neural modulation on a gross
85 tonotopy. We then build a detailed spiking model which incorporates the mechanisms discovered in the rate
86 models. The rate models provide a qualitative intuition of the underlying mechanisms, and the spiking models
87 incorporated these mechanisms in an accessible, open-source codebase for future work. We present the
88 results on testing the three-unit rate model and the detailed spiking model in four distinct auditory paradigms.

89 The model accounted for observed experimental results including the differential role of SSTs and
90 PVs in SSA (Natan et al. 2015), forward suppression (Phillips, Schreiner, and Hasenstaub 2017), tuning-
91 curve adaptation (Natan, Rao, and Geffen 2017), and the effects of PV activation on feedforward functional
92 connectivity (Hamilton et al. 2013). We found that compensating currents between the two types of inhibitory
93 neurons explain experimental findings of differential effects of their modulation on excitatory activity.
94 Furthermore, the model is consistent with existing hypotheses regarding inhibitory and excitatory balance in
95 the cortex. This framework can be used to build and test hypotheses for similar phenomena in other sensory
96 modalities, and studies of upstream or downstream auditory processing in AC.

97

98 Materials and Methods

99 We first built an augmented version of the Wilson-Cowan model, consisting of one iso-frequency unit
 100 of the auditory cortex. The model consisted of one excitatory neural population and two inhibitory neural
 101 subpopulations. Importantly, the single iso-frequency unit model served as the template for all other models
 102 in this paper. By using the results and parameters from this model, we extended our results to the substantially
 103 more complex three-unit rate model and three-unit spiking models. We constrained the parameters using
 104 experimental data from the literature.

105 All code used to generate figures (including model simulations, numerical methods, and analysis
 106 methods) are available on GitHub at https://github.com/geffenlab/park_geffen under the MIT open source
 107 license.

108 In figures in this manuscript, we use blue or black lines to depict Exc activity in absence of optogenetic
 109 manipulation (called “Control”); magenta for SST activity; cyan solid for PV activity; orange for Exc activity
 110 under SST suppression or activation; green for Exc activity under PV suppression or activation.

111 Augmented Wilson-Cowan Model

112 We modeled a single iso-frequency unit as an augmented version of the Wilson-Cowan model (Wilson
 113 and Cowan 1972) by including an additional inhibitory subtype. We emphasize that we drew much of our
 114 understanding of adaptation throughout this paper using this single iso-frequency unit:

$$\begin{aligned}
 \tau_u \frac{du(t)}{dt} &= -u(t) + f(w_{ee}u(t) - w_{ep}p(t) - w_{es}s(t) + qg(t)i(t)), \\
 \tau_p \frac{dp(t)}{dt} &= -p(t) + f(w_{pe}u(t) - w_{pp}p(t) - w_{ps}s(t) + I_{\text{Opt,PV}}(t) + qg(t)i(t)), \\
 \tau_s \frac{ds(t)}{dt} &= -s(t) + f(w_{se}u(t) - w_{sp}p(t) - w_{ss}s(t) + I_{\text{Opt,SST}}(t)),
 \end{aligned} \tag{1}$$

116 where $u(t)$, $p(t)$, and $s(t)$ represent the normalized firing rate (scaled from 0 to 1) of the excitatory population,
 117 PV inhibitory subpopulation, and SST inhibitory subpopulation, respectively (Figure 1. Note that connection
 118 strengths in the rate model are given by w_{ij} , and the spiking model are given by conductances g_{ij} , which are
 119 distinct from the thalamic depression variable $g(t)$. We will keep these distinctions consistent throughout the
 120 text, so there is no ambiguity). An important caveat of the rate equations is that the rescaled firing rates are
 121 a consequence of *dimensionless* equations. All weights, parameters, optogenetic inputs, and thalamic inputs
 122 are also dimensionless in the rate model, with the only exception being time.

123 The parameters $I_{\text{Opt,PV}}(t)$ and $I_{\text{Opt,SST}}(t)$ represent the strength of PV and SST activation or
 124 inactivation, respectively, and w_{ij} and τ_i are synaptic weights and time constants, respectively. All time
 125 constants are $\tau_u = \tau_p = \tau_s = 10\text{ms}$, roughly in agreement with known data (Tsodyks et al. 1997; Natan et al.
 126 2015). The function f is a threshold linear function defined as

$$f(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ rx & \text{if } 0 < x \leq 1/r, \\ 1 & \text{if } x > 1/r \end{cases} \tag{2}$$

128 where the function f roughly approximates a sigmoid that converges to zero for small or negative inputs and
 129 saturates to 1 for large inputs. The parameter $r = 3$ determines the gain of all firing-rate functions and was

chosen roughly to be the same as other modeling studies (Yarden and Nelken 2017; Natan et al. 2015). We included a threshold by subtracting a constant u_{th} from the input, i.e., $f(x - u_{th})$ for some input x , where u_{th} is a positive number (typically in the range from 0 to 1). In all rate model simulations, we chose Exc, PV, and SST thresholds to be $u_{th} = 0.7$, $p_{th} = 1$, and $s_{th} = 1$, respectively. The thresholds indicate the minimum activity required for a neural population to affect postsynaptic neural populations. Because the thresholds are greater than zero, sub-threshold activity does not affect the dynamics of the network.

The input function $i(t)$ consists of blocks of inputs with stimulus duration and interval based on the experimental paradigm. We list the stimulus duration and stimulus interval for each paradigm in Table 2 and detail the paradigm in the text and figures where appropriate. When an auditory input arrives into the Exc and PV populations, the default temporal profile is taken to have an instantaneous rise with amplitude q and exponential decay (Figure 1B, bottom red curve) with time constant $\tau_q = 10\text{ms}$, which roughly agrees with known values (Rankin, Sussman, and Rinzel 2015). The instantaneous rise and exponential decay were chosen for simplicity. The input $i(t)$ is further modulated by a slow synaptic depression term g satisfying the standard model of synaptic depression

$$\frac{dg(t)}{dt} = \frac{g_0 - g(t)}{\tau_{d_1}} - \frac{g(t)i(t)}{\tau_{d_2}}, \quad (3)$$

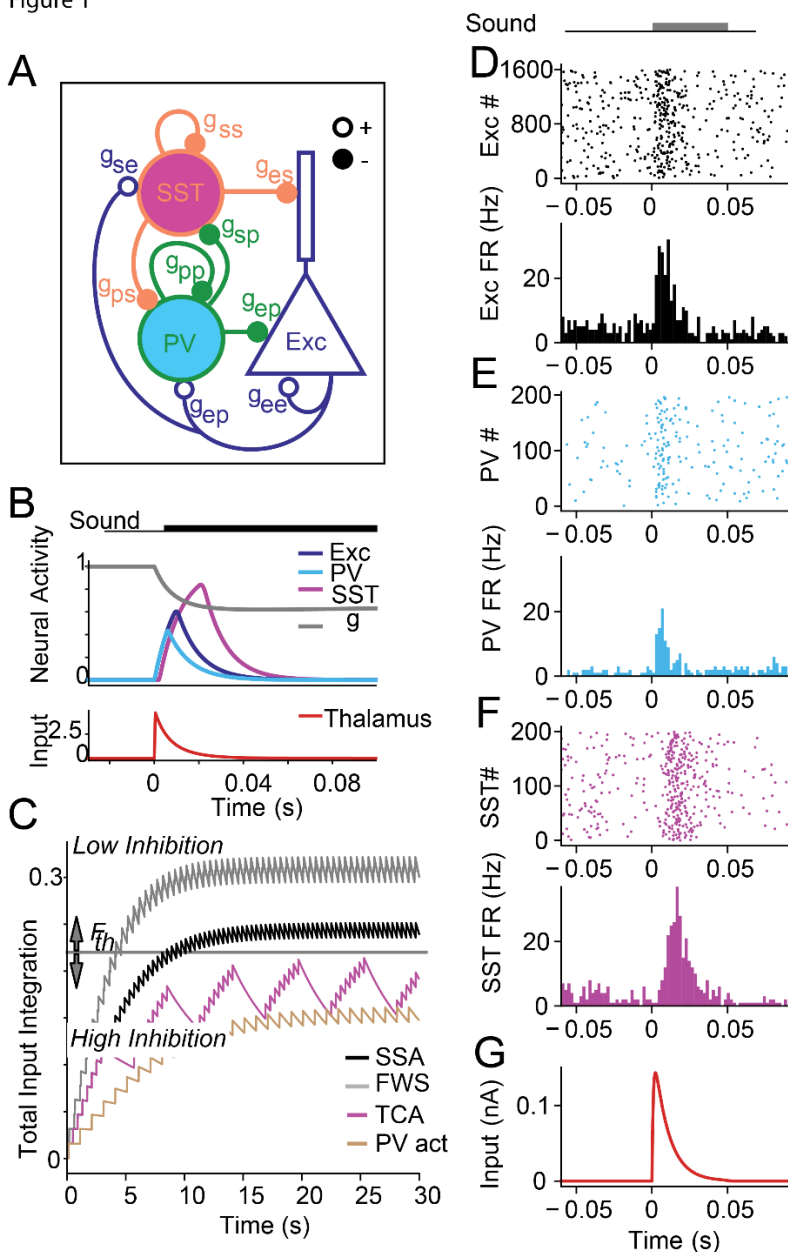
where the time constants are $\tau_{d_1} = 1500\text{ms}$ for replenishment and $\tau_{d_2} = 20\text{ms}$ for depletion (chosen close to reported values (Abbott et al. 1997; Wehr and Zador 2005; Tsodyks et al. 1997; Natan et al. 2015)). The synaptic depression variable g begins at a baseline value of $g_0 = 1$ and when $i(t) > 0$, i.e., when an input arrives into the Exc or PV populations, $g(t)$ decreases on the timescale determined by τ_{d_2} . Because $g(t)$ multiplies the input $i(t)$ to $u(t)$ and $p(t)$ in Equation (1), $g(t)$ serves to modulate the strength of auditory inputs to A1. In the absence of auditory input, $g(t)$ recovers slowly on the order of seconds determined by τ_{d_1} .

We based the proportional strengths of connections in the single-unit model based on existing studies on AC (Pfeffer et al. 2013). The within-unit connectivity is equivalently represented by the matrix,

$$W_1 = \begin{pmatrix} w_{ee} & w_{ep} & w_{es} \\ w_{pe} & w_{pp} & w_{ps} \\ w_{se} & w_{sp} & w_{ss} \end{pmatrix} = \begin{pmatrix} 1.1 & 2 & 1 \\ 1 & 2 & 2 \\ 6 & 0 & 0 \end{pmatrix}. \quad (4)$$

All synaptic weights w_{ij} in the single-unit rate model are constant (Mill et al. 2011), with synaptic depression appearing in the feedforward thalamic inputs (Lee and Sherman 2008). The inhibitory synaptic weights were roughly chosen to agree with known connection types and connection strengths (Womelsdorf et al. 2014; Pfeffer et al. 2013), and the excitatory connections were tuned as free parameters. The constant synapses allowed us to fully understand the model dynamics before transitioning to the more complex three-unit model with depressing and facilitating synapses.

Figure 1



162

Figure 1.A. Model of the PV-SST-Exc circuit (spiking model, where connection strengths are given by conductances g_{ij} . Note that the conductances with subscripts g_{ij} are distinct from $g(t)$, the adaptation variable for the thalamic input. The rate model uses connection strengths w_{ij}). B. Input and response profiles for the single-unit rate and spiking model to a 100 ms long tone. Top: Gray: thalamic depression variable g . Blue: excitatory (Exc) neuron activity. Cyan: PV. Magenta: SST. Bottom: Thalamic input (red). C. Responses to stimulus over the first 30 s after sound onset for the different paradigms modeled in the paper. We used a high and a low inhibition mode of synaptic weights to capture the different results. For SSA (black) and forward suppression (FWS, gray), the variable \bar{F} is higher than threshold \bar{F}_{Th} , resulting in a set of low inhibition parameters. Paradigms for tuning-curve adaptation (purple) and PV activation (g) asymptote at below-threshold levels, resulting in a set of high inhibition parameters. D-G responses of neurons in the spiking model to a 50ms tone. Top: raster plot; Bottom: Firing rate of Exc (D), PVs (E), SSTs (F), and thalamo-cortical input (G).

175

176 **Three-unit Rate Model**

177 Using the single-unit rate model as a template, we arranged copies into three units with lateral cortical and
 178 thalamic connections (Figure 2A. Lateral inhibitory connections are hidden in Figure 2 for clarity). This
 179 arrangement endowed our model with a gross tonotopy, which we used to explore spectrally and temporally
 180 complex auditory inputs. While the three-unit rate model appears to be substantially more complex, the
 181 parameters were strongly constrained by the single-unit rate model. In particular, we aimed for each unit of
 182 the three-unit model to mimic the excitatory and inhibitory currents of the single-unit rate model. We found
 183 that maintaining currents explained many of the known optogenetic experiments in the literature. Before
 184 turning to the spiking model, we briefly describe technical details of the parameter values and functions. Each
 185 unit behaves according to the equations

186

$$\begin{aligned}
 \tau_u \frac{du_i(t)}{dt} &= -u_i(t) + f \left(w_{ee}u_i(t) - (w_{ep} - a(1 - D_i(t)))p_i(t) - w_{es}s_i(t) + J_{1,i}(t) \right), \\
 \tau_p \frac{dp_i(t)}{dt} &= -p_i(t) + f \left(w_{pe}u_i(t) - w_{pp}p_i(t) - w_{ps}s_i(t) + I_{\text{Opt,PV}}(t) + J_{2,i}(t) \right), \\
 \tau_s \frac{ds_i(t)}{dt} &= -s_i(t) + f \left(w_{se}u_i(t) - w_{sp}p_i(t) - w_{ss}s_i(t) + I_{\text{Opt,SST}}(t) + J_{3,i}(t) \right),
 \end{aligned} \tag{5}$$

188 where,

$$J_{1,i}(t) = \begin{cases} -F_i(t)s_2(t) + qI_i(t) + w_{ee}^*u_2(t) & \text{if } i = 1, 3, \\ -F_2(t)(s_1(t) + s_3(t)) + qI_2(t) + \frac{w_{ee}^*(u_1(t) + u_3(t))}{2} & \text{if } i = 2, \end{cases} \tag{6}$$

190 and

$$J_{2,i}(t) = \begin{cases} qI_i(t) + w_{pe}^*u_2(t) & \text{if } i = 1, 3, \\ qI_2(t) + \frac{w_{pe}^*(u_1(t) + u_3(t))}{2} & \text{if } i = 2, \end{cases} \text{ and } J_{3,i}(t) = \begin{cases} w_{se}^*u_2(t) & \text{if } i = 1, 3, \\ \frac{w_{se}^*(u_1(t) + u_3(t))}{2} & \text{if } i = 2. \end{cases} \tag{7}$$

192 The functions $I_i(t)$ are defined as $I_k(t) = g_k(t)i_k(t) + g_2(t)i_2(t)\alpha$, for $k = 1, 3$, and $I_2(t) = (g_1(t)i_1(t) +$
 193 $g_3(t)i_3(t))\alpha + g_2(t)i_2(t)$. We remark that any lower-case letter i that has a subscript and is a function of time,
 194 e.g., $i_1(t), i_2(t), i_3(t)$, represent thalamic inputs. The time-dependent notation for these functions will always
 195 be distinct from the index i . Note that each set of equations are almost identical to the single-unit case, but
 196 with the addition of lateral terms along with facilitating and depressing terms $F_i(t)$ and $D_i(t)$. The lateral terms
 197 are between immediate neighbors and include lateral SST to Exc (facilitating), Exc to Exc, Exc to PV. The
 198 facilitating terms $F_i(t)$ increase from 0 to nonzero values as unit i receives inputs, and the depressing terms
 199 $D_i(t)$ decrease from 1 to lower values as unit i receives inputs. Rather than adjusting weights integrated prior
 200 to facilitating synapses, we controlled the strength of facilitation by changing the time constants in Equation
 201 (8).

202 We chose $\alpha = 0.65$, i.e., 65% of the thalamic inputs to the left or right units reach the center unit.
 203 Likewise, 65% of thalamic inputs to the center unit reach the left and right units. The function f is threshold
 204 linear (Equation 2). The functions $I_k(t)$ are time-dependent inputs with the strongest preference for unit k ,
 205 and the profiles of $i_1(t)$, $i_2(t)$, and $i_3(t)$ are shown in Figure 2F (these profiles are the same as the profile in

the single-unit model, Figure 1B, bottom). Parameters a, b control the strength of depression and facilitation and are chosen to be $a = 0.5, b = 2$. The parameter q controls the strength of all inputs. Each input $I_j(t)$ is modulated by corresponding depression variables $g_k(t)$, where each $g_k(t)$ satisfies Equation 3 independently. The parameters τ_i are membrane time constants and chosen the same as the single-unit model, $\tau_u = \tau_p = \tau_s = 10\text{ms}$ (Tsodyks et al. 1997; Natan et al. 2015). The parameters w_{ij} are within-unit synaptic weights chosen according to Equation 4, while the parameters w_{ij}^* are lateral (between unit) synaptic weights. We chose $w_{ee}^* = 0.667$, $w_{pe}^* = 1.25$, and $w_{se}^* = 0.125$ to reflect the generally weaker lateral synaptic strengths in auditory cortex relative to the within-unit connections (Levy and Reyes 2012).

Figure 2

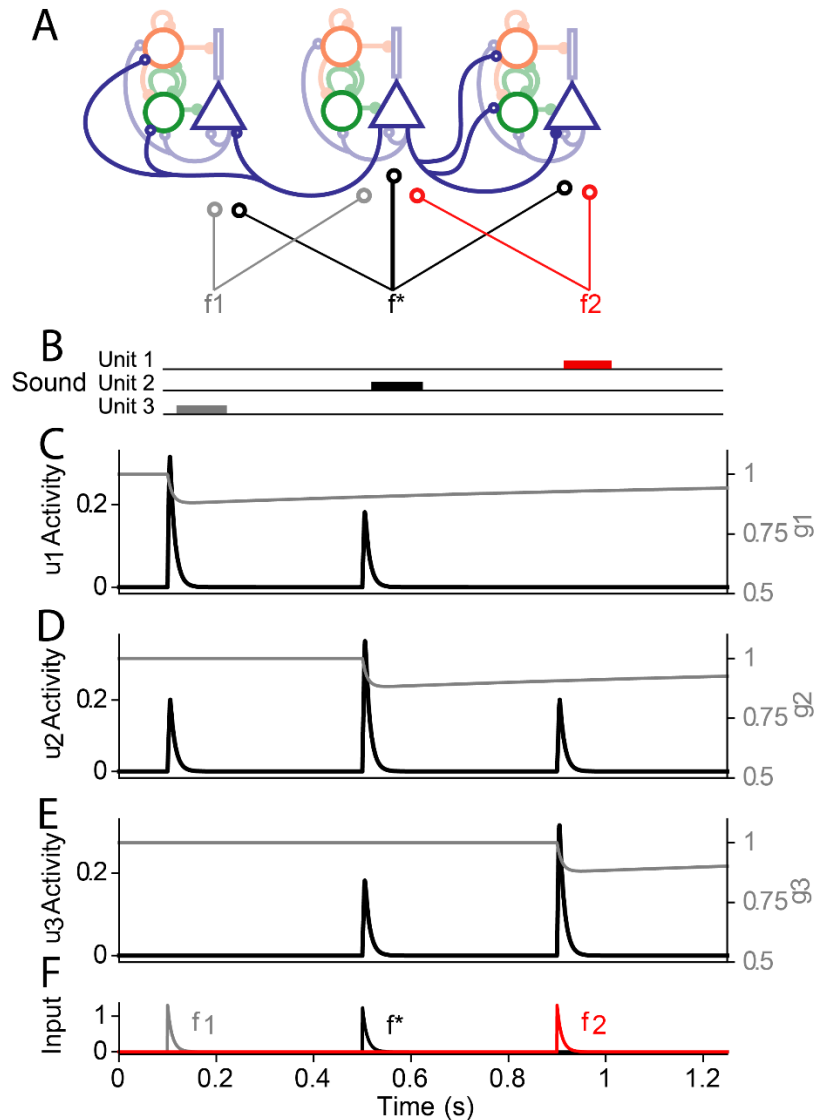


Figure 2. Input and response profiles of the three-unit model. A: The three-unit rate model of the auditory cortex, with three preferred frequencies, f_1 , f^* , and f_2 (the spiking model follows the same motif). Non-excitatory lateral connections have been hidden for clarity. B: 50ms auditory inputs are applied at each frequency in sequence. C—E: Black traces show the excitatory cortical response of the first (u_1), second (u_2), and third (u_3) rate units, respectively. Gray traces show the slow synaptic depression. F: The traces of the thalamic inputs: f_1 (gray), f^* (black), and f_2 (red). Each iso-frequency unit contains lateral excitatory

connections where the Exc population of a given unit synapses laterally onto the neighboring Exc, PVs, and SSTs.

We added facilitating terms $F_i(t)$ in the Exc to SST synapses, and depressing terms $D_i(t)$ in the PV to Exc synapses (Beierlein, Gibson, and Connors 2003). The parameters a and b control the degree of depression and facilitation, respectively, where we chose $a = 0.5, b = 2$ (these values were not taken from the literature). The depressing parameter a was chosen so that the term $(w_{ep} - aD_i(t))$ did not change sign across experimental paradigms. The facilitating variables $F_i(t)$ satisfy

$$\frac{dF_j(t)}{dt} = -\frac{F_j(t)}{\tau_{D_1}} + \frac{i_j(t)}{\tau_{D_2}}, \quad (8)$$

where τ_{D_1} and τ_{D_2} are as in Equation 3. The reuse of the depression time constants τ_{D_i} are an intentional and simplifying choice. By using the same time scales, we were able to use the inputs $i_j(t)$ as a proxy for the excitatory activity $u_j(t)$ and simulate the facilitation variable in terms of the depressing synaptic variable as $F_j(t) = 1 - g_j(t)$. Similarly, the depression variables $D_i(t)$ satisfy

$$\frac{dD_j(t)}{dt} = \frac{1 - D_j(t)}{\tau_{D_1}} - \frac{D_j(t) i_j(t)}{\tau_{D_2}}, \quad (9)$$

and used the thalamic input as a proxy for excitatory activity to simulate the depression variable as $D_i(t) = g_i(t)$. All depressing and facilitating timescales were chosen close to reported values (Abbott et al. 1997; Tsodyks et al. 1997; Wehr and Zador 2005).

The activity of the model is shown in Figure 2, in response to three successive auditory stimuli are applied in order of the frequencies f_1, f^* , and f_2 , stimulating the left, center, and right units, respectively (Figure 2C-E). The center unit ($u_2(t)$) responded equally well to frequencies f_1 and f_2 (Figure 2D), a necessary response for SSA paradigms. For simplicity, activation of an adjacent unit did not affect the thalamic variable, i.e., $g_1(t)$ was left unaffected by $i_2(t)$, $g_2(t)$ was left unaffected by $i_1(t)$ and $i_3(t)$, and $g_3(t)$ was left unaffected by $i_2(t)$. We assumed that the frequency difference between f_1 and f_2 was great enough that auditory inputs at f_1 (f_2) did not affect units responsive to f_2 (f_1).

Table 1: Parameter values of rate models across paradigms. In the optogenetic parameters, numbers without parentheses show optogenetic strengths for use in model reproductions and numbers in parentheses show optogenetic strengths for use in model predictions. Positive optogenetic numbers correspond to activation and negative numbers correspond to inactivation.

| Parameter | SSA (Simple) | SSA | FWS | TCA | PV Act. | Notes |
|-------------------------|--------------|----------|-------------|------------|---------|-----------------------------|
| q | 5 | 5 | 1.3 | 5 | 5 | Fitted, dimensionless, rate |
| $I_{\text{Opt,PV}}(t)$ | -2 | -4 (0.5) | -0.5 (0.25) | -0.5 (1.2) | 2 (2) | Fitted, dimensionless, rate |
| $I_{\text{Opt,SST}}(t)$ | -1 | -2 (1.2) | -0.5 (0.1) | -1 (0.1) | - | Fitted, dimensionless, rate |

| | | | | | | |
|------------------|---|--------|-------|-----|-------|-----------------------------------|
| q | 5 | 5 | 1.3 | 5 | 5 | Fitted, dimensionless, spiking |
| $I_{Opt,PV}(t)$ | - | -0.2nA | 0.1nA | 1nA | 0.5nA | Fitted, dimensionless, spiking |
| $I_{Opt,SST}(t)$ | - | -1nA | 2nA | 1nA | - | Fitted, dimensionless, spiking |

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

When matching the model outputs to experiments in the literature, we found that there were seemingly two different parameter sets that explained all phenomena and we were unable to tweak the rate model to unify the parameters. We unified the disjoint parameter sets by incorporating paradigm-dependent baseline states in the three-unit rate model. The model parameters switch between weak and strong baseline inhibition, where weak inhibition corresponds to high thalamic activity (which corresponds to the parameters in Equation (4)), and strong inhibition corresponds to relatively low thalamic activity (which corresponds to the parameters in Equation (11)). This idea was implemented using an additional facilitating variable,

$$\frac{d\bar{F}(t)}{dt} = -\frac{\bar{F}^2(t)}{\tau_{F_1}} + \frac{\bar{I}(t)}{\tau_{F_2}}, \quad (10)$$

where $\bar{I}(t)$ represents all thalamic inputs of any frequency, and the parameters are chosen to be $\tau_{F_1} = 1500$, and $\tau_{F_2} = 100$. As we applied auditory inputs from particular experiments into the model, the function $\bar{F}(t)$ grew as a function of the excitatory input function $I(t)$, and eventually saturated to different mean values based on the stimulus duration, stimulus interval, and inter-trial interval. Incidentally, we found that the facilitating variable $\bar{F}(t)$ saturated to greater values for forward suppression and SSA, and saturated to relatively lower values for PV activation and tuning-curve adaptation. The two putative disjoint parameter sets correspond exactly to these relatively high and low values of the variable $\bar{F}(t)$. A simulation of Equation 10 is shown in Figure 1C for the various auditory paradigms, with a horizontal line shown where we chose to separate the saturation values.

If $\bar{F}(t)$ is above the threshold, which we chose to be $F_{th} = 0.22$, the system exhibits weak baseline inhibition, and the synapses take baseline values as shown in Equation 4. On the other hand, if $\bar{F}(t) < F_{th}$, the synapses take the strong baseline inhibitory values

$$W_2 = \begin{pmatrix} 1.1 & 3 & 3 \\ 1 & 2 & 2 \\ 6 & 0 & 0 \end{pmatrix}, \quad (11)$$

and the SST activity threshold, $s_{th} = 1$, decreases to $s_{th} = 0$. In our original model, we chose a smooth transition between these parameter sets, i.e., $w_{ep}(t) = 2h(F(t)) + 3(1 - h(F(t)))$, $w_{es}(t) = h(\bar{F}(t)) + 3(1 - h(\bar{F}(t)))$, and $s_{th}(t) = 1h(\bar{F}(t))$ where

$$h(x) = \frac{1}{1 + \exp(-r(x - F_{th}))}, \quad (12)$$

and r , the gain of the sigmoid h was chosen to be steep, e.g., $r = 25$. However, for simplicity, we replaced h with a Heaviside function and assumed that the system already reached either the weak baseline inhibition W_1 (Equation 4), or the strong baseline inhibition W_2 (Equation 11) based on the given experimental paradigm.

279

Table 2: Auditory paradigm parameters. SSA parameters from (Natan et al. 2015). Forward suppression (FWS) parameters from (Brosch and Schreiner 1997; Phillips, Schreiner, and Hasenstaub 2017). Tuning-curve adaptation (TCA) parameters from (Natan, Rao, and Geffen 2017). PV activation parameters from (Hamilton et al. 2013). Optogenetic inhibition was performed 100ms before tone onset and 100ms after tone offset in SSA and FWS. In TCA and PV activation, optogenetic inhibition was turned on at the beginning of the experiment and sustained through the trial. For optogenetic parameters, see Table 2.

| | SSA | FWS | TCA | PV Act. |
|-------------------------|-------|-------|--------|---------|
| Stimulus duration | 100ms | 50ms | 100ms | 50ms |
| Inter-stimulus interval | 300ms | 20ms | 300ms | - |
| Inter-trial interval | - | 380ms | 2400ms | 950ms |
| Stimuli per trial | - | 2 | 8 | 1 |

286

For each paradigm (with paradigm parameters shown in Table 2, we simulated Equation 10 and found that SSA and forward suppression belonged to the weak inhibitory regime ($\bar{F}(t)$ integrated to values above threshold F_{th}), whereas tuning-curve adaptation and PV activation belonged to the strong inhibition regime ($\bar{F}(t)$ integrated to values below threshold F_{th}). All rate models were simulated using the dynamical systems software XPP (Ermentrout 2002), called using PyXPPCALL and visualized using Python (Rossum 1995). The equations were well-behaved enough that a very coarse time step of $dt=0.1$ was sufficient for our purposes. We used the default integration method in XPP, the fourth-order Runge-Kutta.

294

295 Goodness of Fit and Related Metrics

We compared the results obtained with the model as result of optogenetic perturbations to published data qualitatively, by testing whether the model reproduces the increase/decrease of activity that was reported. The limit of large numbers yields theoretically tractable equations such as the Wilson-Cowan models we use in this paper. However, finite size effects may contribute to additional issues that the large-number limit does not address. How do we know that the observed changes would hold consistently with the inclusion of finite and noisy neural populations? To address this question, we built a spiking model constrained by the rate models above. Because the spiking model is derived directly from the simpler rate models, we generally expect that the spiking models will reproduce the rate models results for sufficiently large neural populations, however, we are able to establish that the results are consistent in the presence of noise. Moreover, we built the rate model using Python and brian2, which are both extensively documented and accessible open-source packages. Finally, we include the spiking model as a means for researchers to directly fit parameters from

experimental data if needed. The rate model provides strong qualitative intuition, but does not explicitly account for single-cell interactions. Thus, the spiking model provides an additional, detailed framework for others to modify and extend beyond what the rate model may provide.

Spiking Neuron Dynamics

We use the single-unit rate model as a template for the spiking model, constraining the parameters while preserving the pattern of excitatory and inhibitory currents. All inhibitory neurons consisted of a single somatic compartment, while the excitatory neurons were modeled as two-compartment, “ball-and-stick” models. For each excitatory and inhibitory neuron, we modeled the somatic (ball) compartment as an adaptive exponential integrate-and-fire neuron (Brette and Gerstner 2005; Litwin-Kumar, Rosenbaum, and Doiron 2016):

$$C_m \frac{dV_i^A(t)}{dt} = I_i^A - g_L(V_i^A(t) - E_L) - w_i^A(t) + g_L^A \Delta_T^A e^{\frac{V_i^A(t) - V_T^A}{\Delta_T^A}} \quad (13)$$

where the transmembrane currents are $I_i^A(t) = I_{\text{Syn},i}^A(t) + I_{\text{Baseline}}^A(t) + I_{\text{Thal}}^A(t) + I_{\text{Opto}}^A(t)$, where

$$I_{\text{Syn},i}^A(t) = - \sum_B g_{AB,i}(t)(V_i^A(t) - E^B), \quad (14)$$

and the sum \sum_B in the synaptic current iterates over the presynaptic neurons, $B \in \{e, p, s\}$. If the presynaptic neuron B is excitatory or inhibitory, then $E^B = 0\text{mV}$ or -67mV , respectively. If a synaptic connection existed from PV to Exc, we included a depression variable, $D(t)$, satisfying Equation 9, with $\tau_{D_1} = 1000$, and $\tau_{D_2} = 250$:

$$I_{\text{PV},i}^E(t) = -g_{ep,i}(t)(aD(t))(V_i^E(t) - E^{\text{PV}}), \quad (15)$$

where $a = 1.7$. As in the rate model, the parameter a was chosen such that the sign of $I_{\text{PV},i}^E(t)$ did not change. The additional depression term was necessary to incorporate depression effects that operate well beyond the timescale of inhibitory conductances (Wehr and Zador 2005).

We added transmembrane noise in the form of a white noise process with zero mean and a standard deviation of 20mV to simulate intrinsic and extrinsic current fluctuations. All fixed parameters for each neuron type are shown in Table 3. The parameters that we varied manually were entirely contained in the time-dependent functions $I_{\text{Thal}}^A(t)$ (thalamic inputs) and $I_{\text{Opto}}^A(t)$ (optogenetic parameters). The thalamic input profile, $I_{\text{Thal}}^A(t)$, is determined by

$$I_{\text{Thal}}^A(t) = qI_{\text{Fast}}^A(t)D_{\text{Slow}}^A(t)D_{\text{Fast}}^A(t), \quad (16)$$

where

$$\begin{aligned}
\frac{dD_{\text{Slow}}^A(t)}{dt} &= \frac{1 - D_{\text{Slow}}^A(t)}{\tau_{D_1}} - \frac{D_{\text{Slow}}^A(t)I^A(t)}{\tau_{D_2}} \\
\frac{dD_{\text{Fast}}^A(t)}{dt} &= \frac{1 - D_{\text{Fast}}^A(t) - I^A(t)}{\tau_{D,\text{Fast}}} \\
\frac{dI_{\text{Fast}}^A(t)}{dt} &= \frac{-I_{\text{Fast}}^A(t) + I^A(t)}{\tau_I},
\end{aligned} \tag{17}$$

where $\tau_I = 1\text{ms}$, $\tau_{D,\text{Fast}} = 10\text{ms}$, $\tau_{D_1} = 1000\text{ms}$, and $\tau_{D_2} = 250\text{ms}$. The functions $I^A(t)$ (distinct from $I_{\text{Thal}}^A(t)$) are square wave functions that are active for the duration of the auditory stimulus. Just as in the rate model, the thalamic input function $I_{\text{Thal}}^A(t)$ only appears in Exc and PV neurons. The profile of the thalamic input is shown in Figure 1G. The optogenetic term, $I_{\text{Opto}}^A(t)$, only appears in the PV and SST equations.

Following a presynaptic spike from neuron j , the postsynaptic effect on neuron i appears as an instantaneous spike in the postsynaptic conductance $g_{ij}(t) \rightarrow g_{ij}(t) + g_{ij,\text{max}}/n_X$, where $g_{ij,\text{max}}$ is given by Equation 18, and X stands for the presynaptic neuron type (Exc, PV, or SST). The magnitude of the conductances were chosen to have the same proportion as reported values (Pfeffer et al. 2013), with the same type of connectivity structure as in the rate model .

$$G_{\text{max}} = \begin{pmatrix} g_{ee,\text{max}} & g_{ep,\text{max}} & g_{es,\text{max}} \\ g_{pe,\text{max}} & g_{pp,\text{max}} & g_{ps,\text{max}} \\ g_{se,\text{max}} & g_{sp,\text{max}} & g_{ss,\text{max}} \end{pmatrix} = \begin{pmatrix} 20 & 40 & 20 \\ 2 & 40 & 40 \\ 120 & 0 & 0 \end{pmatrix} \text{nS}. \tag{18}$$

In the absence of presynaptic spikes, the conductances g_{ij} decay exponentially to zero:

$$\frac{dg_{ij}(t)}{dt} = -\frac{g_{ij}(t)}{\tau_{ij}}, \tag{19}$$

where $\tau_{ij} = 1\text{ms}$ for all synapses except for the time constants from excitatory to PVs, $\tau_{pe} = 25\text{ms}$, and excitatory to SSTs, $\tau_{se} = 15\text{ms}$ (Markram, Wang, and Tsodyks 1998). In the spiking model, we switched to the weak inhibitory regime by decreasing the inhibitory inputs into Exc from $g_{ep,\text{max}} = 40$ and $g_{es,\text{max}} = 20$ to $g_{ep,\text{max}} = 38$ and $g_{es,\text{max}} = 19$.

For excitatory neurons ($A = e$), the transmembrane currents are $I_i^A(t) = I_{\text{Syn},i}^A(t) + I^A(t) + I_{\text{Dend},i}(t)$, where

$$I_{\text{Dend},i}(t) = -\frac{g_{sd,i}(t)(1 + bF(t))(V_E(t) - V_{D,i}(t))}{1 - \kappa}. \tag{20}$$

The term $F(t)$ is a dimensionless slow timescale facilitation variable that depends on the thalamic drive, and satisfies Equation 8 (just as in depression, the additional slow timescale allows the model to operate on multiple timescales (Wehr and Zador 2005)). The parameter $b = 3$ modulates the facilitation strength, and $\tau_{F_1} = 1000$, and $\tau_{F_2} = 250$. For simplicity, we allowed $F_i(t)$ to vary continuously over time. The variable w represents spike-frequency adaptation and obeys

$$\tau_w \frac{dw_E(t)}{dt} = a(V_E(t) - E_L) - w_E(t). \tag{ }$$

The dynamics of the dendritic (stick) compartment obey

$$C_m \frac{dV_D(t)}{dt} = -g_L(V_D(t) - E_L) - \frac{g_{sd}(t)(V_D(t) - V_E)}{\kappa} - g_{es}(t)(V_D(t) - E_I), \quad (22)$$

where the parameter $\kappa = 0.3$ is the ratio of somatic to total surface area (Litwin-Kumar, Rosenbaum, and Doiron 2016).

For PV and SST interneurons, the equations are the same as Exc except that there is no dendritic component. Parameters differ marginally between PV and SST neurons (see Table 3). SSTs, unlike PVs, have no incoming synaptic connections from the thalamus and only receive excitatory inputs from Exc. PVs and SSTs include the optogenetic stimulus term $I_{\text{Opto}}^A(t)$, and as mentioned above, only Exc and PVs contain the thalamic input term $I_{\text{Thal}}^A(t)$. These connections reflect the choices made in the rate model.

Table 3: Parameter values of spiking neurons.

| | Exc | Dend | PV | SST |
|----------------------------|-------|-------|------|-------|
| C_m (pF) | 180 | 180 | 80 | 80 |
| E_L (mV) | -60 | -60 | -60 | -60 |
| g_L (nS) | 6.25 | 6.25 | 5 | 5 |
| Δ_T (mV) | 1 | - | 0.25 | 1 |
| V_T (mV) | -40 | - | -40 | -45 |
| V_{reset} (mV) | -60 | - | -60 | -60 |
| g_{sd} (nS) | 18.75 | 18.75 | - | - |
| I_{baseline} (nA) | 0.35 | - | 0.05 | 0.025 |

Three-unit Spiking Model

We introduced the gross tonotopy into the spiking model by copying the single unit spiking model into three units with lateral excitatory connections (Figure 2A). As in the 3-unit rate model, the thalamic inputs of the 3-unit spiking model have weaker lateral connections. For tone responses at frequency f_1 and f_2 , the center unit receives an input of amplitude proportional to 0.85 that of the left and right units, respectively.

The spiking model contains 1600 Exc, 200 PVs, and 200 SSTs. For connection probabilities within units, we chose $E \leftarrow E$ connections to have probability $p^{EE} = 0.1$ and all other probabilities to be the same, $p^{EE} = p^{ES} = p^{PE} = p^{PP} = p^{PS} = p^{SE} = 0.6$. For lateral connection probabilities, we chose $p = 0.1$. The spiking model was constructed using Brian2 (Stimberg, Brette, and Goodman 2019).

Results

Differential effects of interneuron suppression in stimulus-specific adaptation

Almost all neurons (95%) in AC exhibit stimulus-specific adaptation, a phenomenon in which neurons reduce their response selectively to the inputs that is presented frequently in the stimulus (standard tone),

393 while preserving the initial strong response to the less frequent input (deviant, or oddball tone) (Ulanovsky,
394 Las, and Nelken 2003). Previous studies found that following a presentation of the deviant tone, the excitatory
395 neurons adapt over successive presentations of the standard (von der Behrens et al. 2009; Ulanovsky, Las,
396 and Nelken 2003). Similar adaptation was observed in the songbird (Bell, Phan, and Vicario 2015; Lu and
397 Vicario 2017). This phenomenon was largely attributed to feedforward thalamo-cortical depressing synapses
398 (Mill et al. 2012; 2011), but such models could not account for the full range of the effects that were observed
399 (Yarden and Nelken 2017). A recent study found that inhibitory neurons exhibit differential control over SSA
400 (Natan et al. 2015). Suppressing SSTs resulted in disinhibition of the excitatory neurons, such that
401 disinhibition increased with successive presentations of the standard tone, but not the deviant. By contrast,
402 PV inhibition drove equal amount of disinhibition of excitatory neurons in response to both the deviant and
403 the standard. These results suggest that SST inhibition controls adaptation level of excitatory neurons.

404 In order to understand the roles of inhibitory interneurons in modulating SSA responses, we tested
405 whether the single-unit model could reproduce the differential effects of suppressing PVs and SSTs in
406 temporal adaptation. This paradigm involved the application of 8 successive tones to the single iso-frequency
407 circuit with constant synapses and depressing thalamic inputs (Figure 3, Equation 1). We then tuned the
408 parameters of the model in order to achieve the qualitative results for SSA (Natan et al. 2015). We simulated
409 optogenetic suppression of inhibitory neurons by defining the functions $I_{\text{Opt,PV}}(t)$ and $I_{\text{Opt,SST}}(t)$ to turn on
410 100ms before tone onset and turn off 100ms after tone onset, thereby inhibiting the activity of PV or SST
411 neurons as performed in the experimental paradigms. The degree of PV and SST inhibition was chosen as a
412 free parameter, and in the case of this single-unit rate model, had the dimensionless values of $I_{\text{Opt,PV}}(t) = -2$
413 during PV suppression and $I_{\text{Opt,SST}}(t) = -1$ during SST suppression. Through this suppression of inhibition,
414 we sought to reproduce the constant disinhibition from PV inactivation, and increasing disinhibition from SST
415 inactivation. Once the model reproduced this qualitative result, we discovered that the key mechanisms
416 involve the temporal structure of the responses: PVs exhibit a temporally fast tone-evoked response and peak
417 earlier than Exc and SSTs, while SSTs exhibit a temporally delayed and broad tone-evoked response (Figure
418 1B, Figure 3B top left plot), which both agreed with earlier studies (Natan et al. 2015). The SST delay is not
419 hard-coded, but the result of SSTs receiving indirect thalamic excitation through the Exc population (Pfeffer
420 et al. 2013).

421 The simplicity of the model allowed us to understand the mechanism underlying these changes. One
422 important observation is the faster temporal activation of PVs relative to SSTs. With this property, excitatory
423 activity is immediately affected by changes to PV activity and less by changes to SST activity. With PV
424 suppression (Figure 3C middle row), PVs are reduced in activity, leading to greater Exc activity. The SSTs
425 were indeed disinhibited due to a lack of inhibition from PVs and the greater Exc activity, but the inhibition
426 from SST to Exc was not strong enough to compensate for changes in PV activity. Thus, the Exc population
427 received an overall decrease in inhibitory current across all successive tones, resulting in constant
428 disinhibition prior to and following adaptation (Figure 3D green). This result suggests that the temporally faster

PV inhibition of Exc neurons is necessary for modulating excitatory activity, and only a few milliseconds of earlier PV activity results in a substantially stronger inhibitory effect relative to SSTs.

With SST inactivation, Exc activity at the first tone was virtually unaffected because the reduced SST activity resulted in PV disinhibition, and the increased PV activity resulted in no net change to the total inhibitory current entering the Exc population (Figure 3C, bottom left). PVs compensated for the reduced SST activity. Following adaptation, the overall reduced excitatory activity in both thalamus and Exc resulted in reduced PV activity and a net loss of inhibitory current in Exc. The Exc population was therefore disinhibited at the last tone (Figure 3C, bottom right). Combined, these effects drove an increase in disinhibition over successive tones (Figure 3D, orange). These simple mechanisms of disinhibition and compensation can therefore explain the complementary roles of inhibitory interneurons in shaping cortical activity.

Figure 3

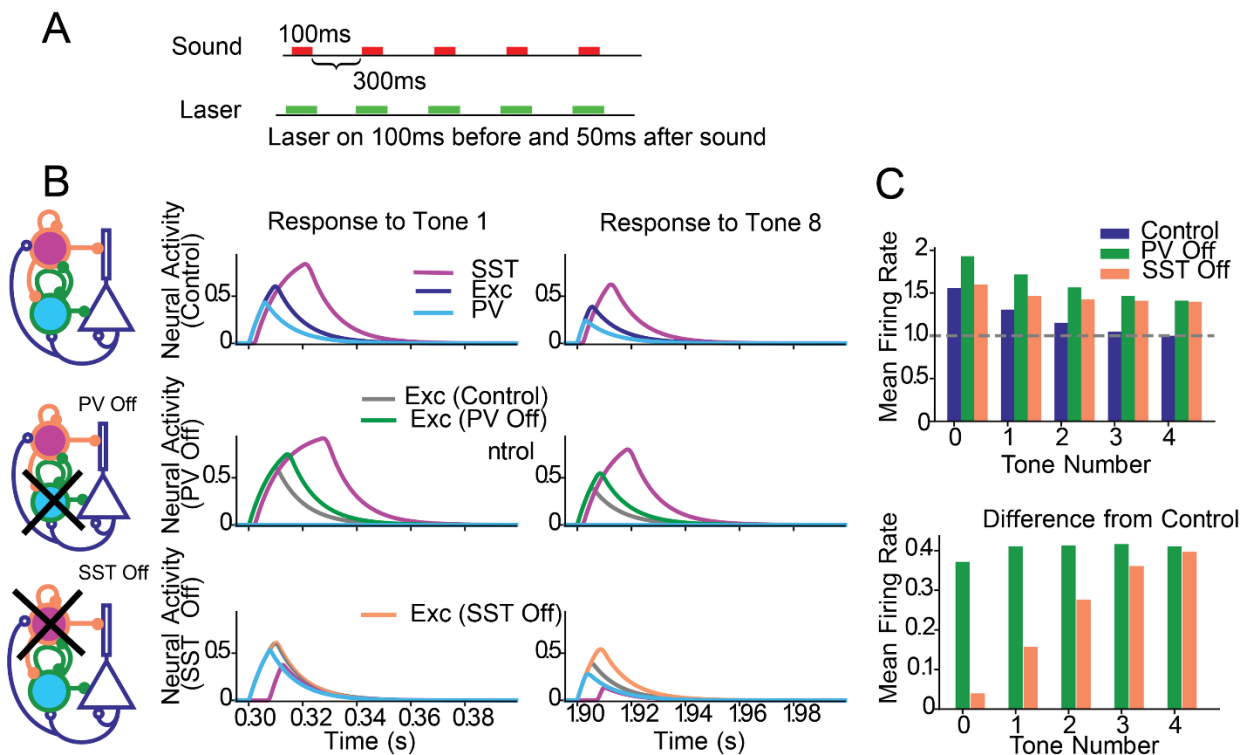


Figure 3. The effect of optogenetic manipulations on adaptation to repeated tones. **A.** Stimulus of repeated tones, with or without concurrent laser stimulation. **B.** Left column: Circuit diagram specifying the inactivation of populations. Top row: no stimulation. Responses of Exc (blue), PV (cyan) and SST (magenta) populations to the first (middle column) and last (right column) tones. Middle row: responses during PV suppression. Green: Responses of Exc under PV suppression, gray: responses of Exc in the control condition. Bottom row: responses during SST suppression. Orange: Responses of Exc under SST suppression, gray: responses of Exc in the control condition. **C.** Top: Mean response of the excitatory population to the repeated tones. Bottom: difference in excitatory responses with and without stimulation. No stimulation (blue), PV suppression (green); SST suppression (orange).

To understand the model dynamics in the presence of inputs at multiple frequencies in the oddball paradigm, we extended the model to a rate and spiking model with three iso-frequency units, in which each microcircuit received inputs of specific preferred frequencies. The three-unit circuitry was based on the single-unit model and the free parameters tuned to reproduce the inhibitory and excitatory currents. For example,

454 an auditory input to the left unit caused lateral excitatory and inhibitory currents to enter the center unit. These
455 currents to the center unit were designed to be similar to the currents entering the single-unit in response to
456 auditory stimuli. We performed the same procedure for auditory inputs to the right unit: lateral excitatory and
457 inhibitory currents from the right were designed to enter the center unit in a manner similar to the single-unit
458 case. Using this procedure, we extended the differential SST and PV inhibition in the single-unit model to
459 work in the case of a tonotopy without the need for exhaustive parameter fitting.

460 The oddball stimulus activated the neighboring units to the center one, whose activity we measured
461 (Stimulus described in Table 2, Figure 4A). We observe thalamic depression in the thalamus which decreases
462 responses to the repeated activation from one unit, and large responses to the unrepeated tone. We simulated
463 PV and SST suppression with the functions $I_{\text{Opt,PV}}(t)$ and $I_{\text{Opt,SST}}(t)$. These functions were turned on 100ms
464 before tone onset and turned off 100ms after tone onset. In the rate model, the inhibition was dimensionless
465 and free parameters chosen to be $I_{\text{Opt,PV}}(t) = -4$ during PV inhibition and $I_{\text{Opt,SST}}(t) = -2$ during SST
466 inhibition. In the spiking model, we chose $I_{\text{Opt,PV}}(t) = -0.2\text{nA}$ during PV inhibition and $I_{\text{Opt,SST}}(t) = -1\text{nA}$
467 during SST inhibition (Figure 4B—G).

A Sound
Unit 1
Unit 2
Unit 3
Laser
100ms
300ms
Deviant (10%)
Standard (90%)
Laser on 100ms before and 50ms after sound

B Rate Model
Normalized Mean FR
Control
PV off
SST off
0 1 2 3 4

C Spiking Model
Normalized Mean FR
0 1 2 3 4

D Experiment (Natan, 2015)
PV Off
FR Change
Post-Deviant Tone #
Rate Model (Reproduction)
FR Change
Post-Deviant Tone #
Spiking Model (Reproduction)
FR Change
Post-Deviant Tone #

E Experiment (Natan, 2015)
SST Off
FR Change
Post-Deviant Tone #
Rate Model (Reproduction)
FR Change
Post-Deviant Tone #
Spiking Model (Reproduction)
FR Change
Post-Deviant Tone #

F Rate Model (Prediction)
Normalized Mean FR
Control
PV On
SST On
0 1 2 3 4

G Rate Model (Prediction)
FR Change
Post-Deviant Tone #

H Rate Model (Prediction)
FR Change
Post-Deviant Tone #

Figure 4. Summary of SSA in the rate and spiking model. A: Oddball stimulus consisted of two tones: standard tones (gray) appear with 90% probability, whereas deviant tones (red) appear with 10% probability. B, C: Average response of the excitatory population to the deviant (red outline) and subsequent standards (gray outline) without stimulation; with PV suppression (green) and with SST suppression (orange). B. Rate model. C. Spiking model. D. Additive change in response of excitatory population due to PV suppression in the rate and spiking models to the deviant (red outline) and standards (gray outline). Left: from published data. Center:

rate model. Right: Spiking model. E. Additive change in response of excitatory population due to SST suppression in the rate and spiking models to the deviant (red outline) and standards (gray outline). Left: from published data. Center: rate model. Right: Spiking model. F. Predictions for the responses to the oddball stimulus with and without interneuron activation in the rate model. Left: Mean responses of the excitatory population to the deviant and subsequent standards (red/gray outline: no activation; green: PV activation; orange: SST activation). Middle: Change in excitatory neuron responses due to PV activation, Right: Identical plot as the middle panel, but for SST activation. PV activation resulted in a near-uniform decrease in FRs, whereas SST resulted in an increase in adaptation.

In the rate and spiking model, the firing rates increased uniformly across all post-deviant tones (Figure 4D,E). In the rate and spiking model, the firing rates exhibited an increase in disinhibition as a function of post-deviant tone number (Figure 4F,G). Both results agree with existing results in SSA (Natan et al. 2015). In order to establish the robustness of these results, we varied several parameters and measured the Common-contrast SSA Index (CSI) (Yarden and Nelken 2017),

$$CSI = \frac{d(f_1) + d(f_2) - s(f_1) - s(f_2)}{d(f_1) + d(f_2) + s(f_1) + s(f_2)}, \quad (23)$$

where $d(f_i)$ is the deviant rate response and $s(f_i)$ is the standard rate response to frequency f_i . For full adaptation, when the standard responses are 0, $CSI = 1$, indicating a high degree of SSA. If the standard responses are equal to the deviant responses, then $CSI = 0$, indicating a low degree of SSA.

We performed a parameter sweep with four key parameters of circuit connectivity (Figure 5): (1) recurrent excitation (w_{ee} , Figure 5A,B,E,F), a key parameter considered in many studies (Yarden and Nelken 2017), especially those related to inhibitory stabilized networks (ISNs) (Tsodyks et al. 1997; Litwin-Kumar, Rosenbaum, and Doiron 2016); (2) timescale of thalamic depression (τ_{d_1} , Figure 5C,D,G,H) whose reported values vary over a large range, from 0.8s to 3s (Yarden and Nelken 2017; Natan et al. 2015); (3) the strength of PV activation or inactivation (Figure 5A,C,E,G); and (4) the strength of SST activation or inactivation (Figure 5B,D,F,H). In all cases, inactivating SSTs reduced the CSI relative to PV inactivation, reflecting the increasing disinhibition over post-deviant tones. The x-axis of each plot corresponds to the strength of the optogenetic laser. Negative values correspond to a decrease in inhibitory activity, and positive values correspond to an increase in inhibitory activity. These effects may be more clearly seen in the model equations (Equations 1, 5, and 13), where we add the terms $I_{Opt,PV}(t)$ and $I_{Opt,SST}(t)$.

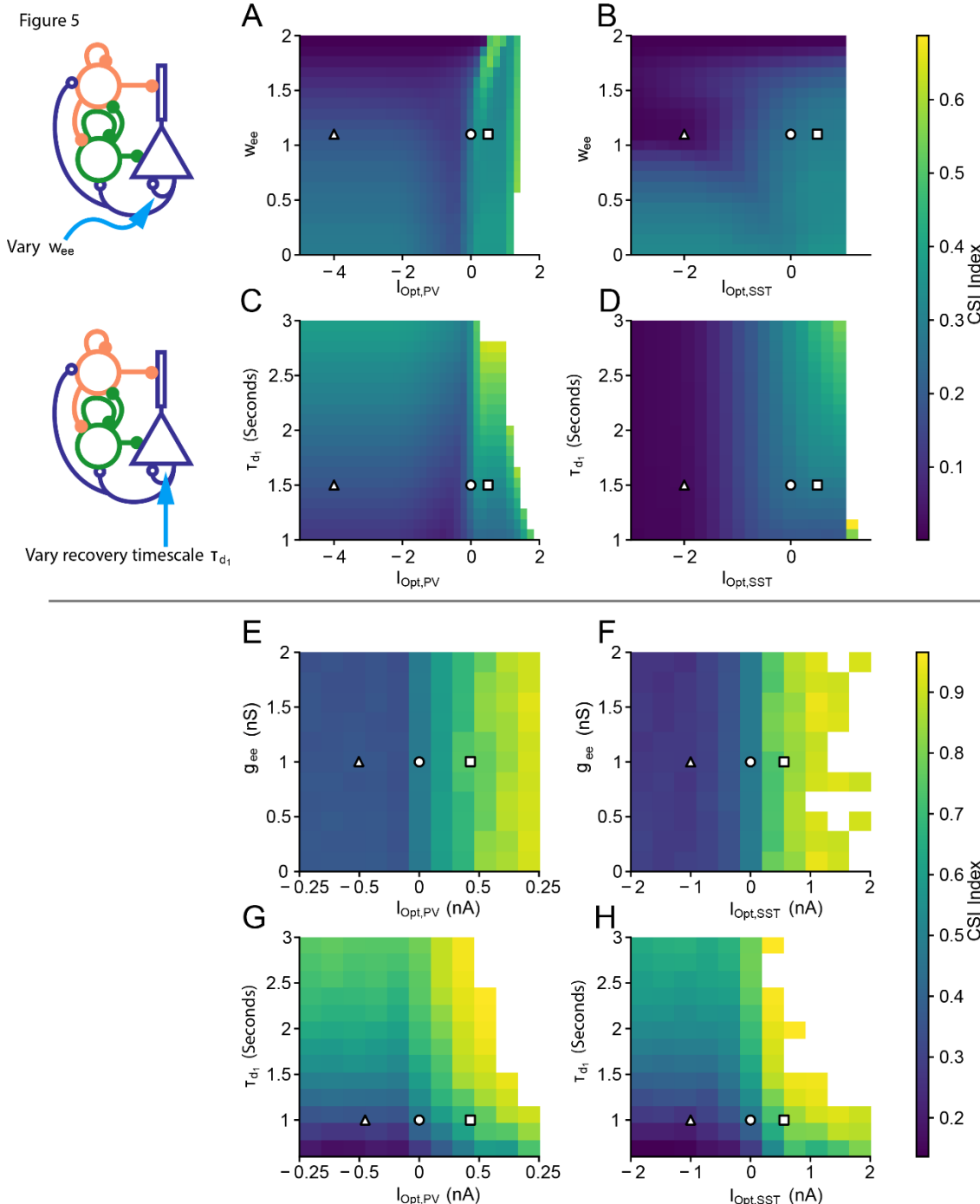


Figure 5. Predicted effects of the key parameters on SSA index (CSI) for the rate model (A-D) and spiking model (E-H). Control parameters are denoted by white circles, PV and SST inactivation parameters are denoted by white triangles, and PV and SST activation parameters are denoted by white squares. The control parameter values, $I_{Opt,PV} = I_{Opt,SST} = 0$, are denoted by white circles. A, E: PV optogenetic parameter vs recurrent excitation (w_{ee}). B, F: SST optogenetic parameter vs recurrent excitation. C, G: PV optogenetic parameter vs thalamic depression time constant τ_{d_1} . D, H: SST optogenetic parameter vs thalamic depression time constant. White regions in all subfigures denote areas where the firing rate (FR) of the standard tone is too low ($FR < 0.1$), or where the excitatory response saturates, making CSI measurements impossible.

This analysis reveals robustness in parameter ranges for given optogenetic modulation strengths. The CSI in the control case (white circle) changed little when the parameters w_{ee} and τ_{d_1} were varied for a given optogenetic strength. In other words, for a fixed value on the x-axis, changing positions in the vertical direction

on each plot did not change the CSI index significantly for a nontrivial range of parameter values. This observation suggests that the cortical model can operate in a broad parameter regime, and precise parameter values may not be important for normal function. In extreme cases, decreasing recurrent excitation removed the decrease in CSI following SST inactivation (Figure 5B), suggesting that sufficient recurrent excitation is an important factor in generating responses in the SSA paradigm. Second, while increasing optogenetic inhibition had little effect on the CSI, increasing optogenetic activation showed an increase in CSI in all cases (CSI= 0.35 for PV activation and CSI= 0.31 for SST activation). Therefore, we predicted that optogenetic activation of PVs and SSTs will generally improve context-dependent cortical responses.

Like the rate model, the spiking model exhibited little sensitivity to changes in w_{ee} and τ_{d1} . However, the spiking model showed almost no dependence on recurrent excitation w_{ee} in the case of SST inactivation (Figure 5E). This effect is likely due to the differences in connectivity between the rate and spiking models. In the rate model, lateral connections depend entirely on excitatory activity, thus SSA results in the rate model are more sensitive to changes in recurrent excitation. In the spiking model, recurrent excitation plays a less important role because the lateral connection probabilities are low ($p = 0.1$), whereas the connection probabilities within units are high ($p = 0.6$).

The three-unit model was developed to reproduce the compensating mechanisms of the single-unit model: PV suppression results in constant disinhibition for repeated tones, and SST suppression results in a compensating effect from PVs before adaptation that weakens as adaptation strengthens. These differential roles explain experimental data in the SSA paradigm to a remarkable degree. We then asked whether this simple mechanism is sufficient to reproduce additional optogenetic experiments. For the remainder of the paper, we use the three-unit rate model with no parameter modifications except for the changes in the inhibition modes and the auditory inputs that depend on the experimental paradigm (Table 2, Figure 1C, and Equation 10).

Differential effects of inhibitory neuron manipulation on cortical forward suppression

Context dependence of auditory responses has been revealed on many time scales. In a well-studied phenomenon termed “forward suppression”, the responses of AC neurons to a tone are suppressed if the tone is preceded by another tone, but the level of suppression depends on the frequency difference between the two tones (Figure 6). In the experiment, the first tone, called the masker, varies in frequency between trials, while the second tone, called the probe, remains fixed at the preferred frequency of the neuron. This phenomenon was explained by feedforward depression, but the inhibitory neurons were recently shown to also control forward suppression (Phillips, Schreiner, and Hasenstaub 2017). PV inactivation (orange) concurrent with the auditory stimulus resulted in a selective increase in forward suppression at the preferred frequency relative to the control case (blue), whereas SST inactivation (green) reduced forward suppression at the preferred frequency relative to the control case (blue) (Figure 4B second row) (Phillips, Schreiner, and Hasenstaub 2017).

Figure 6 A Forward Suppression Stimulus Paradigm

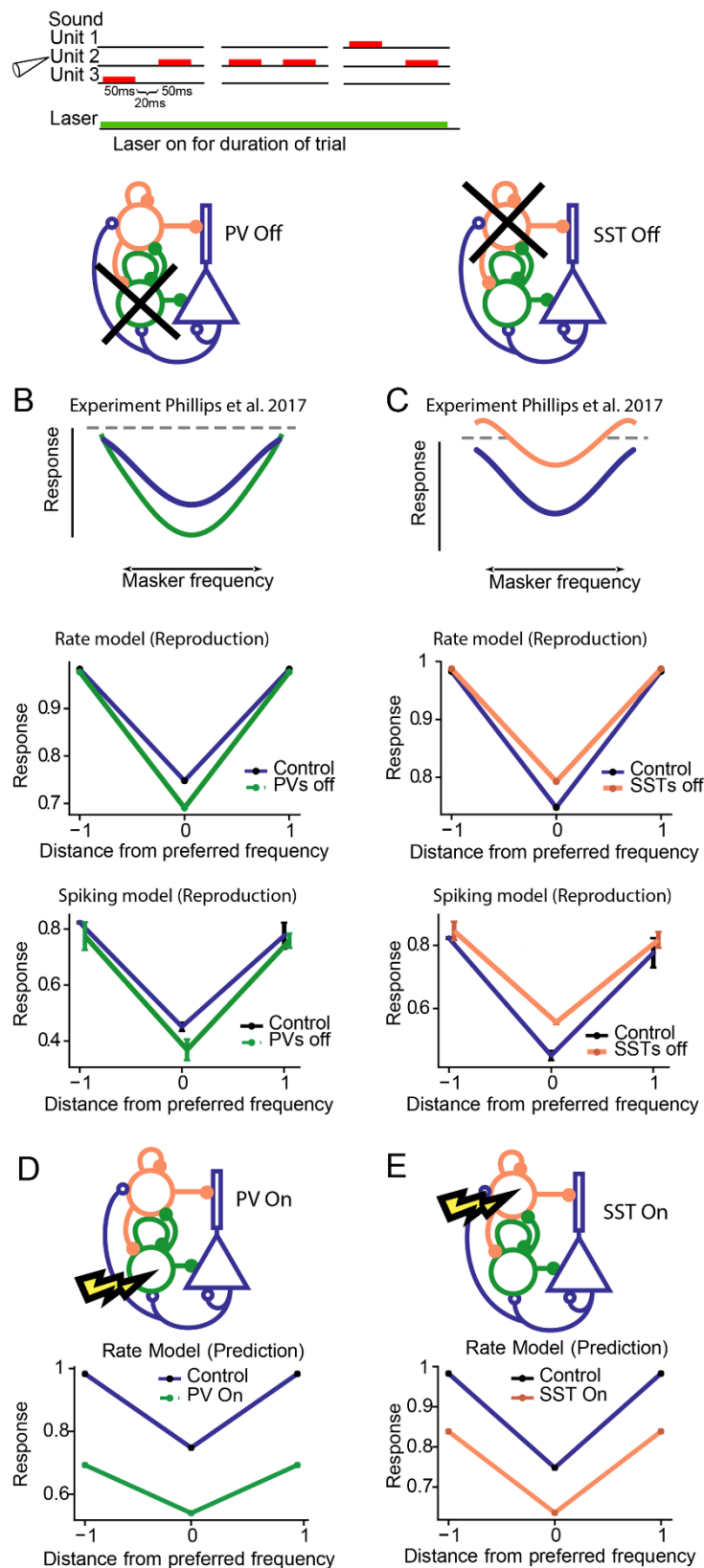


Figure 6. Forward suppression in the rate and spiking model. A. The stimulus consisted of pairs of tones activating either neighboring or the same iso-frequency units. The laser was presented continuously throughout stimulation trials. B,C,D, E: Responses of excitatory neurons to the probe tone as a function of the frequency of the masker. Blue: control. B, C: Top row: Schematic of results from Phillips et al., 2017. Middle row: Results from the rate model. Bottom row: Results from the spiking model. B: Results of PV suppression (green). C: Results of SST suppression (orange). D,E: Rate model prediction for forward suppression during

PV and SST activation. PV and SST activation resulted in enhanced forward suppression. D: Results from PV activation. E. Results from SST activation.

We used the same parameters for connectivity within the circuit as with SSA to reproduce the experimental findings, with only slight changes to the input strength ($q = 1.3$). The stimuli used in the forward suppression paradigm place the baseline state in the strong inhibitory regime (Figure 1C). Both the rate (Figure 6A middle, 6B middle) and spiking models (Figure 6B bottom) yielded the existence of experimentally observed differential effects for PV (Figure 6A) and SST inactivation (Figure 6B): PV inactivation drove a selective decrease in responses whereas SST inactivation drove a suppression of excitatory neuronal responses. We do not expect the scales between the rate and spiking models to match precisely, and we only sought to match the existence of the differential phenomena. Functions $I_{\text{Opt,PV}}(t)$ and $I_{\text{Opt,SST}}(t)$ were turned on for the duration of the simulation to match the experimental protocol. In the rate model, the inhibition was dimensionless and free parameters chosen to be $I_{\text{Opt,PV}}(t) = -4$ during PV inhibition, $I_{\text{Opt,PV}}(t) = 0.5$ during PV activation, $I_{\text{Opt,SST}}(t) = -2$ during SST inhibition, and $I_{\text{Opt,SST}}(t) = 1.2$ during SST activation. In the spiking model, we chose $I_{\text{Opt,PV}}(t) = -0.2\text{nA}$ during PV suppression, $I_{\text{Opt,PV}}(t) =$ during PV activation, $I_{\text{Opt,SST}}(t) = -1\text{nA}$ during SST inhibition, and $I_{\text{Opt,SST}}(t) = -1\text{nA}$ during SST activation.

At first glance, this result seems paradoxical given that PV suppression generally results in excitatory disinhibition as shown in the adaptation and SSA results (Figures 3,4), but the underlying mechanism is straightforward to understand. Following PV suppression, excitatory activity is indeed disinhibited, but the firing rate function of the excitatory population saturates. This behavior means that higher activity neurons are generally disinhibited less strongly than lower activity neurons. Thus, upon receiving the second tone, the input received by the excitatory population is weaker due to thalamic depression, but the disinhibition is greater relative to the disinhibition in the first tone. This phenomenon also requires a special property in the thalamic depression variable $g(t)$, namely that it cannot depress too quickly during the first time or else the excitation in the second tone will be too weak, and although it recovers slowly, it recovers enough such that the input from the second tone is still somewhat strong. In the case of SST suppression, PVs compensate for the loss of inhibition in the first tone, but lose the ability for compensation in the second tone, so Exc are able to respond more strongly relative to the control case. Thus, forward suppression is weakened.

Next, we tested the effects of activating PVs or SSTs (as could be done with ChR2 experimentally) on model responses. The model predicted that both PV and SST activation will result in an increase of forward suppression across preferred and sideband frequencies (Figure 6D,E).

Differential adaptation to repeated tones along the frequency response function

Neurons in A1 adapt to repeated tones (Natan et al. 2015). This adaptation is proportional to the strength of their tone-evoked responses: it is stronger in the center of the frequency response function, and weaker for the sidebands (Natan, Rao, and Geffen 2017). A recent study found that PVs and SSTs exert a differential effect on this form of adaptation: Suppressing PVs drives disinhibition selective to the sidebands in the adapted state, whereas suppressing SSTs drives disinhibition both in the center and at the sidebands

599 of the frequency response function of excitatory neurons (Natan, Rao, and Geffen 2017). To understand how
600 inhibitory neurons affect adaptation across different frequency-tuned inputs, we presented a sequence of 8
601 tones at each frequency to generate adapting tuning curves (Figure 7A), and repeated this stimulus with PV
602 and SST suppression for the model circuit. We found that this auditory paradigm resulted in a below-threshold
603 integration of \bar{F} , so the system switched to a state of strong baseline inhibition (and importantly, the model
604 did not respond in precisely the same way as in SSA and forward suppression).

605 Our model reproduced the differential experimental effects of PV and SST suppression (Figure 7). In
606 the rate model before adaptation, PV and SST inactivation resulted in sideband disinhibition with little to no
607 disinhibition at the preferred frequency (Figure 7D,J left). After adaptation, PV inactivation resulted in sideband
608 disinhibition and no preferred frequency disinhibition (Figure 7D right), whereas SST inactivation resulted in
609 disinhibition across all sideband and preferred frequencies (Figure 7J right). The spiking model closely
610 mirrored these results (Figure 7F,L). The ratio of excitatory responses between light off and light on trials
611 summarize the degree of sideband and preferred frequency disinhibition (Figure 7E,G,K,M).

612 The mechanisms behind these results involve synaptic facilitation and depression and the
613 compensating mechanism discussed in the earlier sections for SSA and adaptation. In the case of PV
614 suppression, SSTs were the only interneurons capable of contributing to Exc inhibition, so only Exc-SST
615 interactions drove the observed effects. In particular, lateral SST to Exc synapses suppressed the center unit
616 over each tone, and facilitation allowed this suppression to persist throughout adaptation. Note that this
617 preferred-frequency effect was not observed in SSA because we never directly stimulated the center unit.
618 Next, in the case of SST suppression, the increasing disinhibition with adaptation at the preferred frequency
619 was a consequence of the same compensating mechanism as in SSA. The tuning curves shown in Figure
620 7B,H are representative cartoons of the changes in the tuning curves observed quantitatively (Natan, Rao,
621 and Geffen 2017), but we aimed to reproduce the light-on and light-off curves (Figure C,H) using our model.
622 The comparison of the model tuning curves to the cartoons from the experiments appear to be significantly
623 different at the sidebands, but in fact our model outputs closely match the desired lines in Figure 7C,I.

624 Turning to predictions, we found that before adaptation, PV activation resulted in a slight decrease at
625 the preferred frequency, whereas SST inactivation reduced overall firing rates across all frequencies (Figure
626 7N,O). After adaptation, PV and SST activation resulted in a subtractive effect. General optogenetic activation
627 and inactivation of PVs and SSTs modulated tuning-curves in combinations of additive, subtractive,
628 multiplicative, and divisive effects (Phillips and Hasenstaub 2016). Our model reproduced one of the key
629 results in these studies as well, since PV and SST inactivation were found to have additive and divisive effects
630 on the frequency response functions of excitatory neurons (Figure E,G,K,M,O,Q).

631 Similar to SSA, the functions $I_{\text{Opt,PV}}(t)$ and $I_{\text{Opt,SST}}(t)$ were turned on 100ms before tone onset and
632 turned off 100ms after tone onset. In the rate model, the inhibition was dimensionless and free parameters
633 chosen to be $I_{\text{Opt,PV}}(t) = -0.5$ during PV inhibition, $I_{\text{Opt,PV}}(t) = 1.2$ during PV activation, $I_{\text{Opt,SST}}(t) = -1$
634 during SST inhibition, and $I_{\text{Opt,SST}}(t) = 0.1$ during SST activation. In the spiking model, we chose $I_{\text{Opt,PV}}(t) =$
635 -1nA during PV inhibition and $I_{\text{Opt,SST}}(t) = -1\text{nA}$ during SST inhibition.

Figure 7

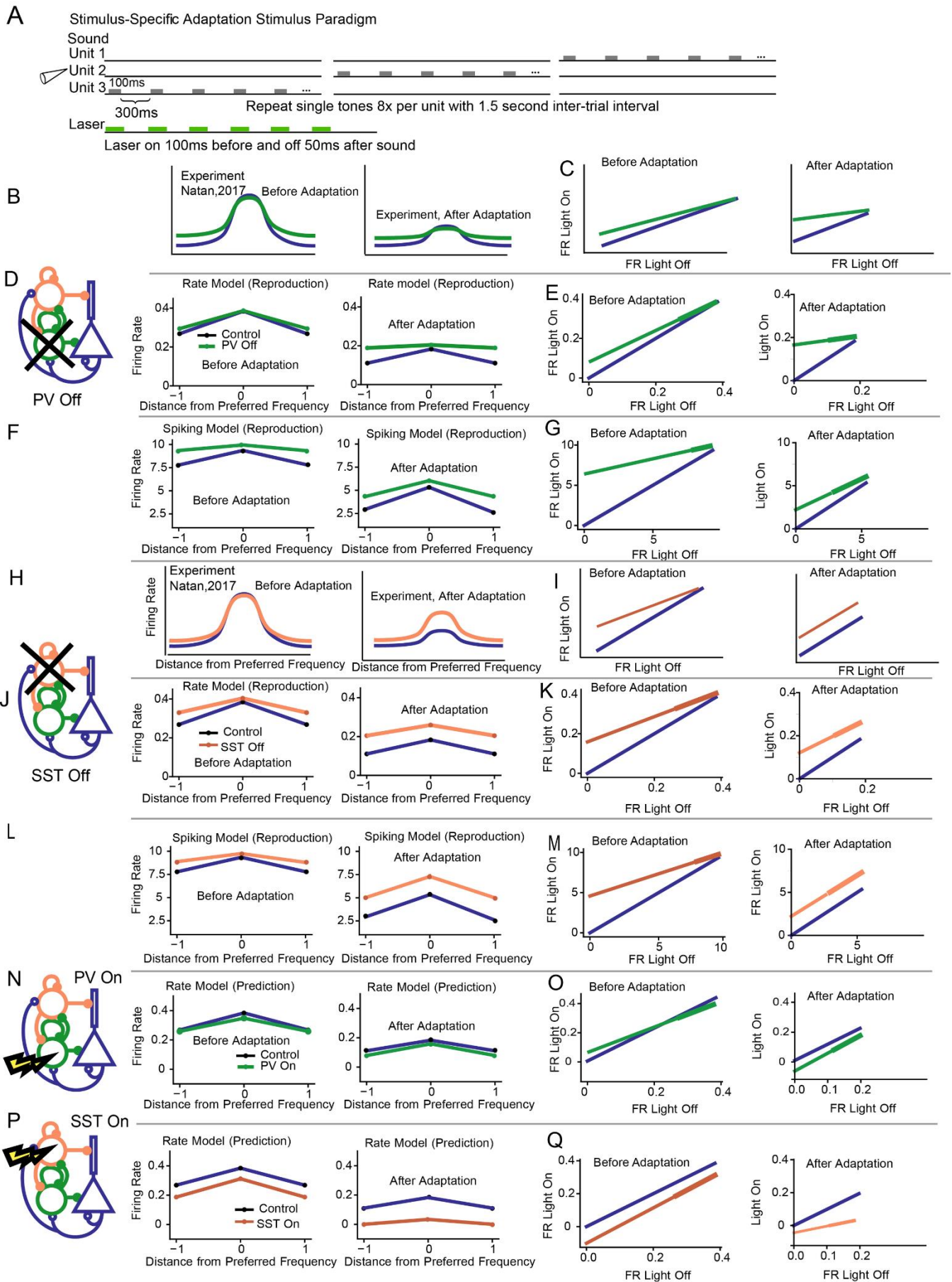


Figure 7. Adaptation to repeated tones along the frequency response function. A. The stimulus consisted of a sequence of repeated tones, presented to each iso-frequency unit. On stimulation trials, the laser overlapped with the sound stimulus. B, D, F, H, J, L, N, P: The responses of excitatory units to the first (left) and last tone (right) as a function of the distance in frequency between the unit and the stimulus without (blue) and with (green: PV suppression; orange: SST suppression) stimulation. C, E, G, I, K, M, O, Q: The response of excitatory neurons to tone 1 (left) and tone 8 (right) on light on and light off trials. The control lines have unit slope because light on and light off in an experimental condition yields no changes to the firing rate. Thicker lines in Figure 7E,G,K,M represent the peak excitatory responses from the first and last simulations taken directly from the simulations, whereas thinner lines are linear extrapolations to assist the visual comparison to the control line (blue). B,C. Experimental results, PV suppression. D,E: Rate Model, PV suppression. F,G. Spiking model, PV suppression. H,I. Experimental results, SST suppression. J,K. Rate model, SST suppression. L,M. Spiking model, SST suppression. N,O. Rate model, PV activation. P,Q. Rate model, SST activation.

PVs Enhance Feedforward Functional Connectivity

Cortical neurons in AC receive inputs from the thalamic auditory nuclei. As the result, neuronal responses in the cortex are correlated with neuronal firing in the thalamus. These interactions can be captured using an Ising model to measure the connection from the thalamus to the cortex. When PVs were activated, the functional coupling between cortical and thalamic responses (Hamilton et al. 2013) became stronger. The specific mechanism underlying this change is unknown.

Using the three-unit model, we identified a candidate mechanism for the enhanced thalamo-cortical correlation following PV activation. We assumed that the feedforward functional connection from the thalamus to the cortex is the same as the anatomical connection, so thalamic inputs directly modulated cortical responses in our model. Following an increase in inhibition, cortical responses became sharper, thus aligning more closely with thalamic inputs and improving feedforward functional connectivity (Figure 8).

Figure 8

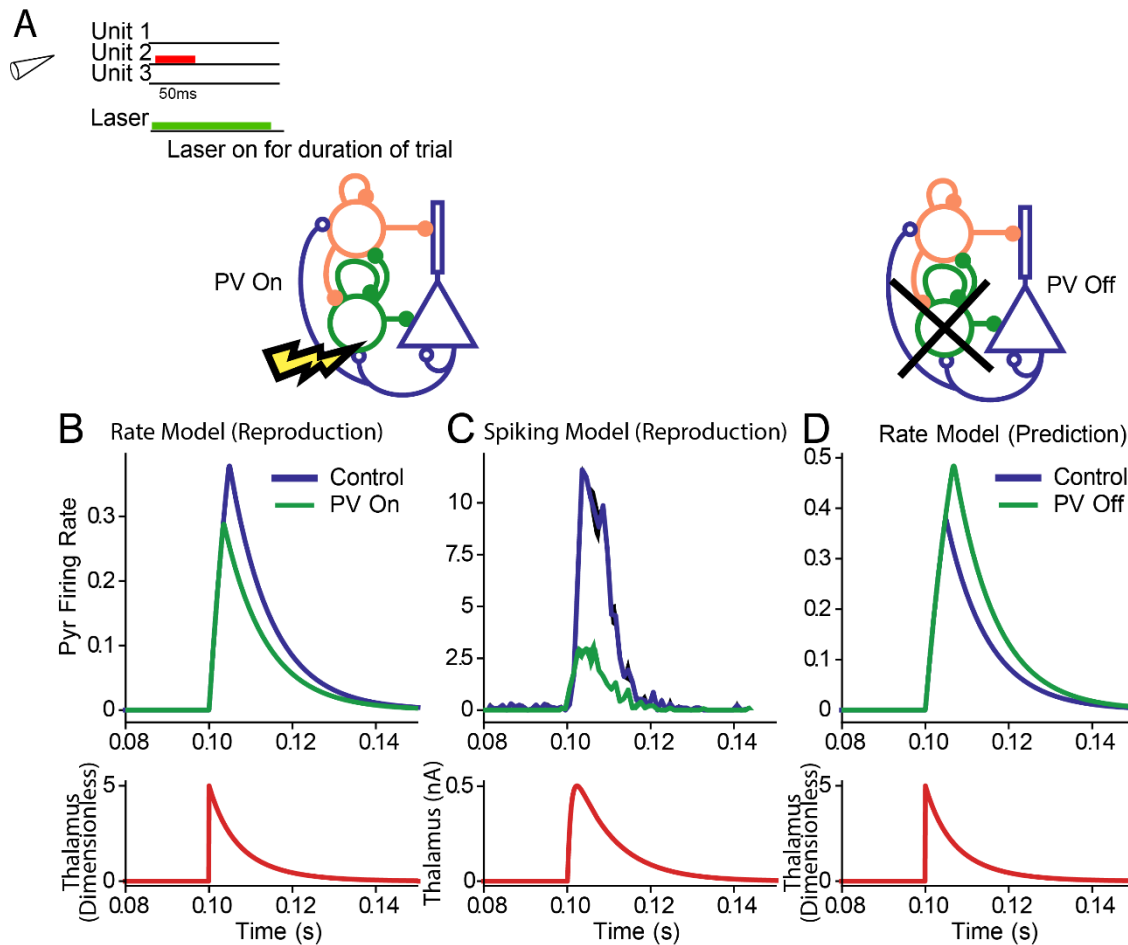


Figure 8. Activation of PVs enhanced feedforward connectivity in the model. A. Stimulus was a single tone accompanied by a laser on stimulation trials presented at 0.1 s. B. Top: Cortical excitatory population responses to tones without (blue) and with PV stimulation (green) in the rate model. Bottom: Thalamic input (red). C. Top: Cortical excitatory population responses to tones without (blue) and with PV stimulation (green) in the spiking model. D: Rate model prediction for effects of PV inactivation.

PV activation (green) in the rate model resulted in an increase in the Pearson correlation between the control (blue) and thalamic inputs (red), from 0.77 and 0.83 (Figure 8B). Thus, whereas inhibitory activation decreased the overall firing rate, the response became more synchronized to the thalamic inputs, resulting in an increase in feedforward functional connectivity. In the spiking model, PV activation resulted in a delayed response of excitatory activity, but we were interested in tested whether PV-activated Exc response profile resembled the thalamic activity more than the control Exc response. To make this comparison, we shifted the PV trace so that the onset of PV-activated Exc activity (green) coincided with the onset of the control curve (blue) (Figure 8C. An equivalent approach would be to measure the peak value of the cross-correlation between excitatory and thalamic activity without shifting the data in time). We observed an increase in the Pearson correlation from 0.87 in the control Exc activity to 0.95 in the PV-activated Exc activity, thus demonstrating a sharpening of excitatory responses, and an increase in feedforward functional connectivity. As in the previous paradigms, we only aimed to show the existence of changes without requiring the magnitude of change to match the experimental data.

These results provide for a simple plausible mechanism for enhanced feedforward functional connectivity: as inhibition reduces the overall cortical inputs, cortical responses better synchronize to thalamic inputs, resulting in stronger correlated activity. We remark that the classic approach to the problem of establishing proper functional connectivity relies on sophisticated models such as Ising models (Ganmor, Segev, and Schneidman 2011), expectation-maximization (Turaga et al. 2013), and subspace identification (Nonnenmacher, Turaga, and Macke 2017), because direct calculations of correlations can lead to false positives when the anatomical connections are not known (Hamilton et al. 2013). Our model has explicit anatomical connections, which eliminates the problem of false-positive correlations. Thus, in this case, the use of the correlation serves as a reliable proxy for the Ising model.

Balanced Networks

Multiple studies postulated that excitatory and inhibitory currents are matched in cortical circuits, contributing to stability of circuit function (Denève and Machens 2016; Zhou and Yu 2018). Therefore, we tested whether our model operated as a balanced network. To make this measurement, we took ratios of inhibitory and excitatory currents during suprathreshold activity, because our model ignores virtually all subthreshold activity. Interestingly, with no tuning, our models showed evidence of operating as a balanced network: as we varied the input strength to the rate and spiking models (Figure 9C,F), the ratio of excitatory to inhibitory inputs to the excitatory population (Figure 9B,E) remained constant. The rate model had an excitatory/inhibitory ratio of 0.37 (Figure 9B), and the spiking model had an excitatory/inhibitory ratio of 2.5 (Figure 9E). These results suggest that excitatory-inhibitory balance a robust, emergent feature of cortical networks. The large differences in scales are to be expected when comparing a dimensional and dimensionless model.

Figure 9

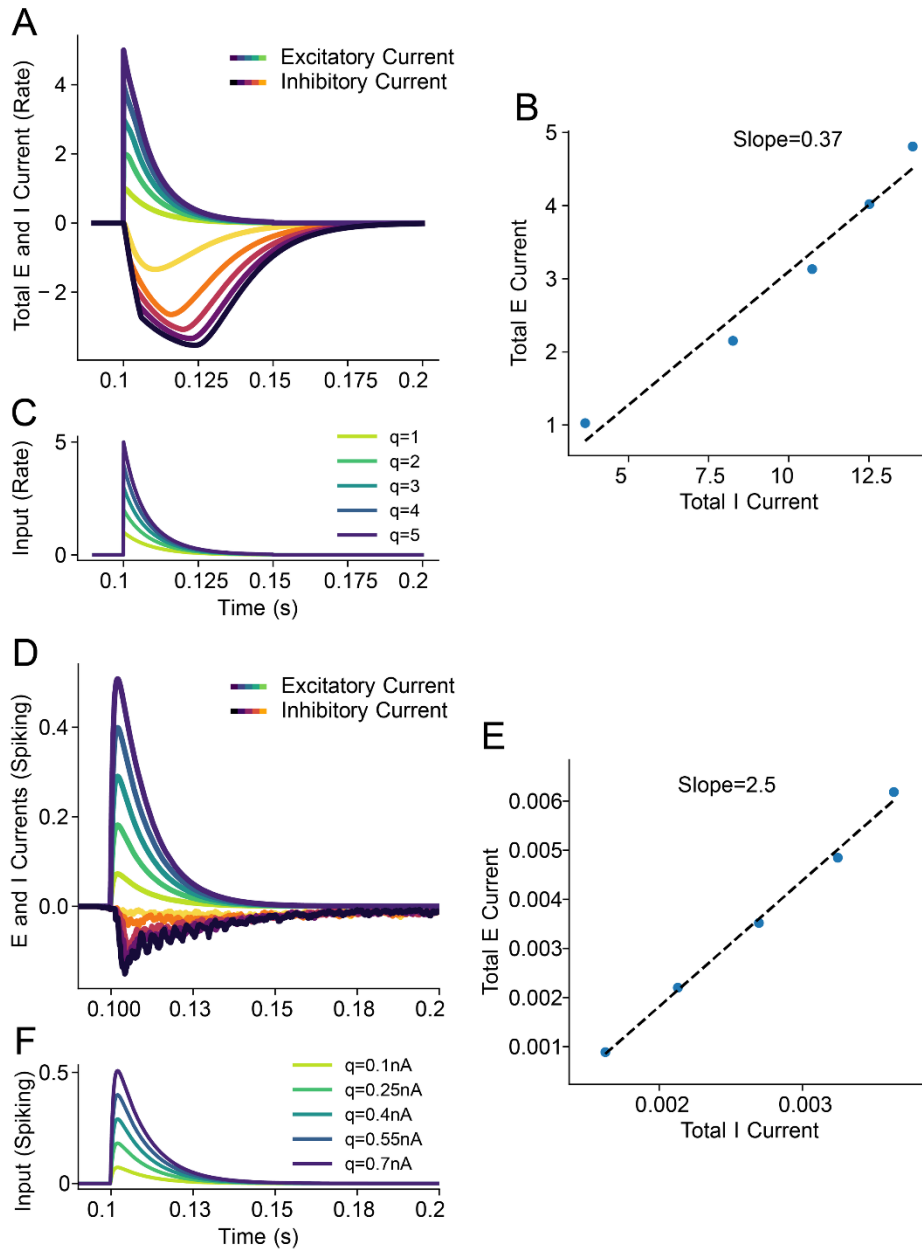


Figure 9. Excitatory-inhibitory balance in the rate and spiking models A. Plot of incoming excitatory and inhibitory currents into the Exc population as a function of different input strengths (C). Darker currents correspond to stronger inputs B. A best-fit line (dashed) accurately captures the ratio of excitatory and inhibitory responses, implying excitatory-inhibitory balance. Equivalent results for the spiking model are shown in panels D, E, and F.

Discussion

A wealth of recent studies provide evidence for distinct function of different types of cortical inhibitory neurons in temporal processing of auditory information. The studies demonstrate that different types of inhibitory neurons, SSTs and PVs, play a differential role in auditory processing, controlling adaptation at different time scales and contexts, and changes to feedforward functional connectivity. Our goal was to integrate the results of these studies to understand whether the observed effects were due to a small set of mechanisms.

719 We built an idealized rate and spiking model that reproduced multiple key results from studies that
720 tested the function of specific inhibitory opsins in specific cells in the auditory cortex. In addition to including
721 different baseline states that modulate the strength of PV-to-Exc and SST-to-Exc synapses, the key
722 mechanisms underlying our models included the fast temporal activation of PVs, the delayed, broad temporal
723 activation of SSTs, the ability for PVs to compensate for weakened SST activity, and dynamic synapses
724 including SST-to-Exc facilitation. These interactions accounted for the differential modulation of cortical
725 responses by interneuron subtypes and suggests that a simplified set of mechanisms can support
726 experimental results.

727 To reproduce the differential function of SSTs and PVs in stimulus-specific adaptation, we built a
728 model loosely based on multiple existing models for SSA and multiple configurations of spiking neuron
729 populations, consisting of inhibitory and excitatory neurons. Previously, a two-layer rate model with synaptic
730 depression was proposed to establish the relationship between the cortical response and the parameters in
731 SSA experiments, such as stimulus frequency differences, probability of deviation, and tone presentation rate.
732 Yarden et al. (2017) successfully used a multi-unit rate model arranged in a coarse tonotopy consisting of
733 inhibitory and excitatory populations to reproduce general deviance detection, but model has not yet been
734 adapted to explain differential interneuron modulation. Another existing model of SSA including differential
735 inhibitory modulation demonstrating similar differential inhibitory effects as in our SSA result (Figure 4), but
736 did not include a tonotopy (Natan et al. 2015). These models only included one type of inhibitory neuron type
737 or did not include tonotopy, and therefore could not account for the observed differential effects of suppression
738 of SSTs and PVs on SSA across multiple frequencies. In the present study, we developed a simple rate and
739 spiking model that accounted for multiple inhibitory cell types and which faithfully reproduced the differential
740 effects of SST and PV inactivation in SSA (Figure 4). In addition, a parameter sweep revealed that both the
741 rate and spiking models were robust to large changes in key parameters commonly explored in the literature,
742 suggesting that SSA is a robust phenomenon (Yarden and Nelken 2017).

743 Existing models that reproduce the enhanced forward suppression from PV inactivation and the
744 reduced forward suppression from SST inactivation (Figure 6) include multiple layers that require both
745 depression and facilitation (Phillips, Schreiner, and Hasenstaub 2017), or rely on depressing recurrent
746 excitation and do not distinguish between inhibitory subtypes (Loebel, Nelken, and Tsodyks 2007). We
747 incorporated depression and facilitation in the model synapses and reproduced the former results with only a
748 single layer, suggesting a surprisingly simple mechanism supporting forward suppression. Furthermore, the
749 models in the present study reproduced tuning-curve adaptation effects previously observed experimentally
750 but not computationally: SSTs exhibited strong preferred-frequency disinhibition following adaptation, while
751 PV disinhibition is independent of the degree of adaptation (Figure 7) (Natan, Rao, and Geffen 2017). These
752 results suggest that the underlying mechanism(s) of the model, namely the PV/SST compensation effect,
753 combined with the facilitating SST-to-Exc synapse, may serve as a general mechanism of adaptation. In
754 addition, our models reproduced changes in feedforward functional connectivity (Figure 8). By increasing PV
755 activity in the models, excitatory activity decreased but became more time-locked to thalamic inputs. This

756 effect agreed with observations in the cortex, where PV activation resulted in enhanced feedforward functional
757 connectivity (Hamilton et al. 2013). The effects of inhibition on sharpening cortical responses have been well-
758 established, thus our models serve as plausible mechanisms for this change (Wehr and Zador 2003; Cardin
759 et al. 2009; Sohal et al. 2009). Finally, our models were shown to operate as a balanced network, where
760 inhibitory and excitatory currents entering neural populations were shown to maintain a consistent ratio across
761 input strengths, suggesting a generality to the theory of balanced networks.

762 One drawback of the model is that it does not feature population spikes, which explain many
763 fundamental cortical responses in AC (Wehr and Zador 2003). In future work, we will seek to reconcile the
764 differences between our models and the population spike model of SSA (Loebel, Nelken, and Tsodyks 2007).
765 Establishing the importance of depression and facilitation in different synapses and extending our model to
766 include population spikes warrants further study.

767 Although we do not explore simultaneous auditory stimuli in this study, it is worth mentioning the
768 response properties of the network due to recent interest in supralinear network models (Rubin et al. 2016).
769 Throughout this paper, neurons operate in a linear manner when above threshold: neurons add inputs linearly,
770 until the maximum rate is reached in the rate models, or until the refractory period saturates spiking rates in
771 the spiking model. The models do not use sub-threshold responses to modulate population activity.

772 Multiple studies from different laboratories revealed the differential effect of distinct inhibitory neurons
773 in auditory processing. We show that a minimalistic model, built on simple mechanisms, is capable of
774 reproducing disparate results in the literature. As inhibitory neurons form similar circuits throughout the
775 mammalian cortex, this model can be readily adapted to test their function and generate predictions (with
776 adjustments for local changes in connectivity) in different sensory modalities.

777

778

779 **Acknowledgments**

780 This work was supported by National Institutes of Health (NIH R01DC014479, NIH R01DC015527), Human
781 Frontier in Science Foundation Young Investigator Award and the Pennsylvania Lions Club Hearing Research
782 Fellowship to MGN. The authors thank the members of the Geffen laboratory and the Hearing Research
783 Center at the University of Pennsylvania, Jason Kim and Lia Papadopoulos, for insightful comments on the
784 manuscript and discussions.

785

786

787 Bibliography

- 788 Abbott, L F, J A Varela, Kamal Sen, and S B Nelson. 1997. "Synaptic Depression and Cortical Gain
789 Control." *Science* 275 (5297): 221–24. <https://doi.org/10.1126/science.275.5297.221>.
- 790 Behrens, Wolfger von der, Peter B  uerle, Manfred K  ssl, and Bernhard H Gaese. 2009. "Correlating
791 Stimulus-Specific Adaptation of Cortical Neurons and Local Field Potentials in the Awake Rat."
792 *Journal of Neuroscience* 29 (44): 13837–49.
- 793 Beierlein, Michael, Jay R Gibson, and Barry W Connors. 2003. "Two Dynamically Distinct Inhibitory
794 Networks in Layer 4 of the Neocortex." *Journal of Neurophysiology* 90 (5): 2987–3000.
- 795 Bell, Brittany A, Mimi L. Phan, and David S. Vicario. 2015. "Neural Responses in Songbird Forebrain
796 Reflect Learning Rates, Acquired Salience, and Stimulus Novelty after Auditory Discrimination
797 Training." *Journal of Neurophysiology*. <https://doi.org/10.1152/jn.00611.2014>.
- 798 Brette, Romain, and Wulfram Gerstner. 2005. "Adaptive Exponential Integrate-and-Fire Model as an
799 Effective Description of Neuronal Activity." *Journal of Neurophysiology* 94 (5): 3637–42.
- 800 Brosch, Michael, and Christoph E Schreiner. 1997. "Time Course of Forward Masking Tuning Curves in
801 Cat Primary Auditory Cortex." *Journal of Neurophysiology* 77 (2): 923–43.
- 802 Cardin, Jessica A, Marie Carl  n, Konstantinos Meletis, Ulf Knoblich, Feng Zhang, Karl Deisseroth, Li-
803 Huei Tsai, and Christopher I Moore. 2009. "Driving Fast-Spiking Cells Induces Gamma Rhythm and
804 Controls Sensory Responses." *Nature* 459 (7247): 663.
- 805 Den  ve, Sophie, and Christian K. Machens. 2016. "Efficient Codes and Balanced Networks." *Nature*
806 *Neuroscience*. <https://doi.org/10.1038/nn.4243>.
- 807 Ermentrout, Bard. 2002. *Simulating, Analyzing, and Animating Dynamical Systems: A Guide to XPPAUT*
808 *for Researchers and Students*. Vol. 14. SIAM.
- 809 Fino, Elodie, Adam M Packer, and Rafael Yuste. 2013. "The Logic of Inhibitory Connectivity in the
810 Neocortex." *The Neuroscientist* 19 (3): 228–37.
- 811 Ganmor, Elad, Ronen Segev, and Elad Schneidman. 2011. "The Architecture of Functional Interaction
812 Networks in the Retina." *Journal of Neuroscience* 31 (8): 3044–54.
- 813 Hamilton, Liberty S, Jascha Sohl-Dickstein, Alexander G Huth, Vanessa M Carels, Karl Deisseroth, and
814 Shaowen Bao. 2013. "Optogenetic Activation of an Inhibitory Network Enhances Feedforward
815 Functional Connectivity in Auditory Cortex." *Neuron* 80 (4): 1066–76.
- 816 Isaacson, Jeffry S, and Massimo Scanziani. 2011. "How Inhibition Shapes Cortical Activity." *Neuron* 72
817 (2): 231–43.
- 818 Lee, Charles C, and S Murray Sherman. 2008. "Synaptic Properties of Thalamic and Intracortical Inputs
819 to Layer 4 of the First-and Higher-Order Cortical Areas in the Auditory and Somatosensory
820 Systems." *Journal of Neurophysiology* 100 (1): 317–26.
- 821 Levy, Robert B, and Alex D Reyes. 2012. "Spatial Profile of Excitatory and Inhibitory Synaptic
822 Connectivity in Mouse Primary Auditory Cortex." *Journal of Neuroscience* 32 (16): 5609–19.
- 823 Litwin-Kumar, Ashok, Robert Rosenbaum, and Brent Doiron. 2016. "Inhibitory Stabilization and Visual
824 Coding in Cortical Circuits with Multiple Interneuron Subtypes." *Journal of Neurophysiology* 115
825 (3): 1399–1409.

826 Loebel, Alex, Israel Nelken, and Misha Tsodyks. 2007. "Processing of Sounds by Population Spikes in a
827 Model of Primary Auditory Cortex." *Frontiers in Neuroscience* 1: 15.

828 Lu, Kai, and David S. Vicario. 2017. "Familiar but Unexpected: Effects of Sound Context Statistics on
829 Auditory Responses in the Songbird Forebrain." *Journal of Neuroscience*.
830 <https://doi.org/10.1523/JNEUROSCI.5722-12.2017>.

831 Markram, Henry, Yun Wang, and Misha Tsodyks. 1998. "Differential Signaling via the Same Axon of
832 Neocortical Pyramidal Neurons." *Proceedings of the National Academy of Sciences* 95 (9): 5323–28.

833 Mill, Robert, Martin Coath, Thomas Wennekers, and Susan L Denham. 2011. "A Neurocomputational
834 Model of Stimulus-Specific Adaptation to Oddball and Markov Sequences." *PLoS Computational
835 Biology* 7 (8): e1002117.

836 ———. 2012. "Characterising Stimulus-Specific Adaptation Using a Multi-Layer Field Model." *Brain
837 Research* 1434: 178–88.

838 Natan, Ryan G, John J Briguglio, Laetitia Mwilambwe-Tshilobo, Sara I Jones, Mark Aizenberg, Ethan M
839 Goldberg, and Maria Neimark Geffen. 2015. "Complementary Control of Sensory Adaptation by
840 Two Types of Cortical Interneurons." *Elife* 4.

841 Natan, Ryan G, Winnie Rao, and Maria N Geffen. 2017. "Cortical Interneurons Ensure Maintenance of
842 Frequency Tuning Following Adaptation." *BioRxiv*, 172338.

843 Nonnenmacher, Marcel, Srinivas C Turaga, and Jakob H Macke. 2017. "Extracting Low-Dimensional
844 Dynamics from Multiple Large-Scale Neural Population Recordings by Learning to Predict
845 Correlations." In *Advances in Neural Information Processing Systems*, 5702–12.

846 Pfeffer, Carsten K, Mingshan Xue, Miao He, Z Josh Huang, and Massimo Scanziani. 2013. "Inhibition of
847 Inhibition in Visual Cortex: The Logic of Connections between Molecularly Distinct Interneurons."
848 *Nature Neuroscience* 16 (8): 1068.

849 Phillips, Elizabeth A K, and Andrea R Hasenstaub. 2016. "Asymmetric Effects of Activating and
850 Inactivating Cortical Interneurons." *Elife* 5.

851 Phillips, Elizabeth A K, Christoph E Schreiner, and Andrea R Hasenstaub. 2017. "Cortical Interneurons
852 Differentially Regulate the Effects of Acoustic Context." *Cell Reports* 20 (4): 771–78.

853 Rankin, James, Elyse Sussman, and John Rinzel. 2015. "Neuromechanistic Model of Auditory
854 Bistability." *PLoS Computational Biology* 11 (11): e1004555.

855 Rossum, Guido. 1995. "Python Reference Manual." Amsterdam, The Netherlands, The Netherlands:
856 CWI (Centre for Mathematics and Computer Science).

857 Sohal, Vikaas S, Feng Zhang, Ofer Yizhar, and Karl Deisseroth. 2009. "Parvalbumin Neurons and
858 Gamma Rhythms Enhance Cortical Circuit Performance." *Nature* 459 (7247): 698.

859 Stimberg, Marcel, Romain Brette, and Dan F M Goodman. 2019. "Brian 2: An Intuitive and Efficient
860 Neural Simulator." *BioRxiv*. <https://doi.org/10.1101/595710>.

861 Tsodyks, Misha V, William E Skaggs, Terrence J Sejnowski, and Bruce L McNaughton. 1997.
862 "Paradoxical Effects of External Modulation of Inhibitory Interneurons." *Journal of Neuroscience*
863 17 (11): 4382–88.

864 Turaga, Srini, Lars Buesing, Adam M Packer, Henry Dalglish, Noah Pettit, Michael Hausser, and Jakob
865 H Macke. 2013. "Inferring Neural Population Dynamics from Multiple Partial Recordings of the

866 Same Neural Circuit." In *Advances in Neural Information Processing Systems*, 539–47.

867 Ulanovsky, Nachum, Liora Las, and Israel Nelken. 2003. "Processing of Low-Probability Sounds by
868 Cortical Neurons." *Nature Neuroscience* 6 (4): 391.

869 Wehr, Michael, and Anthony M Zador. 2003. "Balanced Inhibition Underlies Tuning and Sharpens
870 Spike Timing in Auditory Cortex." *Nature* 426 (6965): 442.

871 Wehr, Michael, and Anthony M Zador. 2005. "Synaptic Mechanisms of Forward Suppression in Rat
872 Auditory Cortex." *Neuron* 47 (3): 437–45.

873 Wilson, Hugh R, and Jack D Cowan. 1972. "Excitatory and Inhibitory Interactions in Localized
874 Populations of Model Neurons." *Biophysical Journal* 12 (1): 1–24.

875 Womelsdorf, Thilo, Taufik A Valiante, Ned T Sahin, Kai J Miller, and Paul Tiesinga. 2014. "Dynamic
876 Circuit Motifs Underlying Rhythmic Gain Control, Gating and Integration." *Nature Neuroscience* 17
877 (8): 1031.

878 Yarden, Tohar S, and Israel Nelken. 2017. "Stimulus-Specific Adaptation in a Recurrent Network Model
879 of Primary Auditory Cortex." *PLoS Computational Biology* 13 (3): e1005437.

880 Zhou, Shanglin, and Yuguo Yu. 2018. "Synaptic E-I Balance Underlies Efficient Neural Coding." *Frontiers*
881 *in Neuroscience*. <https://doi.org/10.3389/fnins.2018.00046>.

882