Title: **A Unifying Mechanistic Model of Excitatory-Inhibitory Interactions in the Auditory Cortex**

Short Title: A Unifying Model of Auditory Cortex

Authors: Youngmin Park[1] and Maria N. Geffen[1,2*]

[1]Department of Otorhinolaryngology: HNS, University of Pennsylvania, Philadelphia, PA, 19104, USA
[2]Department 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), 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.

Abstract

The mammalian sensory cortex is comprised of multiple types of inhibitory and excitatory neurons, which form sophisticated microcircuits for processing and transmitting sensory information. Despite rapid progress in understanding the function of distinct neuronal population, the parameters of connectivity that are required for the function of these microcircuits remain unknown. Recent studies found that two most common inhibitory interneurons, parvalbumin- (PV) and somatostatin-(SST) positive interneurons control sound-evoked responses, temporal adaptation and network dynamics in the auditory cortex (AC). These studies can inform our understanding of parameters for connectivity of the excitatory-inhibitory cortical circuits. Specifically, we asked whether a common microcircuit can account for the disparate effects found in studies in different groups. We built rate and spiking models of the auditory cortex consisting of excitatory, PV and SST neurons, and searched the space of connectivity parameters to identify the set that can account for the experimental findings from multiple groups. We identified microcircuit parameters that accounted for the differential effects of PVs and SSTs in stimulus-specific adaptation, forward suppression and tuning-curve adaptation, as well as the influence of PVs on functional connectivity in the circuit. The unifying mechanisms of the model included a depressing synapse from PVs to excitatory neurons, and a facilitating synapse from excitatory neurons to SSTs. This approach brought together multiple findings from different laboratories and identified a unified circuit that can be used in future studies of upstream and downstream sensory processing.

Significance Statement

The mammalian auditory cortex is comprised of multiple types of inhibitory and excitatory neurons, which form sophisticated microcircuits for processing and transmitting sensory information. Distinct inhibitory neuron subtypes play distinct functions in auditory processing, but it remains unknown whether these phenomena are due to a unified microcircuit or require multiple circuits. Here, we built minimal rate and spiking models and identified a specific set of synaptic mechanisms and parameters that could best reproduce the broad set of experimental results in the auditory cortex. The simplicity of our model provides an understanding of inhibitory cortical processing at the circuit level, unifying the results from different laboratories, and provides for a novel computational framework for future studies of cortical function.

Introduction

Detecting sudden changes in the acoustic environment and extracting relevant acoustic features from noise are important computations for auditory navigation and scene analysis. The mammalian auditory cortex (AC) is a key region for processing temporally patterned sounds. Neurons in AC exhibit adaptation to repeated toned, which may be selective for an overrepresented stimulus (SSA). They furthermore exhibit forward suppression, in which a preceding stimulus masker tone drives a decrease in responses to the subsequent target tone. How these computations are carried out by cortical circuits has been subject of extensive research.

The AC is comprised of tightly coupled networks of excitatory and inhibitory neurons. Recent studies have identified differential involvement of two distinct major classes of inhibitory neurons, parvalbumin-positive (PVs) and somatostatin-positive (SSTs) neurons, in these temporal paradigms. These neurons differ morphologically and physiologically [1,2], and recent studies found that they play differential functions in auditory processing. Specifically, SSTs, but not PVs were shown to facilitate stimulus-specific adaptation [3]. PVs and SSTs played distinct roles in adaptation to repeated tones along the frequency response function of the target neuron [4]. SSTs and PVs drove bi-directional effects on forward suppression [5]. In addition, PVs were shown to enhance feedforward connectivity in the auditory cortex [6]. These experimental results can be used to constrain the set of circuit models of AC, restricting the parameters for connections between PVs, SSTs and Exc. Here, we tested whether these results can be accounted for the same circuit or whether they are due to different regimes of interactions in the cortex.

In the present study, we introduce a unifying mechanistic framework to understand the reported phenomena. We begin with an idealized, single-unit model of AC capable of reproducing detailed SST and PV modulatory effects on excitatory activity in a simulated paradigm similar to stimulus-dependent adaptation. Taking the lessons learned from this simple model, we introduce a tonotopy with synaptic dynamics and find that the model reproduces the full range of optogenetic results in the literature. While the added dynamics increase the complexity of the model, we use our simple starting point as a template along with data from the literature to strongly constrain the parameter values. The identified set of parameters account for the observed results including the differential role of SSTs and PVs in SSA, forward suppression, tuning-curve adaptation, and the effects of PVs on functional connectivity. We find that the key parameters of the model accounting for the disparate results include the mode of inhibition (high versus low), the strength of the thalamic inputs, and the strength of the optogenetic inactivation and activation. The framework that we develop can be used to build and test hypotheses for similar phenomena in other sensory modalities, and studies of upstream or downstream auditory processing in AC.

Materials and Methods

We constructed two primary model types in this paper. We first built an augmented version of the Wilson-Cowan model, consisting of one or three iso-frequency units of the auditory cortex, and included one excitatory neural population and two inhibitory neural subpopulations. This model served as the template for all other models in this paper. By using the results and parameters from this model, we extended our results to the

substantially more complex three-unit rate model and three-unit spiking models. The increased complexity of the models was mitigated by using the single-unit rate model as a template with data from the literature to constrain the parameter values.

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

Augmented Wilson-Cowan Model

We modeled a single iso-frequency unit as an augmented version of the Wilson-Cowan model [7] by including an additional inhibitory subtype. We drew much of our understanding of adaptation in the auditory cortex using this single iso-frequency unit:

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

where u, p, and s represent the normalized firing rate (scaled from 0 to 1) of the excitatory population, PV inhibitory subpopulation, and SST inhibitory subpopulation, respectively (Figure 1). The parameters $I_{\text{Opt,PV}}$ ($I_{\text{Opt,SST}}$) represent the strength of PV (SST) activation or inactivation, and w_{ij} and τ_i are synaptic weights and time constants, respectively. All time constants are $\tau_u = \tau_p = \tau_s = 10$ms, roughly in agreement with known data [3,8]. The function f is a threshold linear function defined as

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

where the function f roughly approximates a sigmoid that converges to zero for small or negative inputs and 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 [3,9]. We included a threshold as $f(x - u_{th})$, where u_{th} is a positive number. 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, and simplifies our modeling study.

The input function $I(t)$ consists of blocks of inputs with stimulus duration and interval based on the experimental paradigm. We show the stimulus duration and stimulus interval for each paradigm in Table 1 and detail the paradigm in the text and figures where appropriate. When an auditory input arrives, 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 $\tau_q = 10$ms, which roughly agrees with known values [10]. The input $I(t)$ is further modulated by a slow synaptic depression term g satisfying the standard model of synaptic depression

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

where the time constants are $\tau_{d_1} = 1500$ms for replenishment and $\tau_{d_2} = 20$ms for depletion (chosen close to reported values [3,8,11,12]). The synaptic depression variable g begins at a baseline value of g_0 and when $I(t) > 0$, i.e., when an input arrives, $g(t)$ decreases on the timescale determined by τ_{d_2} and modulates the peak amplitude of auditory inputs to A1. $g(t)$ then recovers slowly on the order of seconds.

The single-unit model uses the connectivity pattern based on existing studies on AC [13]. The within-unit connectivity is equivalently represented by the matrix,

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

All synaptic weights w_{ij} in the single-unit rate model are constant [14], with synaptic depression appearing in the feedforward thalamic inputs [15]. The inhibitory synaptic weights were roughly chosen to agree with known connection types and connection strengths [13,16], 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. We briefly discuss the insights gained using this simple model and the details of the transition into the three-unit model in the results.

Figure 1

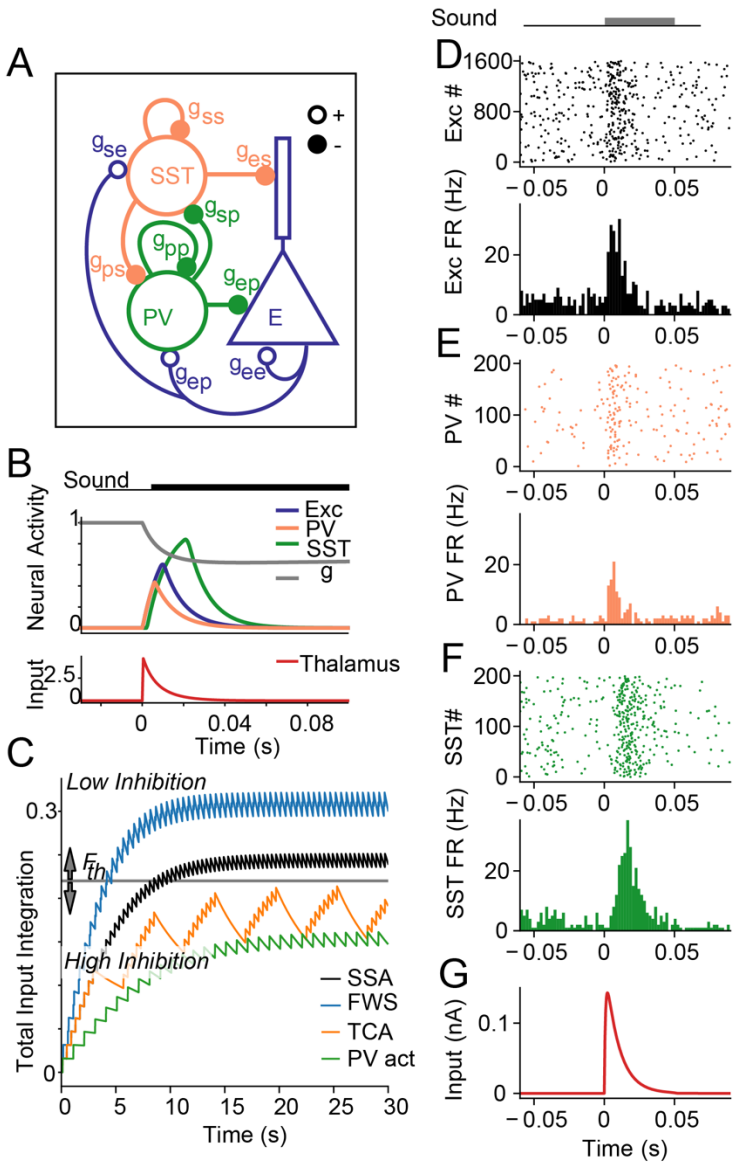

Figure 1.A. Model of PV-SST-Exc circuit. B. Input and response profiles for the single-unit rate and spiking model to a 100 ms long tone. Top: Gray: thalamic depression variable g. Blue: excitatory (Exc) neuron activity. Green: PV. Orange: SST. Bottom: Thalamic input (red). C. Responses to stimulus over 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 (blue), the variable $\overline{F}$ is higher than threshold $\overline{F}_{Th}$, resulting in a set of low inhibition parameters. Paradigms for tuning-curve adaptation (orange) and PV (green) activation 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), SSTs (E), PVs (F), and thalamo-cortical input (G).

Three-unit Rate Model

Using the single-unit rate model as a template, we arranged copies into three units with lateral cortical and thalamic connections (Figure 2A). This arrangement endowed our model with a gross tonotopy, which we used to explore spectrally and temporally complex auditory inputs.

The first, or leftmost unit, satisfies

$$\begin{aligned}
\tau_u u_1{}' &= -u_1 + f(w_{ee}u_1 - (w_{ep} - a(1-D_1))p_1 - w_{es}s_1 - F_1 s_2 + qI_1(t) + w_{ee}^* u_2),\\
\tau_p p_1{}' &= -p_1 + f(w_{pe}u_1 - w_{pp}p_1 - w_{ps}s_1 + I_{\text{Opt,PV}} + qI_1(t) + w_{pe}^* u_2),\\
\tau_s s_1{}' &= -s_1 + f((w_{se} + bF_1)u_1 - w_{sp}p_1 - w_{ss}s_1 + I_{\text{Opt,SST}}),
\end{aligned}$$

where $I_1 = i_1(t) + i_2(t)\alpha$. The second, or center unit, satisfies

$$\begin{aligned}
\tau_u u_2{}' &= -u_2 + f(w_{ee}u_2 - (w_{ep} - a(1-D_2))p_2 - w_{es}s_2 - F_2(s_1+s_3) + qI_2(t) + w_{ee}^*(u_1+u_3)/2),\\
\tau_p p_2{}' &= -p_2 + f(w_{pe}u_2 - w_{pp}p_2 - w_{ps}s_2 + I_{\text{Opt,PV}} + qI_2(t) + w_{pe}^*(u_1+u_3)/2),\\
\tau_s s_2{}' &= -s_2 + f((w_{se} + bF_2)u_2 - w_{sp}p_2 - w_{ss}s_2 + I_{\text{Opt,SST}}),
\end{aligned}$$

where $I_2(t) = (i_1(t) + i_3(t))\alpha + i_2(t)$. Finally, the third, or right unit, satisfies

$$\begin{aligned}
\tau_u u_3{}' &= -u_3 + f(w_{ee}u_3 - (w_{ep} - a(1-D_3))p_3 - w_{es}s_3 - F_3 s_2 + qI_3(t) + w_{ee}^* u_2),\\
\tau_p p_3{}' &= -p_3 + f(w_{pe}u_3 - w_{pp}p_3 - w_{ps}s_3 + I_{\text{Opt,PV}} + qI_3(t) + w_{pe}^* u_2),\\
\tau_s s_3{}' &= -s_3 + f((w_{se} + bF_3)u_3 - w_{sp}p_3 - w_{ss}s_3 + I_{\text{Opt,SST}}),
\end{aligned}$$

where $I_3(t) = i_2(t) + i_3(t)\alpha$. Note that each set of equations are almost identical to the single-unit case, but with the addition of lateral terms along with facilitating and depressing terms F_i and D_i. The lateral terms are between immediate neighbors and include lateral SST to Exc (facilitating), Exc to Exc, Exc to PV, and PV to Exc (depressing). The facilitating terms F_i increase from 0 to nonzero values as unit i receives inputs, and the depressing terms s D_i decrease from 1 to lower values as unit s i receives inputs. While the three-unit rate model appears to be substantially more complex, the parameters are strongly constrained by the single-unit rate model. Each unit of the three-unit model is designed to mimic the excitatory and inhibitory currents of the single-unit rate model. The purpose is to both constrain the model parameters as much as possible and to retain the currents of the single-unit model. The latter reason is important because it explains many of the known optogenetic experiments in the literature. Before turning to the spiking model, we briefly describe technical details of the parameter values and functions.

We chose $\alpha = 0.65$, i.e., 65% of the thalamic inputs to the left or right units reach the center unit. Likewise, 65% of thalamic inputs to the center unit reach the left and right units. The function f is threshold linear (Equation 2). The functions $I_k(t)$ are time-dependent inputs with the strongest preference for unit k, and the profiles of i_1, i_2, and i_3 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 its own depression variable g_j, where each g_j satisfies Equation 3 independently. The parameters τ_i are membrane time constants and chosen the same as the single-unit model, $\tau_u = \tau_p = \tau_s = 10$ms [3,8]. 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 $w_{ee}^* = 1$, $w_{pe}^* = 1.25$, and $w_{se}^* = 0.125$ to reflect the generally weaker lateral synaptic strengths in auditory cortex relative to the within-unit connections [17].

Figure 2

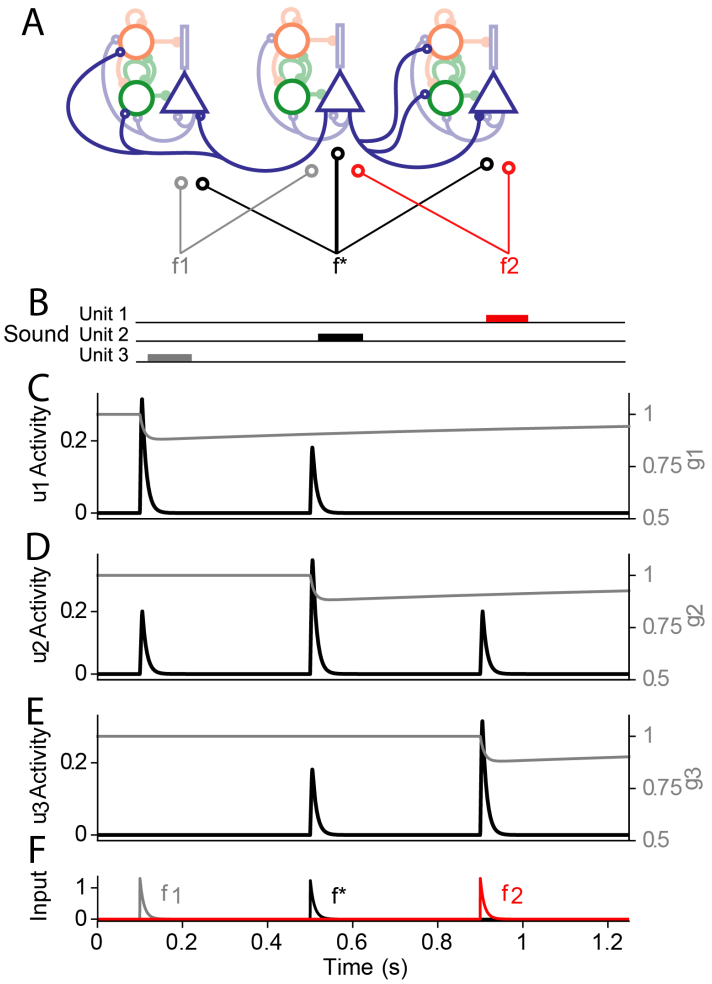

Figure 2. Input and response profiles of the three-unit model. A: The three-unit rate model of the auditory cortex, with three preferred frequencies, f_1, f^*, and f_2 (the spiking model follows the same motif). B: 50ms auditory inputs are applied at each frequency in sequence. C—E: Black traces show 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 in the Exc to SST synapses, and depressing terms D_i in the PV to Exc synapses [18]. The parameters a and b control the degree of depression and facilitation, respectively, and we chose $a = 0.5, b = 2$. The depressing parameter a was chosen carefully such that the term $(w_{ep} - aD_i)$ did not change sign across experimental paradigms. The facilitating variables F_i satisfy

$$F_j' = -\frac{F_j}{\tau_{D_1}} + \frac{i_j(t)}{\tau_{D_2}}, \tag{5}$$

where τ_{D_1} and τ_{D_2} are as in Equation 3. We used the inputs $I_i(t)$ as a proxy for the excitatory activity $u_i(t)$ so we could simulate the facilitation variable in terms of the depression variable as $F_i = 1 - g_i$. Similarly, the depression variables D_i satisfy

$$D_j' = \frac{1 - D_j}{\tau_{D_1}} - \frac{D_j i_j(t)}{\tau_{D_2}}, \tag{6}$$

and used the thalamic input as a proxy for excitatory activity to simulate the depression variable as $D_i = g_i$. All depressing and facilitating timescales were chosen close to reported values [8,11,12].

The activity of the model is shown in Figure 2, where three successive auditory stimuli were 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) responded equally well to both f_1 and f_2 (Figure 2D), which is a necessary response for SSA paradigms. For simplicity, activation of an adjacent unit did not affect the thalamic variable, i.e., g_1, g_2, and g_3 were left unaffected by u_1, u_2, and u_3, respectively. 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).

We incorporated paradigm-dependent baseline states in the three-unit rate model. The parameters switch between weak and strong baseline inhibition, where weak inhibition corresponds to high thalamic activity, and strong inhibition corresponds to relatively low thalamic activity. This idea is captured more precisely by the facilitating variable,

$$\overline{F}' = -\frac{\overline{F}^2}{\tau_{F_1}} + \frac{I(t)}{\tau_{F_2}}, \tag{7}$$

where $I(t)$ is the sum of all thalamic inputs (independent of the tonotopic arrangement), $\tau_{F_1} = 1500$, and $\tau_{F_2} = 100$. As the experimental paradigm progresses, $\overline{F}$ grows and eventually saturates (over the course of approximately 15 seconds). A simulation of Equation 7 is shown in Figure 1C for the various auditory paradigms.

If $\overline{F}$ is above the threshold $F_{th} = 0.22$, the system exhibits weak baseline inhibition, and the synapses take baseline values as shown in Equation 4. On the other hand, if $\overline{F} < F_{th}$, the synapses take the strong baseline inhibitory values

$$W_2 = \begin{pmatrix} 1.1 & 3 & 3 \\ 1 & 2 & 2 \\ 6 & 0 & 0 \end{pmatrix}, \tag{8}$$

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., $w_{ep}(t) = 2g(F(t)) + 3(1 - g(F(t)))$, $w_{es}(t) = g(\overline{F}(t)) + 3(1 - g(\overline{F}(t)))$, and $s_{th}(t) = 1g(\overline{F}(t))$ where

$$g(x) = \frac{1}{1 + \exp\left(-r(x - F_{th})\right)},$$

and r, the gain of the sigmoid g was chosen to be steep, e.g., $r = 25$. However, for simplicity, we replaced g 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 8) based on the given experimental paradigm.

Table 1: Auditory paradigm parameters. SSA parameters from [3]*. Forward suppression parameters from* [5,19]*. Tuning-curve adaptation parameters from* [4]*. PV activation parameters from* [6]*.*

Suppression				
Adaptation				
Stimulus duration	100ms	50ms	100ms	50ms
Inter-stimulus interval	300ms	20ms	300ms	-
Inter-trial interval	-	380ms	2400ms	950ms
Stimui per trial	-	2	8	1

For each paradigm (with paradigm parameters shown in Table 1, we simulated Equation 7 and found that SSA and forward suppression belonged to the weak inhibitory regime ($\overline{F}$ integrated to values above threshold F_{th}), whereas tuning-curve adaptation and PV activation belonged to the strong inhibition regime ($\overline{F}$ integrated to values below threshold F_{th}). All rate models were simulated using the dynamical systems software XPP [20], called using PyXPPCALL and visualized using Python [21].

Spiking Neuron Dynamics

We additionally considered a spiking model equivalent of the rate model. The spiking model is far more complicated than the rate models, but as in the case of the three-unit rate model, we use the single-unit rate model as a template, greatly 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 [22,23]:

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

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

$$I_{\text{Syn},i}^A = -\sum_B g_{AB,i}(t)\left(V_i^A(t) - E^B\right), \tag{9}$$

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 (inhibitory), then $E^B = 0\text{mV}$ (-67mV). If a synaptic connection existed from PV to Exc, we included a depression variable, D , satisfying Equation 6, with $\tau_{D_1} = 1000$, and $\tau_{D_2} = 250$:

$$I^E_{\text{PV},i} = -g_{ep,i}(aD(t))(V^E_i(t) - E^{PV}),$$

where $a = 1.7$. As in the rate model, the parameter a was chosen such that the sign of $I^E_{\text{PV},i}$ did not change. The additional depression term was necessary to incorporate depression effects that operate well beyond the timescale of inhibitory conductances [12].

Noise in our model comes from a white noise process with zero mean and a standard deviation of 20mV, to simulate intrinsic and extrinsic membrane fluctuations. All fixed parameters for each neuron type are shown in Table 2. The parameters that we varied manually were entirely contained in the time-dependent functions $I^A_{\text{Thal}}(t)$ (thalamic inputs) and $I^A_{\text{Opto}}(t)$ (optogenetic parameters). The thalamic input profile, $I^A_{\text{Thal}}(t)$, is determined by

$$I^A_{\text{Thal}}(t) = qI^A_{\text{Fast}}D^A_{\text{Slow}}D^A_{\text{Fast}},$$

where

$$\begin{aligned} dD^A_{\text{Slow}}/dt &= (1 - D^A_{\text{Slow}})/\tau_{D_1} - D^A_{\text{Slow}}I^A(t)/\tau_{D_2} \\ dD^A_{\text{Fast}}/dt &= (1 - D^A_{\text{Fast}} - I^A(t))/\tau_{D,\text{Fast}} \\ dI^A_{\text{Fast}}/dt &= (-I^A_{\text{Fast}} + I^A(t))/\tau_I, \end{aligned}$$

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^A_{Thal}) are square wave functions that are active for the duration of the auditory stimulus. Just as in the rate model, the thalamic input function $I^A_{\text{Thal}}(t)$ only appears in Exc and PV neurons. The profile of the thalamic input is shown in Figure 1G. The optogenetic term, $I^A_{\text{Opto}}(t)$, 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 $g_{ij} \to g_{ij} + g_{ij,\text{max}}/n_{\text{X}}$, where $g_{ij,\text{max}}$ is given by Equation 10, 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 [13], with the same type of connectivity structure as in the rate model .

$$G_{\text{max}} = \begin{pmatrix} g_{ee,\text{max}} & g_{ep,\text{max}} & g_{es,,\text{max}} \\ g_{pe,\text{max}} & g_{pp,\text{max}} & g_{ps,\text{max}} \\ g_{se,\text{max}} & g_{sp,\text{max}} & g_{ss,\text{max}} \end{pmatrix} = \begin{pmatrix} 20 & 40 & 20 \\ 2 & 40 & 40 \\ 120 & 0 & 0 \end{pmatrix} \text{nS}. \tag{10}$$

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

$$\frac{dg_{ij}}{dt} = -g_{ij}/\tau_{ij},$$

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, $\tau_{se} = 15\text{ms}$ [24]. 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 $I^A_i = I^A_{\text{Syn},i} + I^A(t) + I_{\text{Dend},i}$, where

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

The term F is a dimensionless slow timescale facilitation variable that depends on the thalamic drive, and satisfies Equation 5 (just as in depression, the additional slow timescale allows the model to operate on multiple timescales [12]). The parameter $b = 3$ modulates the facilitation strength, and $\tau_{F_1} = 1000$, and $\tau_{F_2} = 250$. For simplicity, we allowed F_i to vary continuously over time. The variable w represents spike-frequency adaptation and obeys

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

The dynamics of the dendritic (stick) compartment obey

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

where the parameter $\kappa = 0.3$ is the ratio of somatic to total surface area [23].

For PV and SST interneurons, the equations are the same as Exc except that there is no dendritic component, and parameters differ marginally (see Table 2). SSTs, unlike PVs, have no incoming synaptic connections from the thalamus, PVs and other SSTs and only receives excitatory input from Exc. Both PVs and SSTs include the optogenetic term $I^A_{\text{Opto}}(t)$, and as mentioned above, only Exc and PVs contain the thalamic input term $I^A_{\text{Thal}}(t)$. These connections reflect the choices made in the rate model.

Table 2: 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). Like the rate model, the thalamic inputs have weaker lateral connections. For tone responses at frequency f_1 (f_2), the center unit receives an input of amplitude proportional to 0.85 that of the left and right units.

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 [25].

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 frequency in the stimulus (standard tone in an oddball), while preserving the initial strong response to the less frequent input (deviant tone) [26]. Previous studies found that following a presentation of the deviant tone, the excitatory neurons adapt over successive presentations of the standard [26,27]. This phenomenon was largely attributed to feedforward thalamo-cortical depressing synapses [14,28], but such models could not account for the full range of the effects that were observed [9]. A recent study found that inhibitory neurons exhibit differential control over the stimulus-specific adaptation [3]. Suppressing SSTs resulted in disinhibition of the excitatory neurons, such that disinhibition increased with successive presentations of the standard tone, but not the deviant. By contrast, PV inhibition drove equal amount of disinhibition of excitatory neurons in response to both the deviant and the standard. These results suggest that SST inhibition increases with adaptation level of excitatory neurons.

We tested whether our model could reproduce the differential effects of suppressing PVs and SSTs on stimulus-specific adaptation in AC. We first modeled adaptation to 8 successive tones in a single iso-frequency circuit with constant synapses and depressing thalamic inputs (Figure 3, Equation 1). A key mechanism behind this result is the temporal structure of the responses: PVs exhibit a temporally fast tone-evoked response and peak earlier than Exc and SSTs, while SSTs exhibit a temporally delayed and broad tone-evoked response (Figure 1B, Figure 3B top left plot) in agreement with earlier studies [3]. Moreover, the SST delay is not hard-coded, but the result of SSTs receiving indirect thalamic excitation through the Exc population [13].

With PV inactivation (Figure 3C middle row), the Exc population received an overall decrease in inhibitory current across all successive tones, resulting in constant disinhibition prior to and following adaptation (Figure 3D green). 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). 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, causing Exc disinhibition (Figure 3C, bottom right). This disinhibition increased over the successive tones (Figure 3D, orange). Strikingly, Figure 3D, bottom is precisely the type of result found in optogenetic results in stimulus-specific adaptation [3]. These simple mechanisms of disinhibition and compensation can therefore explain the complementary roles of inhibitory interneurons in shaping cortical activity.

Figure 3

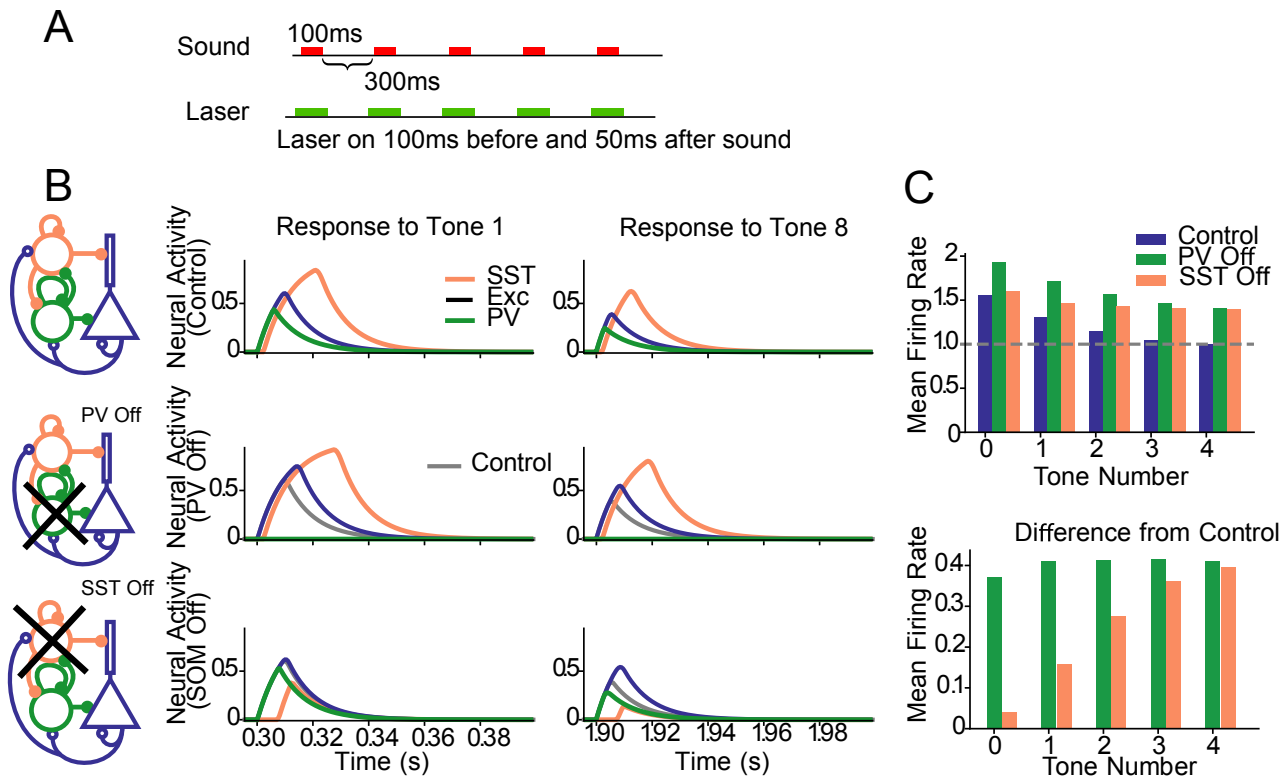

Figure 3. The effect of optogenetic manipulations on adaptation to repeated tones. A. Stimulus of repeated tones, with or without concurrent laser stimulation. B. Left: Circuit diagram specifying the inactivation of populations. Responses of Exc (blue), PV (green) and SST (orange) populations to the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

Although the single-unit rate model provided an important insight, the model lacked the ability to distinguish between auditory frequencies, which is a necessary aspect of stimulus-specific adaptation and many other optogenetic paradigms. To further test the inhibitory mechanisms discovered using the single-unit model, 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. As mentioned above, the three-unit circuitry was based on the single-unit model and the parameters chosen to reproduce the inhibitory and excitatory currents. For example, an auditory input to the left unit caused lateral excitatory and inhibitory currents to enter the center unit. These currents to the center unit were designed to be similar to the currents in the case of the single-unit. Due to symmetry, we easily performed the same procedure for auditory inputs to the right unit: lateral excitatory and inhibitory currents from the right were designed to enter the center unit in a manner similar to the single-unit case. Using this procedure, we were able to extend the differential SST and PV inhibition in the single-unit model to work in the case of a tonotopy without the need for exhaustive parameter fitting.

The procedure for reproducing SSA is as follows: record from the center unit and apply the standard tone to the right unit using the stimulus interval and duration in Table 1 (Figure 4A). The synaptic depression in the thalamus then adapts the responses, and the center unit responds with a mean firing rate similar to Figure

3D, tone number 4. With a 10% chance, apply the oddball tone to the left unit. At this event, the center unit responds with a firing rate similar to Figure 3D, tone number 1 because the thalamic input of the left unit has not adapted. In this brief time, the stimulus has skipped the right unit, and the depression variable of the right unit recovers more than usual. The next several standard tones applied to the right unit produce center-unit responses similar to Figure 3D tone numbers 2—4, because the depression variable has had a little time to recover, but not so much that the first response is like the oddball response. Continued tones to the right unit evoke center-unit responses similar to Figure 3D tone number 4. This process is repeated with PV and SST suppression. Results are shown in Figure 4B—G.

Figure 4

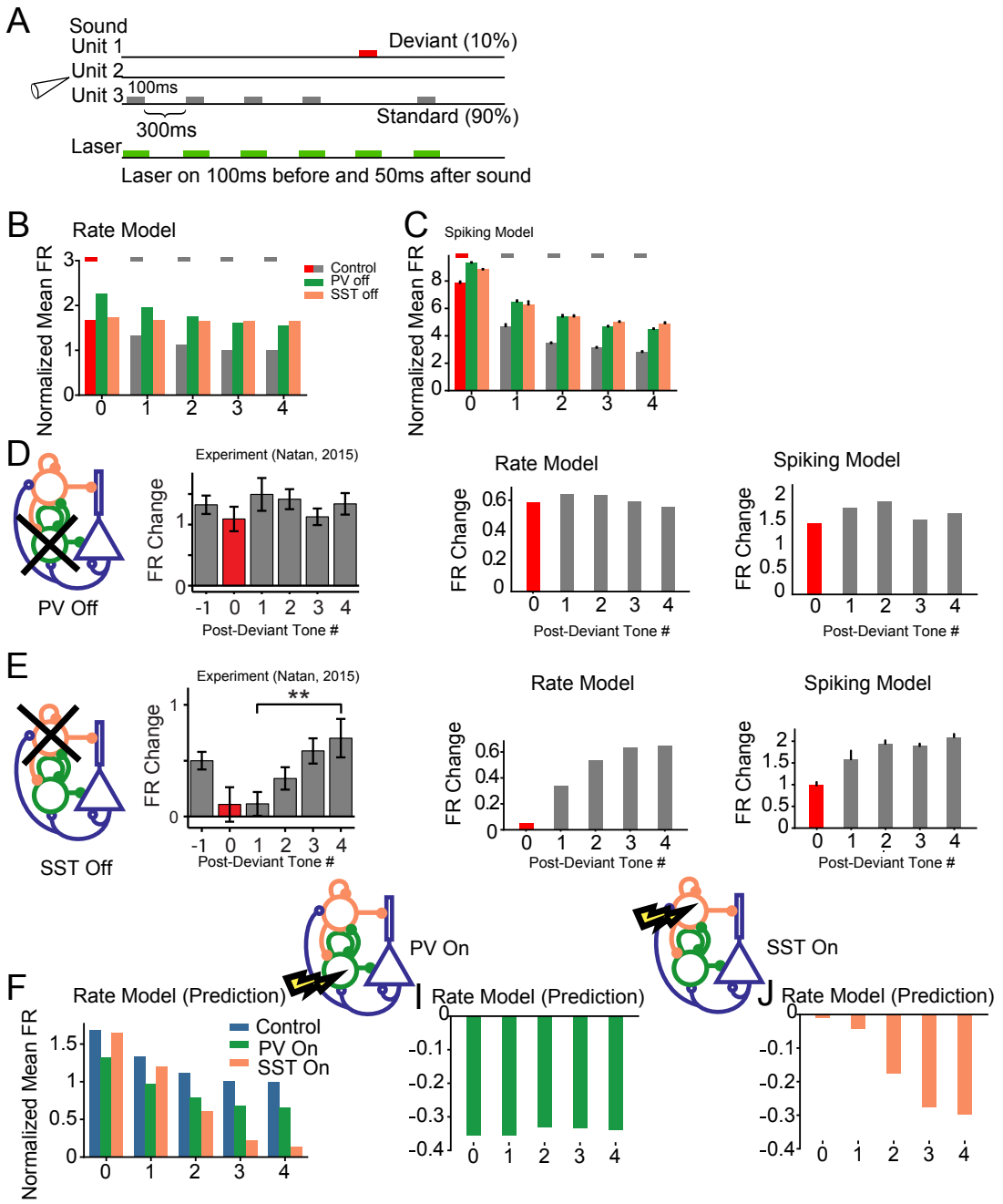

Figure 4. Summary of SSA in the rate and spiking model. A: Oddball stimulus consisted of two tones: standard tones (gray) appear with 90% probability, whereas deviant tones (red) appear with 10% probability. B, C: Average response of the excitatory population to the deviant (red) and subsequent standards (gray) without stimulation;

with PV suppression (green) and with SST suppression (orange). B. Rate model. C. Spiking model. D. Change in response of excitatory population due to PV suppression in the rate and spiking models to the deviant (red) and standards (gray). Left: from published data. Center: rate model. Right: Spiking model. E. Change in response of excitatory population due to SST suppression in the rate and spiking models to the deviant (red) and standards (gray). Left: from published data. Center: rate model. Right: Spiking model. F. Predictions for the responses to the oddball stimulus with and without interneuron activation. Left: Mean responses of the excitatory population to the deviant and subsequent standards (red/gray: no activation; green: PV activation; orange: SST activation). Change in excitatory neuron responses due to PV activation (middle), and SST activation (right). 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 [3]. In order to establish the robustness of these results, we varied several parameters and measured the Common-contrast SSA Index (CSI) [9],

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

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). For the first parameter, we chose recurrent excitation (w_{ee}, Figure 5A,B,E,F), because it is a key parameter considered in many studies [9], especially those related to inhibitory stabilized networks (ISNs) [8,23]. For the second parameter, we chose the timescale of thalamic depression (τ_{d_1}, Figure 5C,D,G,H) because reported values vary over a large range, from 0.8s to 3s [3,9]. Finally, we chose the remaining two parameters to be the strength of PV activation or inactivation (Figure 5A,C,E,G), and the strength of SST activation or inactivation (Figure 5B,D,F,H). These choices allowed us to generate experimentally testable predictions. Control parameters are denoted by a white circle, PV and SST inactivation parameters are denoted by white triangles, and PV and SST activation parameters are denoted by white squares. In all cases, inactivating SSTs had a much greater effect on decreasing the CSI, reflecting the increasing disinhibition over post-deviant tones.

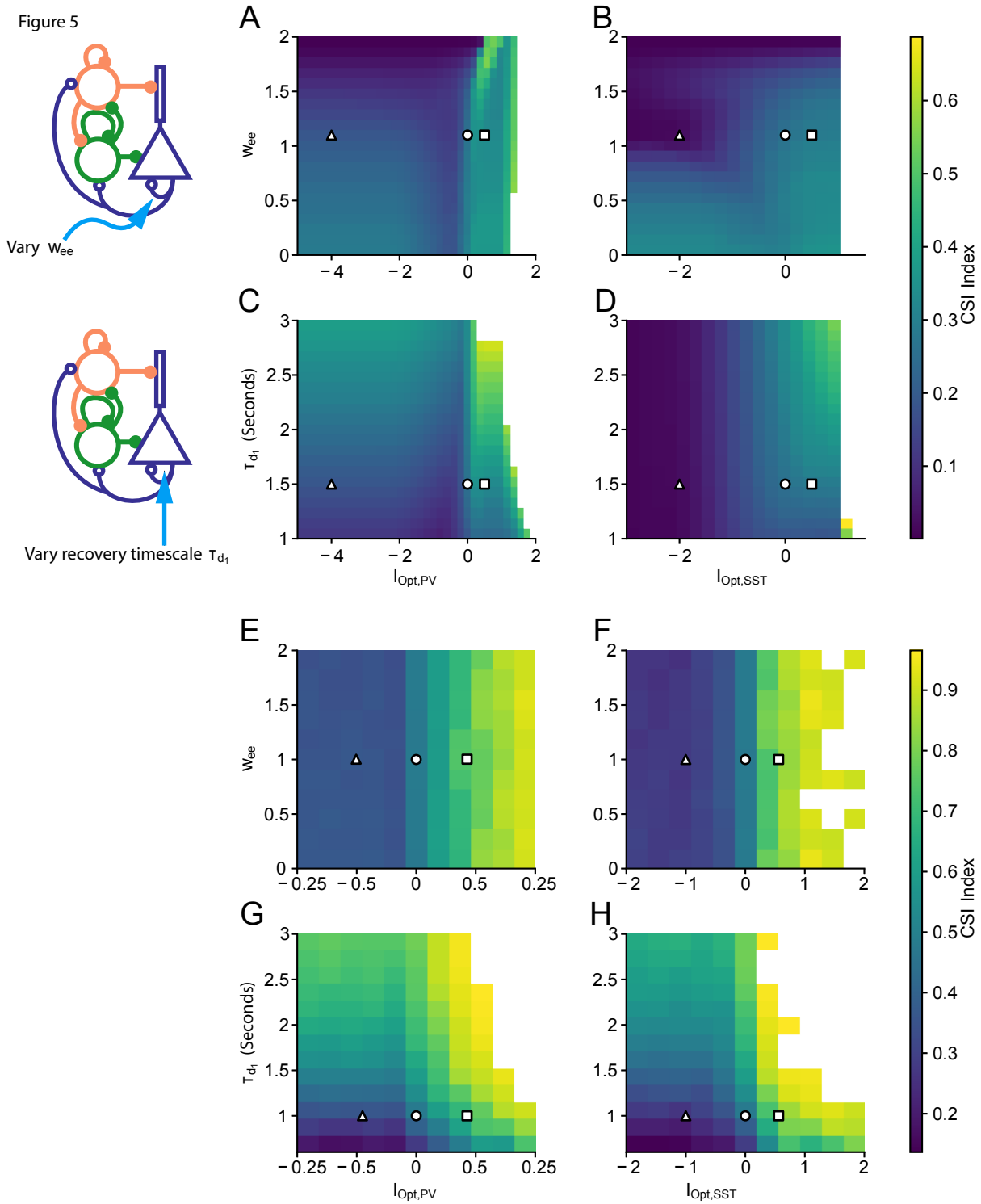

Figure 5. The effect of the key parameters on SSA index (CSI) for rate model (A-D) and spiking model (E-H). Optogenetic inactivation parameters are marked by white triangles, and optogenetic activation parameters are marked by white squares. The control parameter values, $I_{Opt,PV} = I_{Opt,SST} = 0$, are marked by the 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.

These plots reveal 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 (i.e., shifting the white circle up and down). This result 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 (i.e., shifting the white triangle right) had little effect on the CSI, increasing optogenetic activation (i.e., shifting the white square left) 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_{d_1}. 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 ($p = 0.6$). Moreover, more of the excitation in the spiking model comes from the thalamus. The goal of the present study is not a thorough examination of the parameter space, so we leave the details of this question to future studies.

We reiterate that 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. As we have seen, 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 1, Figure 1C, and Equation 7).

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-studies 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 suppression, but the inhibitory neurons were recently shown to also control forward suppression. 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) [5].

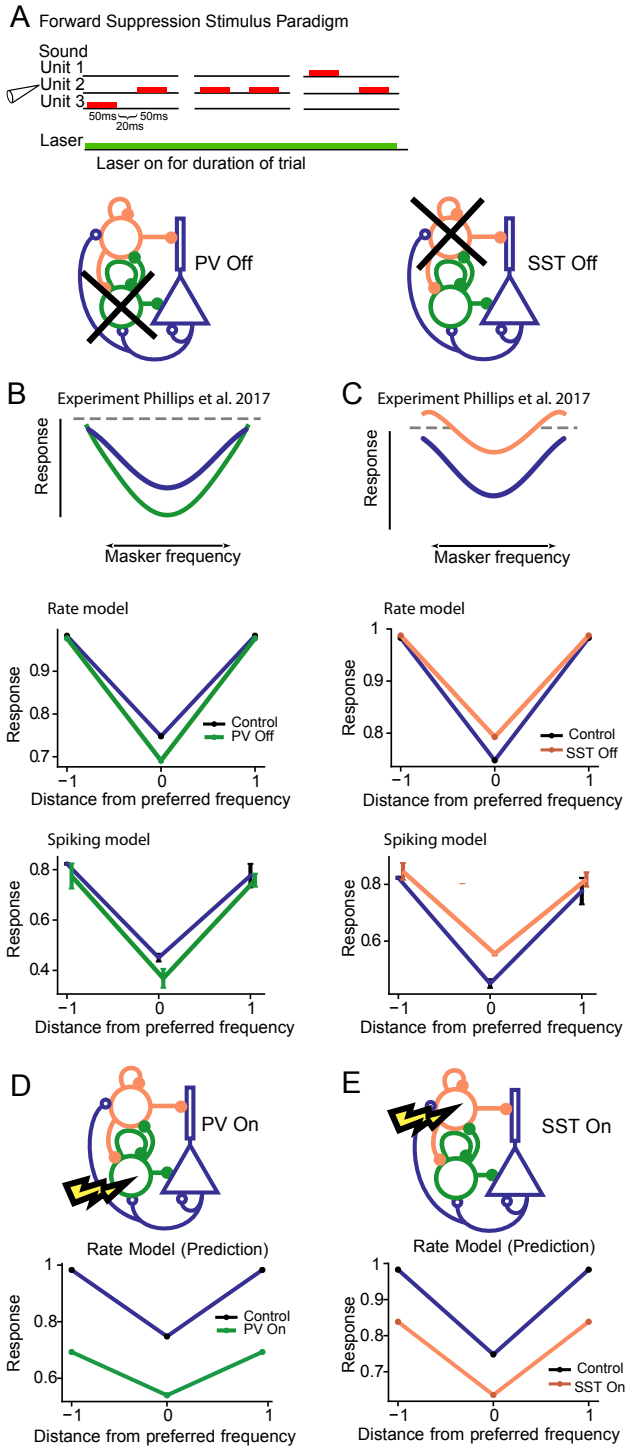

Figure 6. Forward suppression in the rate and spiking model. A. The stimulus consisted of pairs of tones activating either neighboring or the same iso-frequency units. The laser was presented continuously throughout stimulation trials. B,C, D, E: Responses of excitatory neurons to the probe tone as a 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. 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 6B bottom) yielded the experimentally measured 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.

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 this disinhibition forces the thalamic variable g to decrease more compared to the control case. Upon receiving the second tone, the input received by the excitatory population is weaker, in turn weakening the excitatory response. This weakening in the light-on trial is proportionally greater compared to the control case, so forward suppression is strengthened. 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 [3]. 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 [4]. 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 of the frequency response function of excitatory neurons [4]. To understand how inhibitory neurons affect adaptation across different frequency-tuned inputs, we presented a sequence of 8 tones at each frequency to generate adapting tuning curves (Figure 7A), and repeated this process with PV and SST inactivation for the model circuit. We found that this auditory paradigm resulted in a below-threshold integration of $\overline{F}$, so the system switched to a state of strong baseline inhibition (and importantly, the model did not respond in precisely the same way as in SSA and forward suppression).

Our model reproduced the differential experimental effects of PV and SST suppression (Figure 7). In the rate model before adaptation, PV and SST inactivation resulted in sideband disinhibition with little to no disinhibition at the preferred frequency (Figure 7B left, F left). After adaptation, PV inactivation resulted in sideband disinhibition and no preferred frequency disinhibition (Figure 7B right), whereas SST inactivation resulted in disinhibition across all sideband and preferred frequencies (Figure 7F right). The spiking model closely mirrored these results (Figure 7D,H). The ratio of excitatory responses between light off and light on trials summarize the degree of sideband and preferred frequency disinhibition (Figure 7 C, E, G, and I). Thicker lines

in Figure 7C, E, G and I 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).

The mechanisms behind these results involve the synaptic dynamics and the compensating mechanism discussed in the earlier sections for SSA and adaptation. In the case of PV suppression, SSTs were the only interneurons capable of contributing to Exc inhibition, so only Exc-SST synaptic dynamics drove the observed effects. In particular, lateral SST to Exc synapses suppressed the center unit over each tone, and facilitation allowed this suppression to persist throughout adaptation. Note that this preferred-frequency effect was not observed in SSA because we never directly stimulated the center unit. Next, in the case of SST suppression, the increasing disinhibition with adaptation at the preferred frequency was a consequence of the same compensating mechanism as in SSA. Our model is the first to provide a microcircuit mechanism for the observed experimental results.

Our model predicted that before adaptation, PV activation resulted in a slight decrease at the preferred frequency, whereas SST inactivation reduced overall firing rates across all frequencies (Figure 7J, L). After adaptation, PV and SST activation resulted in a subtractive effect. General optogenetic activation and inactivation of PVs and SSTs have been shown to modulate tuning-curves in combinations of additive, subtractive, multiplicative, and divisive effects [29]. Our model reproduced one of the key results these studies as well, since PV and SST inactivation were found to have additive and divisive effects on the frequency response functions of excitatory neurons (Figure C, E, G, I, K, M).

Figure 7

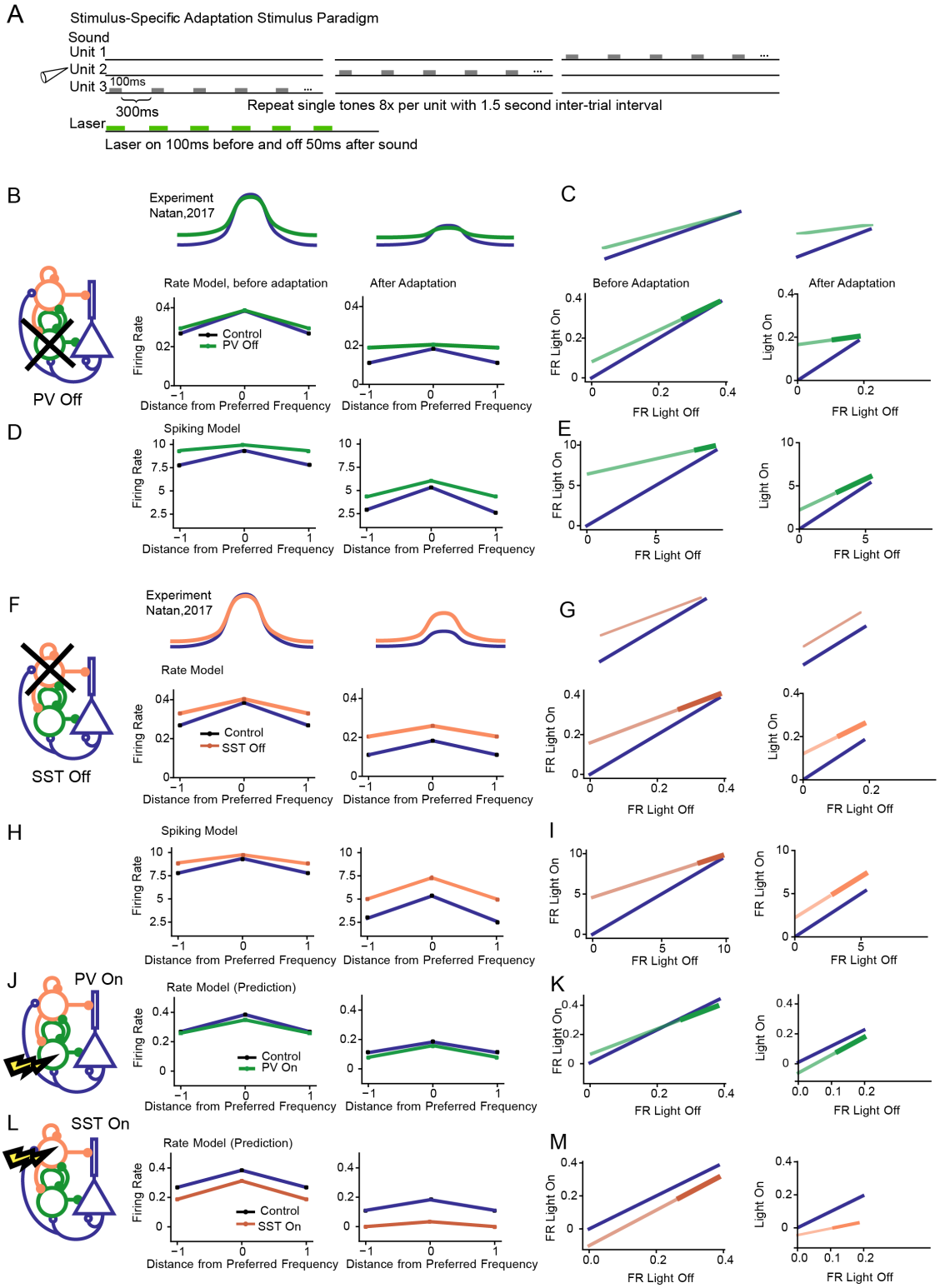

Figure 7. Adaptation to repeated tones along the frequency response function. A. The stimulus consisted of a sequence of repeated tones, presented to each iso-frequency unit. On stimulation trials, the laser overlapped with the sound stimulus. B, D, F, H: The responses of excitatory units to the first (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: The response of excitatory neurons to tone 1 (left) and tone 8 (right) on light on and light off trials. B. Rate Model, PV suppression. D. Spiking model, PV

suppression. F. Rate model, SST suppression. H. Spiking model, SST suppression. J. Rate model, PV activation. L. Rate model, SST activation.

PVs Enhance 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 [6] 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 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 functional connectivity (Figure 8).

Figure 8

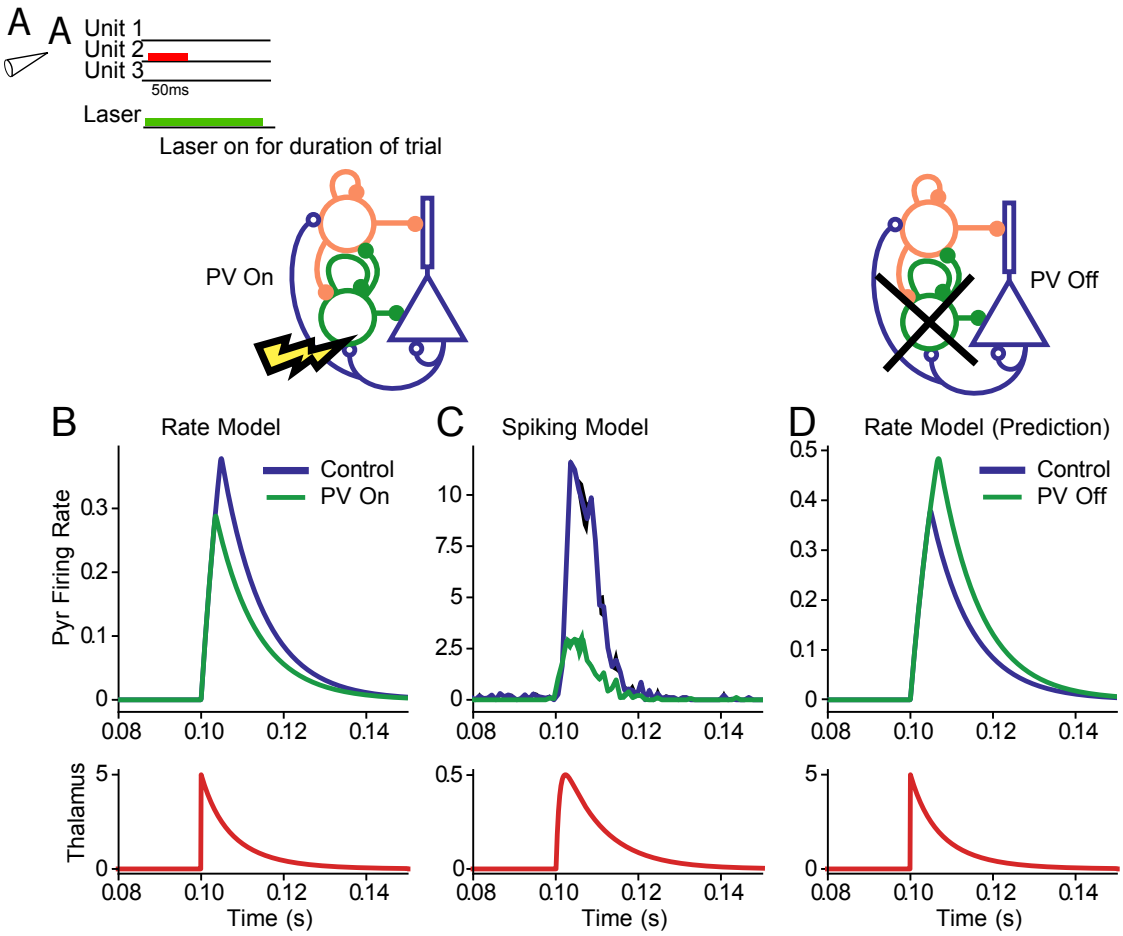

Figure 8. Activation of PVs enhanced feedforward connectivity in the model. A. Stimulus was a single tone accompanied by a laser on stimulation trials presented at 0.1 s. B. Top: Cortical excitatory population responses to tones without (blue) and with PV stimulation (green) in the rate model. Bottom: 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 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, but we shifted the data for simplicity). 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 functional connectivity.

These results provide for a simple plausible mechanism for enhanced functional connectivity in the cortex: as inhibition reduces the overall cortical inputs, cortical responses better synchronize to thalamic inputs, resulting in stronger correlated activity. We remark that whereas correlations in general do not measure functional connections as the Ising model, 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.

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, such as SSTs and PVs, play a differential role in auditory processing, controlling adaptation at different time scales and contexts, as well a functional connectivity. Our goal was to integrate the results of these studies to understand whether the observed effects are due to a specific unified circuit, or to different modes and connectivity patterns potentially found across AC.

We built a unifying rate and spiking model that reproduced multiple key results from studies that tested the function of specific inhibitory opsins in specific cells in the auditory cortex. In addition to including different baseline states that modulate the strength of PV-to-Exc and SST-to-Exc synapses, the key mechanisms underlying our models are the fast temporal activation of PVs, the delayed, broad temporal activation of SSTs, SST-to-Exc facilitation, and PV-to-Exc depression. These interactions account for the differential modulation of cortical responses by interneuron subtypes and proposed a unified set of connectivity parameters that support experimental results.

To reproduce the differential function of SSTs and PVs in stimulus-specific adaptation, we built a model loosely based on multiple existing models for SSA and multiple configurations of spiking neuron populations, consisting of inhibitory and excitatory neurons. Previously, a two-layer rate model with synaptic depression was proposed to establish the relationship between the cortical response and the parameters in SSA experiments, such as stimulus frequency differences, probability of deviation, and tone presentation rate. Yarden et al. (2017)

successfully used a multi-unit rate model arranged in a coarse tonotopy consisting of inhibitory and excitatory populations to reproduce general deviance detection, but model has not yet been adapted to explain differential interneuron modulation. Another existing model of SSA including differential inhibitory modulation demonstrating similar differential inhibitory effects as in our SSA result (Figure 4), but did not include a tonotopy [3]. These models only included one type of inhibitory neuron type or did not include tonotopy, and therefore could not account for the observed differential effects of suppression of SSTs and PVs on SSA across multiple frequencies. In the present study, we developed a simple rate and spiking model that accounted for multiple inhibitory cell types and which faithfully reproduced the differential effects of SST and PV inactivation in SSA (Figure 4). In addition, a parameter sweep revealed that both the rate and spiking models were robust to large changes in key parameters commonly explored in the literature, suggesting that SSA is a robust phenomenon [9].

Figure 9

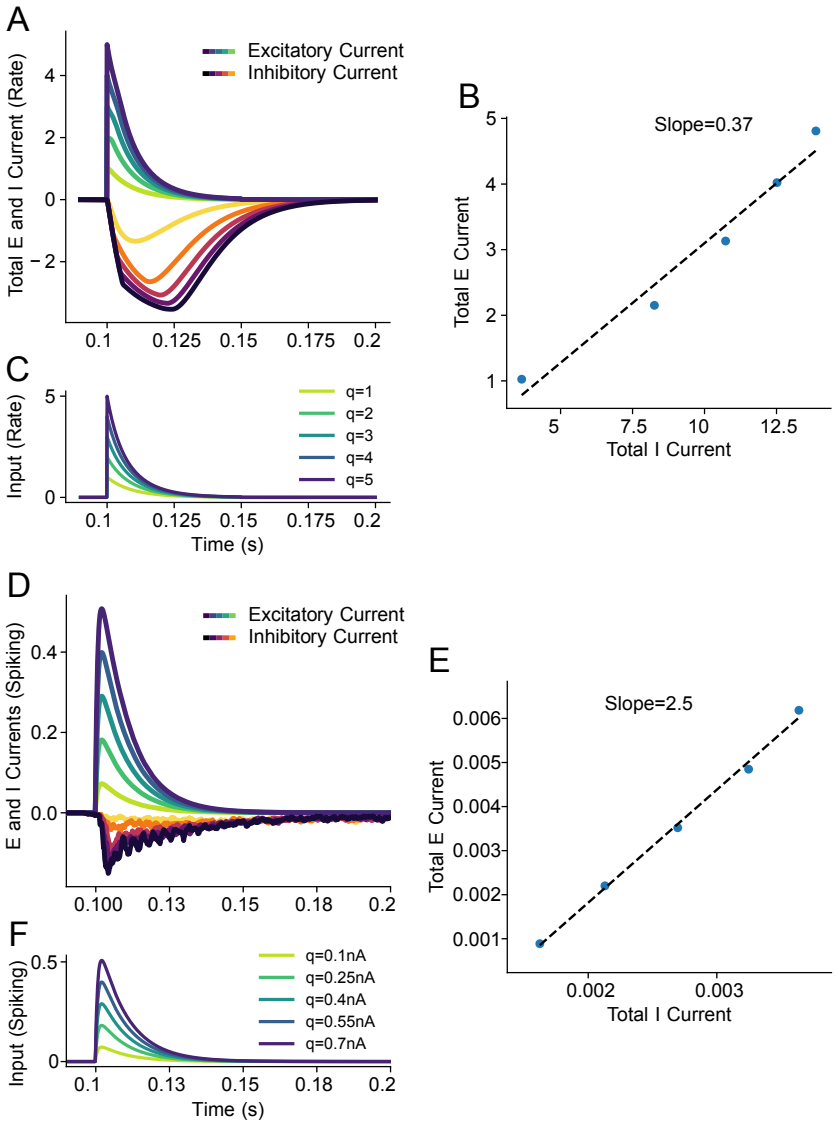

Figure 9. Excitatory-inhibitory balance in the rate and spiking models A. Plot of incoming excitatory and inhibitory currents into the Exc population as a function of different input strengths (C). Darker currents correspond to stronger inputs B. A best-fit line (dashed) accurately captures the ratio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

Existing models that reproduce the enhanced forward suppression from PV inactivation and the reduced forward suppression from SST inactivation (Figure 6) include multiple layers that require both depression and facilitation [5], or rely on depressing recurrent excitation and do not distinguish between inhibitory subtypes [30]. We incorporated depression and facilitation in the model synapses and reproduced the former results with only a single layer, suggesting a surprisingly simple mechanism supporting forward suppression. Furthermore, the models in the present study reproduced tuning-curve adaptation effects previously observed: SSTs exhibited strong preferred-frequency disinhibition following adaptation, while PV disinhibition is independent of the degree of adaptation (Figure 7) [4]. These results suggest that the underlying mechanism(s) of the model, namely the PV/SST compensation effect, combined with the facilitating SST-to-Exc synapse and depressing PV-to-Exc synapse, may serve as a unifying mechanism of adaptation.

Finally, our models reproduced changes in functional connectivity in the cortex (Figure 8). By increasing PV activity in the models, excitatory activity decreased but became more time-locked to thalamic inputs. This effect agrees with observations in the cortex, where PV activation results in enhanced functional connectivity [6]. The effects of inhibition on sharpening cortical responses have been well-established, thus our models serve as plausible mechanisms for this change [31–33].

One drawback of the models is that they do not feature population spikes, which explain many fundamental cortical responses in AC [31]. In future work, we will seek to reconcile the differences between our models and the population spike model of SSA [30]. Establishing the importance of depression and facilitation in different synapses and extending our model to include population spikes warrants further study.

Our models show evidence of operating as a balanced network: as we vary 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) remains constant: the rate model has an excitatory/inhibitory ratio of 0.37 (Figure 9B), and a ratio of 2.5 for the spiking model (Figure 9E). These results demonstrate the potential importance of inhibitory-excitatory balance in the cortex. In future studies, we will explore whether inhibitory-excitatory balance is a necessary or sufficient condition for the results in this study.

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

Multiple studies from different laboratories revealed the differential effect of distinct inhibitory neurons in auditory processing. Strikingly, a minimalistic model, built on simple mechanisms, produced differential information processing by various subtypes of inhibitory neurons, and has unified these disparate studies. As inhibitory neurons form similar circuits throughout the mammalian cortex, this model can be readily adapted to test their function and generate predictions (with adjustments for local changes in connectivity) in different sensory modalities.

6

Acknowledgments

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

Bibliography

1. Fino E, Packer AM, Yuste R. The logic of inhibitory connectivity in the neocortex. Neurosci. 2013;19: 228–237.

2. Isaacson JS, Scanziani M. How inhibition shapes cortical activity. Neuron. 2011;72: 231–243.

3. Natan RG, Briguglio JJ, Mwilambwe-Tshilobo L, Jones SI, Aizenberg M, Goldberg EM, et al. Complementary control of sensory adaptation by two types of cortical interneurons. Elife. 2015;4.

4. Natan RG, Rao W, Geffen MN. Cortical interneurons ensure maintenance of frequency tuning following adaptation. bioRxiv. 2017; 172338.

5. Phillips EAK, Schreiner CE, Hasenstaub AR. Cortical interneurons differentially regulate the effects of acoustic context. Cell Rep. 2017;20: 771–778.

6. Hamilton LS, Sohl-Dickstein J, Huth AG, Carels VM, Deisseroth K, Bao S. Optogenetic activation of an inhibitory network enhances feedforward functional connectivity in auditory cortex. Neuron. 2013;80: 1066–1076.

7. Wilson HR, Cowan JD. Excitatory and inhibitory interactions in localized populations of model neurons. Biophys J. 1972;12: 1–24.

8. Tsodyks M V, Skaggs WE, Sejnowski TJ, McNaughton BL. Paradoxical effects of external modulation of inhibitory interneurons. J Neurosci. 1997;17: 4382–4388.

9. Yarden TS, Nelken I. Stimulus-specific adaptation in a recurrent network model of primary auditory cortex. PLoS Comput Biol. 2017;13: e1005437.

10. Rankin J, Sussman E, Rinzel J. Neuromechanistic model of auditory bistability. PLoS Comput Biol. 2015;11: e1004555.

11. Abbott LF, Varela JA, Sen K, Nelson SB. Synaptic Depression and Cortical Gain Control. Science (80-). 1997;275: 221–224. doi:10.1126/science.275.5297.221

12. Wehr M, Zador AM. Synaptic mechanisms of forward suppression in rat auditory cortex. Neuron. 2005;47: 437–445.

13. Pfeffer CK, Xue M, He M, Huang ZJ, Scanziani M. Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. Nat Neurosci. 2013;16: 1068.

14. Mill R, Coath M, Wennekers T, Denham SL. A neurocomputational model of stimulus-specific adaptation to oddball and Markov sequences. PLoS Comput Biol. 2011;7: e1002117.

15. Lee CC, Sherman SM. Synaptic properties of thalamic and intracortical inputs to layer 4 of the first- and higher-order cortical areas in the auditory and somatosensory systems. J Neurophysiol. 2008;100: 317–326.

16. Womelsdorf T, Valiante TA, Sahin NT, Miller KJ, Tiesinga P. Dynamic circuit motifs underlying rhythmic gain control, gating and integration. Nat Neurosci. 2014;17: 1031.

17. Levy RB, Reyes AD. Spatial profile of excitatory and inhibitory synaptic connectivity in mouse primary auditory cortex. J Neurosci. 2012;32: 5609–5619.

18. Beierlein M, Gibson JR, Connors BW. Two dynamically distinct inhibitory networks in layer 4 of the neocortex. J Neurophysiol. 2003;90: 2987–3000.

19. Brosch M, Schreiner CE. Time course of forward masking tuning curves in cat primary auditory cortex. J Neurophysiol. 1997;77: 923–943.

20. Ermentrout B. Simulating, analyzing, and animating dynamical systems: a guide to XPPAUT for researchers and students. SIAM; 2002.

21. Rossum G. Python Reference Manual. Amsterdam, The Netherlands, The Netherlands: CWI (Centre for Mathematics and Computer Science); 1995.

22. Brette R, Gerstner W. Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. J Neurophysiol. 2005;94: 3637–3642.

23. Litwin-Kumar A, Rosenbaum R, Doiron B. Inhibitory stabilization and visual coding in cortical circuits with multiple interneuron subtypes. J Neurophysiol. 2016;115: 1399–1409.

24. Markram H, Wang Y, Tsodyks M. Differential signaling via the same axon of neocortical pyramidal neurons. Proc Natl Acad Sci. 1998;95: 5323–5328.

25. Stimberg M, Brette R, Goodman DFM. Brian 2: an intuitive and efficient neural simulator. bioRxiv. 2019. doi:10.1101/595710

26. Ulanovsky N, Las L, Nelken I. Processing of low-probability sounds by cortical neurons. Nat Neurosci. 2003;6: 391.

27. von der Behrens W, Bäuerle P, Kössl M, Gaese BH. Correlating stimulus-specific adaptation of cortical neurons and local field potentials in the awake rat. J Neurosci. 2009;29: 13837–13849.

28. Mill R, Coath M, Wennekers T, Denham SL. Characterising stimulus-specific adaptation using a multi-layer field model. Brain Res. 2012;1434: 178–188.

29. Phillips EAK, Hasenstaub AR. Asymmetric effects of activating and inactivating cortical interneurons. Elife. 2016;5.

30. Loebel A, Nelken I, Tsodyks M. Processing of sounds by population spikes in a model of primary auditory cortex. Front Neurosci. 2007;1: 15.

31. Wehr M, Zador AM. Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. Nature. 2003;426: 442.

32. Cardin JA, Carlén M, Meletis K, Knoblich U, Zhang F, Deisseroth K, et al. Driving fast-spiking cells induces gamma rhythm and controls sensory responses. Nature. 2009;459: 663.

33. Sohal VS, Zhang F, Yizhar O, Deisseroth K. Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. Nature. 2009;459: 698.