Inhibitory cell type heterogeneity in a spatially structured mean-field model of V1

Soon Ho Kim* and Hannah Choi*
School of Mathematics
Georgia Institute of Technology
Atlanta, GA 30332-0160
(Dated: March 13, 2025)

Inhibitory interneurons in the cortex are classified into cell types differing in their morphology, electrophysiology, and connectivity. Although it is known that parvalbumin (PV), somatostatin (SST), and vasoactive intestinal polypeptide-expressing neurons (VIP), the major inhibitory neuron subtypes in the cortex, have distinct modulatory effects on excitatory neurons, how heterogeneous spatial connectivity properties relate to network computations is not well understood. Here, we study the implications of heterogeneous inhibitory neurons on the dynamics and computations of spatially-structured neural networks. We develop a mean-field model of the system in order to systematically examine excitation-inhibition balance, dynamical stability, and cell-type specific gain modulations. The model incorporates three inhibitory cell types and excitatory neurons with distinct connectivity probabilities and recent evidence of long-range spatial projections of SST neurons. Position-dependent firing rate predictions are validated against simulations, and balanced solutions under Gaussian assumptions are derived from scaling arguments. Stability analysis shows that while long-range inhibitory projections in E-I circuits with a homogeneous inhibitory population result in instability, the heterogeneous network maintains stability with long-range SST projections. This suggests that a mixture of short and long-range inhibitions may be key to providing diverse computations while maintaining stability. We further find that conductance-based synaptic transmissions are necessary to reproduce experimentally observed cell-type-specific gain modulations of inhibition by PV and SST neurons. The mechanisms underlying cell-type-specific gain changes are elucidated using linear response theory. Our theoretical approach offers insight into the computational function of cell-type-specific and distance-dependent network structure.

I. INTRODUCTION

Cortical processing involves a precise interplay between synaptic excitation and inhibition [1–3]. Inhibitory neurons tune cortical oscillations, modulate gain, and gate excitatory neuronal responses [4–7]. Although cortical circuits are commonly modeled as recurrent neural networks [8], biological circuits have two features often overlooked in model networks: spatially-dependent connectivity [9] and the existence of cell types [10, 11].

Biological neural networks are spatially embedded, and distance plays an important role in governing connection probabilities [12]. To describe their effects on network dynamics, mean-field models of spatially distributed spiking neural networks have been developed in previous studies [13–15]. These studies have found that broader lateral projections can result in distinct pairwise correlation structures [15]. Furthermore, stability analysis of these models shows that the distance of projections has a critical effect on network stability: if inhibitory projections are broader than the excitatory projections, the fixed points of the network can become unstable [13], which can result in unreliable tracking of inputs [14]. These studies, however, have not examined the consequence of having cell types and assumed homogeneous inhibitory populations.

In cortical networks, diverse inhibitory subtypes including parvalbumin (PV), somatostatin (SST), and vasoactive intestinal polypeptide-expressing neurons (VIP) form circuits with excitatory (E) neurons with distinct connectivity statistics [10, 16, 17]. Several theoretical works have examined networks with heterogeneous inhibitory cell types. Wilson-Cowan-type models incorporating cell types have been used to study inhibitory stabilization and subtractive/divisive modulations [18] and to describe visual processing of objects superimposed on a background [19]. A rate-based circuit model has shown that PV and SST neurons may each be involved in encoding the mean and variance of predictions in prediction error neurons [20]. These works have suggested the role of cell types in network computations, but did not address realistic, spatially-structured connectivity among different cell types and in particular did not address distance-dependent modulations.

Understanding how cell types and their specific distance-dependent connectivity influences neural computations is an important aspect of cortical processing that is not well understood. The computational functions of inhibitory neurons can be understood through their arithmetic operations on neuronal input-output functions [21]. In the mouse primary visual cortex (V1), local and moderate activation of PV neurons divisive operations on excitatory response curves to stimuli while the

* To whom correspondence may be addressed. Email: soonhokim@gatech.edu or hannahch@gatech.edu

effect of SST neurons is subtractive [22–26]. However, recent evidence shows that when activated distally, SST neurons inhibit excitatory neurons receiving visual stimuli in a divisive manner while PV neurons exert subtractive modulations [27]. Anatomical data showed that SST neurons form broader axonal projections than PV neurons do [27–29]. Furthermore, a computational model of spiking neurons suggested that this long-range inhibition is necessary to reproduce the effects [27]. On the other hand, VIP neurons may increase the response gain, possibly through disinhibition [7, 30].

While spatially-embedded spiking neural networks of cortical populations incorporating cell types have recently been developed [27, 31], the nature of large spiking network models has limited these previous studies to simulation-based analyses, lacking theoretical insights. A systematic analysis of how position-dependent activation of PV, SST and VIP neurons modulates neural activity, therefore, is missing. A biologically-informed computational model of V1 enabling such analysis *in silico* would provide mechanistic insight into these experimental results. A mean-field model in particular can be tractable under some assumptions, allowing direct computation of important network features such as the eigenspectrum. Furthermore, such models would provide a platform to understand the role of cell-type-specific structure in information processing across cortical networks.

Here we develop a mean-field model of spatially-extended spiking neural networks in V1. We first construct the model, incorporating distinct, cell-type-specific connectivity and derive conditions that achieve excitatory-inhibitory balance in the thermodynamic limit. We use an approach based on the Fokker-Planck equation to compute the position-dependent stationary firing rates of the model and validate them against spiking neural network simulations. We examine stability features of the network, specifically how long-range SST projections affect network stability. Finally, we examine gain modulations during activation of specific inhibitory cell types both by examining shifts in the stationary firing rates with input strength and using linear response theory.

II. RESULTS

A. Network Model with Inhibitory Cell Types

We construct a spatially-organized network of spiking neurons and incorporate cell-type-specific connectivity. The model consists of N neurons that are uniformly distributed on a 1-dimensional space with periodic boundary conditions, $\Gamma = (0, 1]$ (i.e., a ring). The network consists of $N_e = q_e N$ excitatory and $N - N_e$ inhibitory neurons which are subdivided into $N_p = q_p N$ PV, $N_s = q_s N$ SST, and $N_v = q_v N$ VIP neurons with $q_e + q_p + q_s + q_v = 1$. We set $q_e = 0.8$, $q_p = 0.1$, and $q_s = q_v = 0.05$, approximately following anatomical evidence [11] (see Figs. 1a, b for model schematics).

We first assume that the neurons follow leaky integrate-and-fire (LIF) dynamics with current-based synapses (see equations in Methods). The probability with which the jth neuron of cell type β at position $y = j/N_\beta$ (where β takes on indices e, p, s, and v for E, PV, SST, and VIP neurons) connects to the kth neuron of type α at $x = k/N_\alpha$ is given by

$$k^{\Gamma}_{\alpha\beta}(x - y) = \bar{k}_{\alpha\beta} g(x - y; 0, \sigma_\beta). \tag{1}$$

Here $\bar{k}_{\alpha\beta}$ controls the overall connectivity probability from β to α (we use $\bar{\cdot}$ to denote the spatial average, i.e., $\bar{k}_{\alpha\beta} = \int_\Gamma k^{\Gamma}_{\alpha\beta}(x')dx'$). We select relative values to match recent anatomical data from layer 2/3 of V1 in mice [17] (Fig. 1c; see also Table I in Methods). Additionally, we scale all probabilities so that the overall connectivity of the network is 1%, imposing sparse connectivity. The distance-dependence is given by the wrapped Gaussian function $g(x - y; 0, \sigma_\beta)$, where σ_β is the characteristic projection distance of β-type neurons (see Methods IV A). The values for σ_β are also chosen based on recent anatomical evidence; motivated by the experimental evidence of broad SST axonal projections and more local PV and VIP projections [27–29, 32], we set $\sigma_s = 0.2$ and $\sigma_p = \sigma_v = 0.1$ (Fig 1d). If a synapse is formed from a presynaptic neuron of type β to a postsynaptic one of type α, the weight of the synapse is given by $J_{\alpha\beta}$, where the synaptic weight matrix J is tuned to achieve a balanced state as outlined below.

B. Derivation of balanced solution in the thermodynamic limit

The kth neuron of call type α receives inputs $I_{\alpha,k}$ given by a sum of recurrent synaptic inputs, described above, and a cell-type- and position-dependent external current $F_\alpha(x)$. We consider a case where $F_\alpha(x)$ have Gaussian spatial profiles. Current inputs are given by a weighted sum of a constant term and spatially structured term, $F_\alpha(x) = \bar{F}_\alpha[(1 - p_\alpha) + p_\alpha g(x; x_f, \sigma_f)]$, where the stimulus is centered at $x_f = 0.5$ and has a broadness given by $\sigma_f = 0.25$. The relative weight of the spatially-structured input is given by p_α, and $\bar{F}_\alpha = \int_\Gamma F_\alpha(x)dx$ is the mean current input to population α. Following studies of networks with E-I balance [8, 13], we assume that the synaptic weights follow scaling $J_{\alpha\beta} = j_{\alpha\beta}/\sqrt{N}$ and inputs follow $F_\alpha = f_\alpha\sqrt{N}$ with network size N, where $j_{\alpha\beta}$ and f_α do not depend on N. We assume that other parameters, such as the threshold and reset potentials, do not depend on N. Under these assumptions, a neuron receives $O(N)$ excitatory inputs but requires $O(\sqrt{N})$ presynaptic neurons to activate during an integration window to produce a spike [33]. In order for a stable stationary firing rate solution to exist, a dynamic balance between excitation and inhibition must be maintained [8, 13, 34].

The kth neuron of type α (at position $x = k/N$) has mean firing rate $\nu_\alpha(x) = \langle E[s_{\alpha,k}]\rangle$, where $E[\cdot]$ denotes

the average over network configurations, $\langle\cdot\rangle$ the time average, and $s_{\alpha,k} = \sum_{t'_{\alpha,k}} \delta(t - t'_{\alpha,k})$ is the spike train of the neuron. The mean input current is given by

$$\begin{aligned}\mu_\alpha(x) &\equiv \langle E[I_{\alpha,k}(t)]\rangle \\ &= \sqrt{N}\left[\sum_\beta (w_{\alpha\beta} * \nu_\beta)(x) + f_\alpha(x)\right] \end{aligned} \tag{2}$$

where $w_{\alpha\beta}(x') \equiv q_\beta j_{\alpha\beta} k^\Gamma_{\alpha\beta}(x')$ and "$*$" indicates circular convolution, i.e., $(w_{\alpha\beta} * \nu_\beta)(x) = \int_\Gamma w_{\alpha\beta}(x')\nu_\beta(x - x')dx'$. The balanced state exists when the mean input currents obey $O(1/\sqrt{N})$ scaling,

$$\sum_\beta (w_{\alpha\beta} * \nu_\beta)(x) + f_\alpha(x) = O(1/\sqrt{N}). \tag{3}$$

Taking $N \to \infty$ (which we call the thermodynamic limit) gives a Fredholm equation of the first kind, which in the Fourier domain becomes

$$\sum_\beta \tilde{w}_{\alpha\beta}(n)\tilde{\nu}_\beta(n) + \tilde{f}_\alpha(n) = 0 \tag{4}$$

where the tilde indicates a Fourier transform, i.e. $\tilde{h}(n) \equiv \int_\Gamma e^{-2\pi x n i} h(x)dx$, and n is the Fourier mode. The solution is given by

$$\tilde{\nu}_\alpha(n) = \frac{\sum_\beta \tilde{f}_\beta \tilde{A}_{\alpha\beta}}{\det(\tilde{W})} \tag{5}$$

where $\tilde{W}$ is a matrix whose elements are $\tilde{w}_{\alpha\beta}$ and $A_{\alpha\beta}$ are cofactors of $\tilde{W}$. Eq. 5 must hold for every n for which $\det \tilde{W} \neq 0$. If the determinant equals 0 at some Fourier mode, then for a solution to exist, $\sum_\beta \tilde{f}_\beta \tilde{A}_{\alpha\beta}$ must also equal zero at that Fourier mode.

The solutions are viable only if $\tilde{\nu}_\alpha$ has a well-defined inverse Fourier transform, such that

$$\lim_{n\to\infty} \tilde{\nu}_\alpha(n) = 0$$

for each $\alpha = e, p, s, v$. Each term $(w_{\alpha\beta} * \nu_\beta)(x)$ in Eq. 2 a convolution of two Gaussian functions, which is itself a Gaussian with spatial variance equal to the sum of those of $w_{\alpha\beta}$ and ν_β. This implies that external current inputs must be broader than the recurrent projections of each population α (i.e., $\sigma_\alpha < \sigma_f$) to satisfy Eq. 4 [13]. Under these conditions, the stationary firing rates can be derived by taking the inverse Fourier transform of Eq. 5 to get

$$\nu_\alpha(x) = (1 - p_\alpha)\bar{\nu}_\alpha + p_\alpha \bar{\nu}_\alpha g(x; x_f, \sqrt{\sigma_f^2 - \sigma_\alpha^2}), \tag{6}$$

where $\bar{\nu}_\alpha$ is the spatial mean of the firing rate. Synaptic weights $j_{\alpha\beta}$ and input values $\bar{f}_\alpha$ that achieve network balance can be generated through numerical means (see Table II in Methods). Spiking neural network simulations using balanced parameters are shown in Fig. 1e.

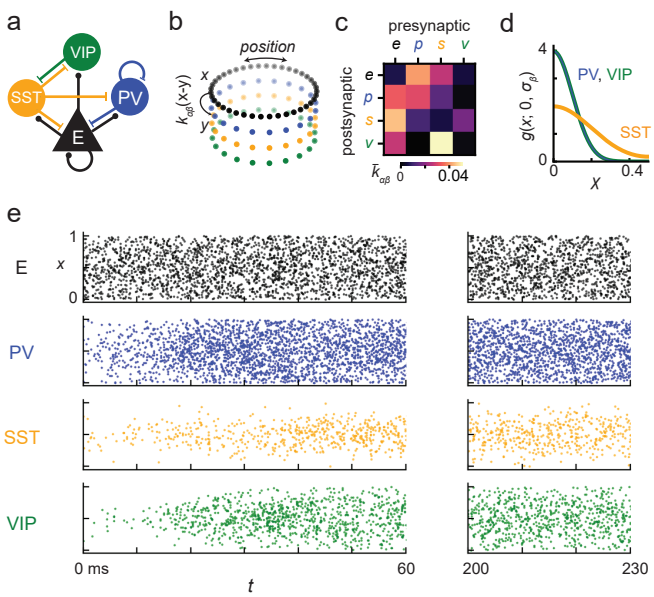

FIG. 1. Model schematic and spiking dynamics. (a) The cortical network model consists of excitatory (E) neurons and PV, SST, and VIP inhibitory neurons. The schematic shows the major projections within and between cell types. (b) Cartoon of the spatial structure of the network. Neurons are placed with even spacing in a ring configuration. The probability of connection depends on the relative position $x - y$ between the position of the presynaptic and postsynaptic neurons. (c) Cell-type-specific connectivity probabilities. Brighter color represents a greater probability that the corresponding presynaptic cell type will project to the postsynaptic cell type. Indices e, p, s, and v correspond to E, PV, SST, and VIP neurons, respectively. (d) Spatial projection kernel of each inhibitory cell type. (e) Raster plot of example spiking neural network run with a network size of $N = 5 \times 10^5$. The vertical axes represents the position and the horizontal axes the time of the spike.

C. Finite-N solutions

Under mean-field assumptions, firing rates for networks of finite size N can be approximated with a semi-analytic approach. The current input $I_{\alpha,k}$ is replaced by a diffusion term $\mu_\alpha(x) + \sqrt{2D_\alpha(x)}\xi(t)$ where $\mu_\alpha(x)$ is the mean and $D_\alpha(x)$ the variance of the input to neurons of type α, and $\xi(t)$ is a zero-mean white noise. Given $\mu_\alpha(x)$ and $D_\alpha(x)$, the stationary firing rates $\nu_\alpha(x)$ can be computed by numerically integrating the Fokker-Planck equation as introduced in previous works [35, 36]. Since $\mu_\alpha(x)$ and $D_\alpha(x)$ themselves depend on $\nu_\alpha(x)$, this forms self-consistent equations from which an iterative approach can be used to numerically compute the stationary firing rates (details in Methods).

Comparison of the stationary firing rates with LIF network simulations shows that the mean-field description

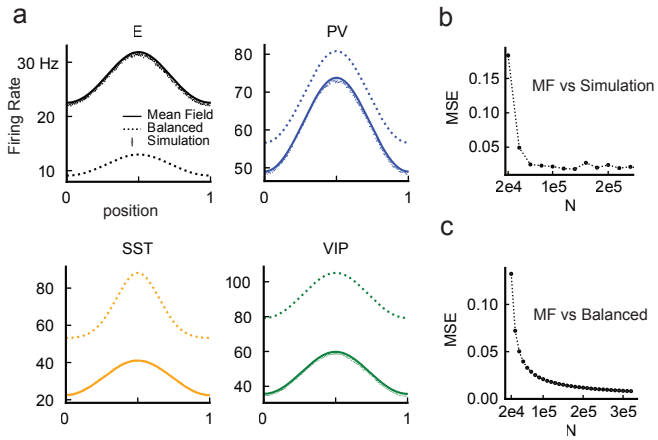

FIG. 2. Firing rates in the four-population (E, PV, SST, and VIP) model. (a) Stationary firing rate from mean-field theory (solid lines), analytical solutions in the thermodynamic limit (dashed lines), and firing rates calculated from simulations (error bars corresponding to mean ± sem over 10 independent network initializations) from spiking networks with $N = 5 \times 10^5$. (b) Normalized mean square error (NMSE) between mean-field theory and simulated firing rates as a function of network size N. (c) NMSE between the mean-field model with varied network sizes and the balanced solution in the thermodynamic limit.

accurately predicts the stationary firing rates (Fig. 2a). The normalized mean square error (NMSE) between the mean-field fixed points and rates computed from simulations decreases with network size (Fig. 2b). Furthermore, we see that as N increases the network converges to the analytical solutions satisfying the balanced state in the thermodynamic limit in Eq. 6 (Fig. 2c).

D. Stability Analysis

One of the key features we include in the model is the long-range projections of SST neurons, i.e. $\sigma_s > \sigma_p, \sigma_v$. This is based on anatomical evidence as well as fact that long-range SST projections are necessary in spiking neural network models to reproduce the divisive effects of lateral inhibition by distal SST neurons [27]. However, previous work on spatially distributed E-I networks with a single type of inhibitory population showed that inhibitory projections must be strictly shorter than excitatory projections ($\sigma_i < \sigma_e$) to maintain a stable balanced state [13, 14]. Here we examine whether long-range SST projections in a network composed of different inhibitory subtypes cause instability in our proposed model.

To do this, we linearize the firing rate dynamics about the fixed point,

$$\tau_\alpha \frac{\partial \nu_\alpha}{\partial t} = -\nu_\alpha + \gamma \mu_\alpha \Theta(\mu_\alpha). \tag{7}$$

where $\gamma = 1$ is the gain of the neuron and Θ is the Heaviside step function. Under this approximation, the fixed point is stable whenever the matrix $C(n) = \tilde{W} - \epsilon I$ (where $\epsilon = (\gamma\sqrt{N})^{-1}$ and I is the identity matrix) has eigenvalues λ_n with negative real parts $\mathrm{Re}(\lambda_n)$ for each Fourier mode n.

We first consider the stability of a network of size $N = 10^8$ with a homogeneous inhibitory neuron population by setting $q_i \equiv q_p = 0.2$ and $q_s = q_v = 0$. This is similar to the model in Ref. [13], but with inhibitory neuron parameters specifically derived from PV neurons. We computed the maximum real part of the eigenvalues for varying excitatory and inhibitory projection range σ_e and σ_i (Fig. 3a). The dashed line indicates the boundary between stable and unstable networks. When $\sigma_e = 0.1$ and $\sigma_i = 0.2$ (* on Fig. 3a), the network lies in the unstable region.

We then considered the stability diagram of the network with heterogeneous inhibitory neurons and vary σ_s along with σ_e while keeping $\sigma_p = \sigma_v = 0.1$. The boundary of stability has a significantly different profile. When $\sigma_s = 0.2$ and $\sigma_e = 0.1$ (* on Fig. 3b), stability is maintained.

Increasing N for both networks increases the range of instability (Fig. 3c). For the homogeneous model, the stability-instability boundary trends towards the diagonal line $\sigma_e = \sigma_i$. This indicates that as $N \to \infty$, the network is unstable as long as $\sigma_i > \sigma_e$. However, for the heterogeneous network the boundary converges to a horizontal line, and a region exists for which the network remains stable when $\sigma_s > \sigma_e$.

We examined how heterogeneity affects local firing rate readouts. The average firing rates over a local patch of neurons were computed for the networks with homogeneous and heterogeneous inhibitory populations with long-range inhibitory projections (Fig. 3d). To see how the rates fluctuate about its mean value, we normalized the firing rates by dividing them by the respective time averages. The excitatory neurons in the stable heterogeneous network (Fig. 3d, top) closely tracked the time-averaged value with deviations under 1% of the average (black line). The unstable homogeneous network var-

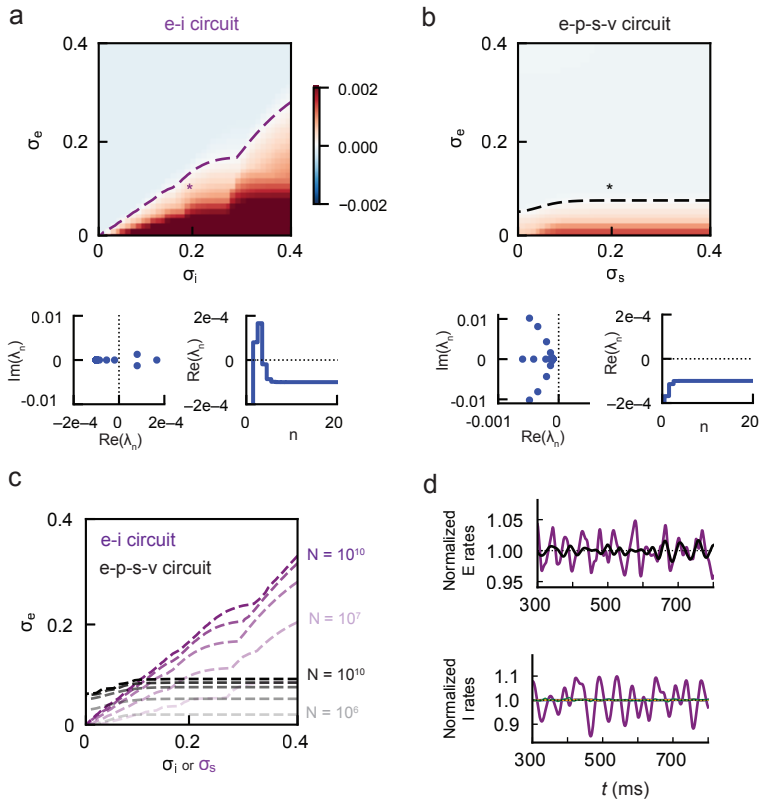

FIG. 3. Stability with long-range inhibitory projections with heterogeneous inhibitory populations. (a) Top: Stability diagram of an E-I circuit with a homogeneous inhibitory neuron population consisting of PV neurons only ($N = 10^8$). The color indicates the maximum value of $\mathrm{Re}(\lambda_n)$ for the corresponding inhibitory projection distance σ_i and excitatory projection distance σ_e. Red indicates positive eigenvalues (unstable) and blue negative values (stable). The purple dashed line traces the border of stability. Bottom: Eigenspectrum λ_n on the complex plane (left) and $\mathrm{Re}(\lambda_n)$ as a function of the Fourier node n when $\sigma_e = 0.1$ and $\sigma_i = 0.2$, indicated with an asterisk on top panel. (b) Same as A, but with heterogeneous inhibitory neuron populations (PV-SST-VIP). Black dashed line indicates stability boundary. (c) Stability border for the two circuit models with varying network size N. (d) Firing rates normalized against the time average ($\int_\Delta \nu(x,t)dx / \int\int_\Delta \nu(x,t)dxdt$, dimensionless units) in a patch of neurons (specifically, neurons in the range $x \in \Delta = [0.2, 0.25]$) from the networks with projection profiles marked with asterisks in (a) and (b). Top panel shows readouts from E neurons in E-I (purple) and E-PV-SST-VIP (black) networks, respectively. Bottom panel shows inhibitory neuron readouts from the homogeneous inhibitory population in the E-I network (purple) and from each inhibitory cell type in the heterogeneous network (blue, gold, and green). Firing rates are obtained by convolving the spike train with a Gaussian with $\sigma = 5$ ms.

ied up to $\sim 5\%$ of the average (purple line). The firing rates of inhibitory neurons followed a similar trend, exhibiting a greater degree of time-varying responses in the network with a homogeneous inhibitory population than in the network with with inhibitory subtypes (Fig. 3d, bottom).

E. Cell-type-specific gain modulations

Next, we examined the gain modulations induced by stimulating specific cell types in the mean-field model. We define gain modulations as changes in the slope of the visual response curves in excitatory neurons: a decrease in the slope indicates a divisive modulation, while a vertical shift without slope change indicates a purely subtractive modulation [21, 25, 27]. We presented a Gaussian-shaped input current to excitatory and PV neurons centered at $x = 0.5$ with varying intensities to simulate visual inputs of different contrast levels. Cell-type-specific modulatory signals activating inhibitory neurons were induced with an additional Gaussian-shaped input current centered at x_C which was either proximal ($x_C = 0.5$) or distal ($x_C = 0$) to the site of the visual stimulus input (Fig. 4a; details in Methods). The spatial structure of the E response with PV (blue), SST (gold) and no stimulation (black) showed distinct shapes with proximal stimulations (Fig. 4b, left). However, the slope of excitatory responses to varying stimulus intensity did not decrease when paired with stimulation of any types of inhibitory neurons, i.e. no divisive gain changes were observed (Fig. 4b, right). Different spatial response profiles were

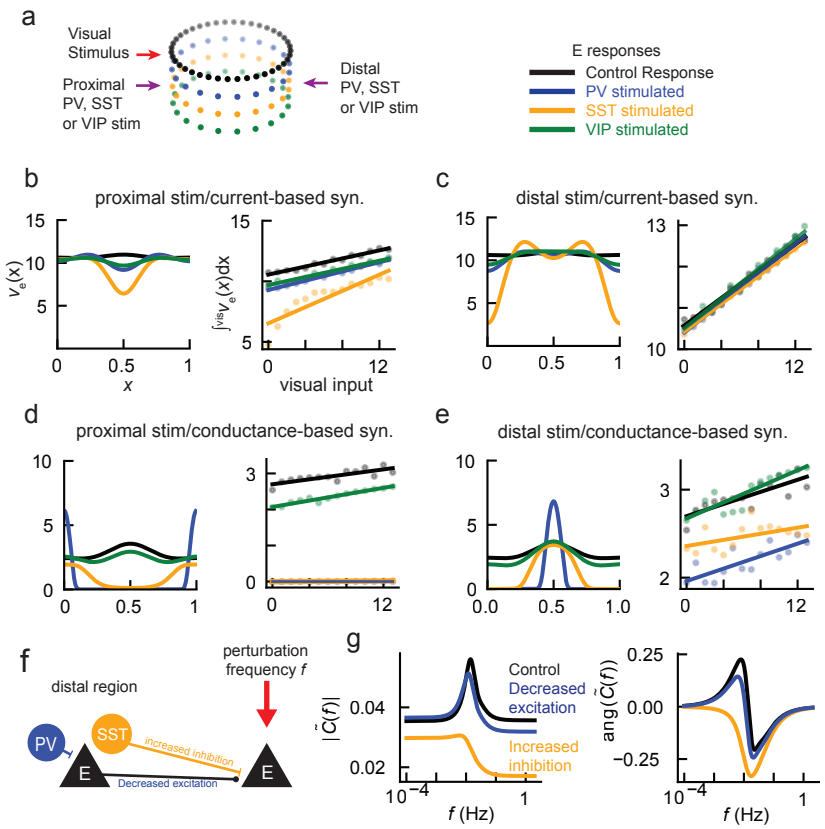

FIG. 4. Cell-type-specific gain modulations. (a) Schematic of stimulations. Visual stimulus is given as input to E and PV neurons centered at $x = 0.5$. For proximal modulations, PV, SST, or VIP neurons at $x = 0.5$ were stimulated. For distal modulations, neurons at $x = 0$ were stimulated. (b) Left: Response profiles of E neurons in the current-based model with PV (blue), SST (gold), VIP (green) proximal stimulation and without activation of inhibitory neurons (black). Right: E responses averaged across the region of visual input as a function of the strength of the visual stimulus. (c) Same as (b) but for distal PV, SST, and VIP stimulation. (d, e) Same as (b, c) but with the conductance-based model. (f) Schematic of linear response. Oscillatory perturbation is applied to the visual input. The linear response with increased long-range inhibition (gold) and decreased long-range excitation (blue) are computed. (g) Linear response amplitude (left) and phase shift (right).

also observed with distal stimulation of inhibitory subtypes, but again, with no divisive modulations (Fig. 4c).

Experimental evidence suggests that both PV and SST neurons can exert divisive gain modulations (i.e., negative slope changes) depending on the stimulation conditions and the location of their stimulation relative to the site of visual input [22–24, 27]. We reasoned that the lack of divisive gain modulations in the model may be due to simplifications in the model, in particular due to the current-based synapses assumed in the model. This is manifested in the balanced solution Eq. 5, which is linear in external inputs $\tilde{f}_\beta$. We thus incorporated conductance-based synapses into the model; mean-field firing rates can be computed with a similar approach to the current-based case with modified parameters and a voltage-dependent diffusion term (see Methods). We first verified that the mean-field results with the conductance-based synapses agreed with the corresponding spiking network simulations (Supplementary Fig. 1). We repeated the stability analysis across spatial scales of excitatory and inhibitory projections to account for voltage-dependent synaptic transmissions and found similar results to the current-based model (Supplementary Fig. 2).

Having verified that the conductance-based mean-field model generates the same results on spatial response profiles (Supplementary Fig. 1) and projection-dependent stability (Supplementary Fig. 2) as with the current-based synapses model, we next simulated proximal and distal activations of inhibitory neurons using the mean-field model with conductance-based synapses. When proximal stimulation was given to the conductance-based network, again distinct profiles emerged—PV and SST stimulation completely suppressed excitatory responses at the site of visual input, while VIP stimulation exerted a subtractive effect (Fig. 4d). With distal stimulation, the spatial profile of the excitatory response sharpened with PV stimulation. In addition, distal PV stimulation caused a subtractive effect on the E response to visual input strengths (no slope change), while distal SST stimulation caused a divisive gain change (decrease in slope). Distal VIP stimulation caused slight multiplicative gain change (increase in slope, Fig. 4e). These distance- and

cell-type dependent modulations are consistent with experimentally observed effects of both proximal and distal PV or SST stimulation in mouse V1: both resulted in strong divisive gain modulations in excitatory neurons when the activation of the inhibitory population was proximal, and only SST activation resulted in divisive operations when distal [27]. The positive gain change induced by VIP neurons is also consistent with a previous experimental study [30].

We turned to linear response theory to clarify the mechanisms of the long-range effects of PV and SST stimulation. Distal SST neurons, due to their long-range projections, can directly inhibit E neurons in the visual field of stimuli. On the other hand, distal PV neurons primarily act on E neurons in the visual field polysynaptically by inhibiting presynaptic E neurons, thereby decreasing excitation (Fig. 4f). To test effects of these distinct mechanisms, we analyzed the response of an E neuron at the center of the visual field ($x = 0.5$) to a frequency-dependent perturbation to the visual input (details in Methods). We examined the amplitude and angle of the response with a 2-fold increased inhibition (simulating effects of distal SST stimulation) and 10-fold decreased excitation (simulating effects of distal PV stimulation), and compared them to the base case (Fig. 4g). The amplitude of the response in the control case showed a peak at about 10^{-2} Hz, indicating a resonant frequency, and there was an associated positive phase shift (black curve, Fig. 4g). When excitation was decreased, the response was similar to the control case, exhibiting the same resonant frequency and phase shift (blue curve). However, when inhibition was increased, the amplitude of the response was shifted downward, indicating a lower gain (gold curve). Moreover, the resonant frequency disappeared with increased inhibition. These results indicate that increased inhibition induces a negative gain change in the excitatory neurons across all stimulus frequencies, which may act as a mechanism for divisive effects of distal SST stimulation. Distal PV stimulation decreases excitation without increasing direct inhibition, resulting in a subtractive effect. Meanwhile, a strong local stimulation of either PV or SST can induce divisive operations through direct inhibition.

III. DISCUSSION

We derived mean-field equations for a spatially-organized spiking neural network with excitatory and three inhibitory sub-populations. The neurons have distance-dependent connectivity rules that depend on pre- and post-synaptic cell types, informed by anatomical data from mouse primary visual cortex. Given the cell-type specific connectivity, we derived conditions for E-I balance in the network. The model revealed dynamical properties conferred by inhibitory heterogeneity that may indicate computational benefits. The heterogeneous network maintained stable dynamics even with long-range SST inhibition; this was in contrast to the network with homogeneous inhibitory neurons, whose projections were restricted to be shorter than those of excitatory neurons in order to maintain stability as suggested by a previous theoretical study [13]. Thus, our result, by introducing inhibitory subtypes, closes the gap between the previous theoretical prediction where inhibitory projections are restricted to be local [13] and contradictory experimental observations of long-range inhibition [27–29]. Furthermore, activation of each inhibitory cell type performed distinct modulations on excitatory neurons processing visual input, in agreement with recent experimental observations [24, 27, 30]. This demonstrates that there are multiple modes of computations that can be activated flexibly depending on spatial relation between inhibitory modulatory signals and visual stimuli.

Our work sheds light on the computational role of neuronal cell type diversity and their distinct spatial connectivity in the cortex. The balance of excitation and inhibition in the cortex is believed to be the basis for irregular spiking activity and heterogeneity of responses [8]. By extending previous work on spatially-distributed E-I networks [13, 15], we demonstrated that the network composed of excitatory and three main subtypes of inhibitory neurons—PV, SST, and VIP—can achieve balance to each of its populations. Our analytical findings give us a simple description to determine balance in a multi-population network depending on connectivity and synaptic weights. Understanding how balanced states can be altered by changes in connectivity may be important in understanding network activity in neurological disorders [37–39].

Our mean-field approach further allowed us to directly compute the eigenspectrum of the network's fixed state. Our model shows that neural heterogeneity enhances network stability, allowing stable firing rate readouts even in the presence of long-range inhibition. In contrast, networks with excitatory and homogeneous inhibitory populations displayed metastable to unstable dynamics and high response variability with long-range inhibition. This result thus implicates possible computational advantages of cell type diversity.

Identifying neuronal cell types and their functional roles has emerged as a major problem in neuroscience [17, 40, 41]. Computational studies have played an indispensable role in understanding the role of inhibitory neuron subtypes in dynamics and computations [18–20, 31, 42–44]. Our work complements these studies by providing additional mechanisms by which networks may perform divisive and subtractive computations. For example, Wilmes et al. [20] showed that divisive and subtractive operations on excitatory rates by PV and SST neurons, respectively, may account for computations needed to encode the variance and mean of a variable. Our work suggests that PV and SST neurons can perform various degrees of divisive or subtractive modulations depending on the distance between their activation and the stimulus input sites. Thus, inhibitory neurons, by utilizing

their heterogeneous projection patterns across the cortical network, may perform a range of distinct operations on spatially organized stimuli in an effective and robust manner.

Our results highlight the computational benefits of a heterogeneous neuronal population in recurrent neural networks. A recent line of work has found that the degree of heterogeneity within a cell type can play a major role in network computations [43, 44]. While our work focused on heterogeneity in intrinsic properties and projection patterns across cell types and assumed no variation in parameters within each cell type, it would be interesting to investigate the interaction of within- and between-population heterogeneity. In another work, Winston et al. found that heterogeneous cell types emerge when training spiking neural networks to perform temporal processing tasks [42]. It is an open question whether PV, SST, and VIP-like structures may emerge from network optimization under certain objective functions and constraints.

In future studies, the current work can be expanded in several ways. First, we will investigate how the network can take advantage of the computational properties explored here by employing them in visual processing tasks. In addition, further biophysical details neglected in the present model, such as dendritic integration [27], can be incorporated to make the mean-field model more realistic.

IV. METHODS

A. Neuronal dynamics and spatial projections

The membrane potential $V_{\alpha,k}$ of the kth neuron of type α ($= e$, p, s, or v) is given by leaky integrate-and-fire (LIF) equations

$$\tau_\alpha \frac{dV_{\alpha,k}}{dt} = g_{L,\alpha}(E_L - V_{\alpha,k}) + I_{\alpha,k}(t), \tag{8}$$

where τ_α is the membrane time constant, set to 30 ms for $\alpha = e$ and 20 ms otherwise; $g_{L,\alpha}$ is the leak conductance, set uniformly to 1 (dimensionless units); E_L is the leak reversal potential at 0; and $I_{\alpha,k}(t)$ is the input current, given by

$$I_{\alpha,k}(t) = \sum_{\beta\in\{e,p,s,v\}} \sum_{j=1}^{q_\beta N} J^{k,j}_{\alpha\beta} s_{\beta,j}(t) + F_\alpha(x). \tag{9}$$

Here $J^{k,j}_{\alpha\beta}$ is the synaptic weight from presynaptic neuron j of type β, $s_{\beta,j}(t) = \sum_{i'} \delta(t - t^{i'}_{\beta,j})$ is its spike train, $x = k/N_\alpha$ is the position of the kth neuron, and $F_\alpha(x)$ is a static, position-dependent external input. The membrane potential variable V and voltage-related parameters are given in dimensionless units.

TABLE I. Cell-type-specific connectivity probabilities $\bar{k}_{\alpha\beta}$ used in the model.

		presynaptic cell type β			
		E	PV	SST	VIP
postsynaptic	E	0.06	0.35	0.23	0.05
cell type	PV	0.42	0.39	0.18	0.02
α	SST	0.35	0.09	0.05	0.14
	VIP	0.16	0.00	0.30	0.01

TABLE II. Synaptic weights $j_{\alpha\beta}$ and external current input parameters $\bar{f}_\alpha = \bar{F}_\alpha/\sqrt{N}$ and p_α used in Results. Values of $j_{\alpha\beta}$ and $\bar{f}_\alpha$ were chosen to achieve balance in the thermodynamic limit as described in the Methods.

		presynaptic cell type β				external	
		E	PV	SST	VIP	$\bar{f}_\alpha$	p_α
postsynaptic	E	0.34	1.34	1.13	2.00	5×10^{-4}	0.25
cell type	PV	0.60	0.77	0.20	2.00	1×10^{-5}	0.25
α	SST	0.28	1.52	2.00	0.20	7×10^{-5}	0.2
	VIP	0.51	0.77	0.68	2.00	1×10^{-5}	0.2

When the membrane potential reaches the threshold $V_{\rm th}$, a spike is emitted and $V_{\alpha,k}$ is reset to the reset voltage $V_{\rm re}$ on the next time step. We set $V_{\rm th} = 1$ and $V_{\rm re} = 0$; numerical integration is performed using the Euler method with time step $\Delta t = 0.1$ ms

The synaptic weight $J^{k,j}_{\alpha\beta}$ from neuron j of cell type β (at position y) to neuron k of cell type α (at position x) is given by

$$J^{k,j}_{\alpha\beta} = \begin{cases} J_{\alpha\beta}, & \text{with probability } k^\Gamma_{\alpha\beta}(x-y), \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

Here $k^\Gamma_{\alpha\beta}(x-y)$ is the distance-dependent connectivity kernel.

The synaptic weight $J_{\alpha\beta}$ is assumed to scale with the network size as $J_{\alpha\beta} = j_{\alpha\beta}/\sqrt{N}$ where $j_{\alpha\beta}$ is a constant independent of N. External inputs scale according to $F_\alpha(x) = \sqrt{N} f_\alpha(x)$, while other parameters are constant with respect to network size. The wrapped Gaussian projection kernel is given by

$$g(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} \sum_{n=-\infty}^{\infty} e^{-(x+n-\mu)^2/2\sigma^2}. \tag{11}$$

Current inputs are given by

$$f_\alpha(x) = p_\alpha \bar{f}_\alpha g(x; x_f, \sigma_f) + (1-p_\alpha)\bar{f}_\alpha. \tag{12}$$

Here $\bar{f}_\alpha$ is the mean input to cells of type α and p_α determines the proportion of the input spatially localized at x_f.

B. Mean-field equations in the continuum limit

In the mean-field model, we take the continuum limit

$$\tau_\alpha \frac{dV_\alpha}{dt}(x) = g_{L,\alpha}(E_L - V_\alpha(x)) + I_\alpha(x,t) \tag{13}$$

where the input current from Eq. 8 is replaced by a position-dependent effective field, modeled as a random process $I_\alpha(x,t) = \mu_\alpha(x) + \sqrt{2D_\alpha(x)}\xi(t)$ where $\mu_\alpha(x)$ is the mean input current (given by Eq. 2) and

$$D_\alpha(x) \approx \frac{1}{2}\lim_{\delta\to 0}\delta^{-1}\left\langle E\left[\int_t^{t+\delta} I_\alpha(x,s) - \mu_\alpha(x)ds\right]\right\rangle = \sum_\beta \frac{j_{\alpha\beta}}{2}(w_{\alpha\beta} * \nu_\beta)(x). \quad (14)$$

is the input variance, and $\xi(t)$ is zero-mean delta-correlated white noise. Here $w_{\alpha\beta}(x) \equiv q_\beta j_{\alpha\beta} k^\Gamma_{\alpha\beta}(x)$ is the distance (x)-dependent weighted connectivity, $\nu_\alpha(x) \equiv \langle E[s_{\alpha,k}]\rangle$ is the population- and position (x)-dependent mean firing rate, and $*$ indicates circular convolution, i.e. $w_{\alpha\beta} \equiv \int_\Gamma w_{\alpha\beta}(x-x')v_\beta(x)dx'$.

Given input mean and variance, the firing rate of neurons of type α in position x is computed using a Fokker-Planck approach [35, 36]. For a given α and x, the variables $p_{0,\alpha}(x,V) = P_{0,\alpha}(x,V)/\nu_{0,\alpha}(x)$ and $j_{0,\alpha}(x,V) = J_{0,\alpha}(x,V)/\nu_{0,\alpha}(x)$, where $P_{0,\alpha}(x,V)$, $J_{0,\alpha}(x,V)$, and $\nu_{0,\alpha}(x)$ are the stationary probability distribution of the membrane potential, the current operator, and the firing rate, respectively, are described by

$$\frac{\partial p_{0,\alpha}}{\partial V}(x,V) = \frac{1}{D_\alpha(x)}(M_{0,\alpha}(x,V)p_{0,\alpha}(x,V) - j_{0,\alpha}(x,V)), \quad (15)$$

$$\frac{\partial j_{0,\alpha}}{\partial V}(x,V) = \delta(V - V_{\rm re}). \quad (16)$$

Here $M_{0,\alpha}(x,V) = \tau_\alpha^{-1}[g_{L,\alpha}(E_L - V) + \mu_\alpha(x)]$ is the drift coefficient. We integrate Eqs. 15, 16 using a numerical integration scheme previously introduced [35], specifically using Simpson's rule [36]. The firing rate is then given by

$$\nu_{0,\alpha}(x) = \left[\int_{V_{\rm lb}}^{V_{\rm th}} p_{0,\alpha}(x,V)dV\right]^{-1}, \quad (17)$$

where the membrane potential lower bound $V_{\rm lb}$ is selected so that $P_{0,\alpha}(x,V_{\rm lb})$ is negligible (here set to $V_{\rm lb} = -1$).

This numerical integration scheme defines a mapping Φ from the input mean $\mu_\alpha(x)$ and variance $D_\alpha(x)$ to the position-dependent firing rate, defining a set of self-consistent equations $\nu_\alpha(x) = \Phi(\mu_\alpha(x), D_\alpha(x))$. Fixed points are found using an iterative scheme as described in Ref. [36]. Starting from an initial guess $\nu_\alpha^{(0)}(x)$ for each α, $\mu_\alpha^{(0)}(x)$ and $D_\alpha^{(0)}(x)$ are computed using Eq. 2 and 14 and the next guess is computed iteratively,

$$\nu_\alpha^{(j)}(x) = \Phi(\mu_\alpha^{(j-1)}(x), D_\alpha^{(j-1)}(x)), \quad (18)$$

until the L2 deviation between $\nu_\alpha^{(j)}(x)$ and $\nu_\alpha^{(j-1)}(x)$ is less than a threshold, which we set to 10^{-6} Hz.

C. Balanced state in the thermodynamic limit

Theories of randomly connected neural networks sought to explain asynchronous spiking activity with models of a balanced state in which excitatory and inhibitory inputs dynamically achieve balance [34]. In this state, firing rates of neurons can be expressed in analytic form which becomes exact in the large-N limit. Here we generalize the approach by Rosenbaum and Doiron [13] which implemented two populations—one excitatory and one inhibitory—for the four-population model where the inhibitory population is split into three subtypes.

The formulation described above gives $O(1)$ distance between the resting potential and spike threshold and $J_{\alpha\beta} \sim O(1/\sqrt{N})$, $q \sim O(1)$, $k_{\alpha\beta} \sim O(1)$, and $f_\alpha(x) \sim O(\sqrt{N})$ as $N \to \infty$. Under these conditions, a neuron receives recurent input from $O(N)$ excitatory neurons but requires only $O(\sqrt{N})$ excitatory inputs to be active in the integration window to produce a spike. Finite, unsaturated firing rates are achieved in the continuum limit through a dynamically stable balance between excitation and inhibition.

The stationary firing rate solutions in the Fourier domain are given by Eq. 5 and the derivation is detailed in Section II B. To constrain the firing rates to be non-negative, we need

$$\mathrm{sgn}\left(\sum_\beta \tilde{f}_\beta A_{\alpha\beta}\right) = \mathrm{sgn}\left(\det[-\tilde{W}]\right). \quad (19)$$

In the "two-population model" composed of an excitatory population and a homogeneous inhibitory population [13], this condition requires that the external inputs must be broader than recurrent connections for the balanced solution to exist. However, in the "four-population model" composed of one excitatory and three distinct inhibitory populations, not every inhibitory neuronal population's projections need to be narrower than those of the excitatory population.

Synaptic weights $j_{\alpha\beta}$ and external input intensities $\bar{f}_\alpha$ that achieve balance are generated numerically as follows. Synaptic weights are constrained to be in the range $[0.2, 2]$ and external inputs in the range $[10^{-5}, 10^{-1}]$. Initial values are generated randomly within the bounds and then the parameters are adjusted to minimize the mean square deviation between $\bar{\nu}_\alpha$ and target firing rates set at 50 Hz using the Sequential Least Squares Programming algorithm. If the resulting firing rates are all positive, the parameters are accepted. One such set of synaptic weights and external input parameters was used in the simulations and are listed in Table II.

D. Conductance-based model

When simulating conductance-based neurons, Eq. 9 becomes

$$I_{\alpha,k}(t) = \sum_{\beta\in\{e,p,s,v\}} \sum_{j=1}^{q_\beta N} J^{k,j}_{\alpha\beta}(E_\beta - V)s_{\beta,j}(t) + F_\alpha(x), \tag{20}$$

where E_β is the reversal potential, set to $E_e = 1.5$ for excitatory and $E_\beta = E_i = -0.5$ for inhibitory presynaptic neurons. For conductance-based equations, we replace E_L in Eq. 15 with the effective leak voltage $E'_{L,\alpha}$ which is dependent on cell type [35],

$$E'_{L,\alpha}(x) = \frac{E_L + \tau_\alpha \sum_\beta R_{\alpha\beta}(x)(b_{\alpha\beta} + b^2_{\alpha\beta})E_\beta}{1 + \tau_\alpha \sum_\beta R_{\alpha\beta}(x)(b_{\alpha\beta} + b^2_{\alpha\beta})} \tag{21}$$

and the membrane conductance $g_{L,\alpha}$ with

$$g'_{L,\alpha} = g_{L,\alpha}\left(1 + \tau_\alpha \sum_\beta R_{\alpha\beta}(x)(b_{\alpha\beta} + b^2_{\alpha\beta})\right) \tag{22}$$

where $b_{\alpha\beta} = 1 - e^{-J_{\alpha\beta}}$ and $R_{\alpha\beta}$ is the input spike rate from population β, given by

$$R_{\alpha\beta}(x) = Nq_\beta(k^\Gamma_{\alpha\beta} * \nu_\beta)(x). \tag{23}$$

We use the same neuronal and synaptic parameters $g_{L,\alpha}$, τ_α, and $j_{\alpha\beta}$ as in the current-based case (Table II). We integrate Eqs. 15 and 16 as before, now with a voltage-dependent diffusion

$$D_\alpha(x, V) = \frac{1}{2}\left(\sum_\beta R_{\alpha\beta}(x)b^2_{\alpha\beta}(V - E_\beta)^2\right). \tag{24}$$

Due to inputs being voltage-dependent, the stability matrix $C(n) = \tilde{W} - \epsilon I$, where $\tilde{W}$ is the matrix with Fourier components $\tilde{w}_{\alpha\beta}(n)$, is also modified. We make the simplifying assumption that the mean membrane potential is independent of x, yielding

$$C(n) = \begin{pmatrix} -\epsilon + \tilde{w}_{ee}v^E_e & -\tilde{w}_{ep}v^I_e & -\tilde{w}_{es}v^I_e & -\tilde{w}_{ev}v^I_e \\ \tilde{w}_{pe}v^E_p & -\epsilon - \tilde{w}_{pp}v^I_p & -\tilde{w}_{ps}v^I_p & -\tilde{w}_{pv}v^I_p \\ \tilde{w}_{se}v^E_s & -\tilde{w}_{sp}v^I_s & \epsilon - \tilde{w}_{ss}v^I_s & -\tilde{w}_{sv}v^I_s \\ \tilde{w}_{ve}v^E_v & -\tilde{w}_{vp}v^I_v & -\tilde{w}_{vs}v^I_v & \epsilon - \tilde{w}_{vv}v^I_v \end{pmatrix} \tag{25}$$

where $v^E_\alpha = \bar{V}_\alpha - E_e$, $v^I_\alpha = \bar{V}_\alpha - E_i$, and $\bar{V}_\alpha = \int_\Gamma V_\alpha(x)dx$.

E. Modulation of visual response

Modulations by specific inhibitory neuronal populations are measured by changes in the excitatory neurons' response to a visual stimulus. We first replace the position-dependent external current in Eq. 12 with a uniform input $f_\alpha(x) = \bar{f}_\alpha$, whose values are re-tuned to result in firing rates within 10 Hz, which are the observed spontaneous firing rates in mouse V1 in the absence of visual stimuli [27]. The baseline currents are given by $\bar{f}_e = 1.5 \cdot 10^{-4}$, $\bar{f}_p = 1 \cdot 10^{-5}$, $\bar{f}_s = 7 \cdot 10^{-5}$, and $\bar{f}_v = 6 \cdot 10^{-5}$. Visual stimuli and modulatory signals are given in addition to the baseline input.

The visual input is incorporated by an additional term $I_{v,\alpha}(x)$ to the total input given to the mean input Eqs. 9 and 20. For the current-based model, this is given by

$$I_{v,\alpha}(x) = a_v\bar{v}_\alpha g(x; x_v, \sigma_v) \tag{26}$$

where a_v controls the overall strength of the visual input, $\bar{v}_\alpha$ is a population-dependent term controlling the relative strength of the visual input to population α, $x_v = 0.5$ is the center of the stimulus and $\sigma_v = 0.3$ is the spatial spread of the input. Feedforward visual inputs are known to be fed to E and PV neurons [25], so we set $\bar{v}_e = \bar{f}_e$, $\bar{v}_p = \bar{f}_p$, and $\bar{v}_s = \bar{v}_v = 0$.

For the conductance-based model, the visual stimuli are represented by conductance-based synaptic inputs. This results in an additional term in each of the summations over β in Eqs. 21, 22, 24 to include inputs from an external layer T where we set $R_{\alpha T}(x) = a_v r_\alpha g(x; x_v, \sigma_v)$ for $\alpha \in e, p$ and $R_{\alpha T} = 0$ otherwise. We set $r_e = 4\times10^{-4}$ and $r_p = 1\times10^{-4}$, values tuned to achieve realistic firing rate responses [27]. Synaptic weight parameters are set to $b_{eT} = 1 - e^{-0.2}$ and $b_{pT} = 1 - e^{-0.02}$.

To measure modulations, we first measure the excitatory response of the system to visual stimuli of varying intensity without additional activation of inhibitory neurons. This is done by computing the cell-type and position-dependent stationary firing rates $\nu_\alpha(x)$ for a_v varying from 0 to 12. Then the process is repeated with an additional current input given to a specific inhibitory cell type, similar to optogenetic stimulations in experimental studies [24, 27, 30]. The inhibition-activating stimulus is given by a Gaussian current input centered at x_C, which is set to $x_C = 0.5$ for proximal stimulations (Figs. 4b,d) and 0 for distal stimulations (Figs. 4c,e). The spatial spread of the inhibition-activating input is given by the standard deviation $\sigma_C = 0.2$ and the magnitude is set equal to $\bar{f}_\alpha$, i.e., $f_C(x) = \bar{f}_\alpha g(x; x_C, \sigma_C)$.

F. Linear response theory

Linear responses to input perturbations are computed with the methods developed in Refs. [35, 36] and briefly outlined here. Linear responses are measured for excitatory neurons at position $x = 0.5$; cell type indices and positions are suppressed in the following equations. If a weak noise term $\sqrt{2D}\xi(t)$ is added to the visual input $I_v(t)$ (we set $D = 0.1$), the cross-spectral density is given by

$$\tilde{C}(f) = 2DS_\nu(f) \tag{27}$$

where $S_\nu(f)$ is the susceptibility function of the firing rate. We consider a perturbation to the visual input

$$I_v(t) = I_{v,0}(t) + \varepsilon e^{2\pi i f t} \tag{28}$$

where $I_{v,0}(t)$ is the unperturbed input to the neuron, f is the frequency of the modulation and ε is a small parameter controlling the amplitude of the modulation. Setting $\varepsilon = 0$ recovers the fixed-point solution. To first order in ε, the response of the neuron is periodic with frequency f,

$$\begin{aligned} P(V,t) &= P_0(V) + \varepsilon P_1(V) e^{2\pi i f t} \\ \nu(t) &= \nu_0 + \varepsilon S_\nu(f) e^{2\pi i f t} \end{aligned}$$

where $P_0(V)$ and ν_0 are the steady-state probability density function and the firing rate, respectively, and $P_1(V)$ and $S_\nu(f)$ are the susceptibility functions. From here a time-dependent Fokker-Plank equation for the probability density $P(V,t)$ can be derived with a drift coefficient

$$M(V,t) = M_0(V) + \varepsilon M_1 e^{2\pi i f t} \tag{29}$$

where $M_0(V)$ is the time-independent drift coefficient in Eq. 15 and M_1 is the time-dependent drift coefficient, given by $M_1 = g_{L,\alpha}\tau_\alpha$. We can decompose the probability density and flux functions as $P_1(V) = S_\nu p_\nu(V) + p_1(V)$ and $J_1(V) = S_\nu j_\nu(V) + j_1(V)$ which satisfy the equations

$$\begin{aligned} -\frac{\partial p_1}{\partial V} &= -\frac{1}{D} M_0 p_1 + \frac{1}{D}[j_1 - M_1 p_0] \\ -\frac{\partial j_1}{\partial V} &= 2\pi i f p_1 \\ -\frac{\partial p_\nu}{\partial V} &= -\frac{1}{D} M_0 p_\nu + \frac{1}{D} j_\nu \\ -\frac{\partial j_\nu}{\partial V} &= 2\pi i f p_\nu - \delta(V - V_{\mathrm{re}}) \end{aligned} \tag{30}$$

with boundary conditions $p_\nu(V_{\mathrm{th}}) = p_1(V_{\mathrm{th}}) = j_1(V_{\mathrm{th}}) = 0$ and $j_\nu(V_{\mathrm{th}}) = 1$. We integrate these equations with Simpson's rule to obtain the susceptibility,

$$S_\nu = -\frac{j_1(V_{\mathrm{lb}})}{j_\nu(V_{\mathrm{lb}})} \tag{31}$$

where V_{lb} is the lower bound of the membrane potential, set to -1. The cross-spectral density is then computed using Eq. 27.

Data, Materials, and Software Availability

All code used in this study will be made available upon publication on GitHub at https://github.com/HChoiLab/CellTypeMeanField.

ACKNOWLEDGMENTS

This work was supported by the National Eye Institute of the National Institutes of Health under Award Number R00 EY030840. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

[1] B. Haider, A. Duque, A. R. Hasenstaub, and D. A. McCormick, Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition, Journal of Neuroscience **26**, 4535 (2006).
[2] M. Okun and I. Lampl, Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities, Nature neuroscience **11**, 535 (2008).
[3] B. V. Atallah and M. Scanziani, Instantaneous modulation of gamma oscillation frequency by balancing excitation with inhibition, Neuron **62**, 566 (2009).
[4] T. P. Carvalho and D. V. Buonomano, Differential effects of excitatory and inhibitory plasticity on synaptically driven neuronal input-output functions, Neuron **61**, 774 (2009).
[5] J. A. Cardin, Inhibitory interneurons regulate temporal precision and correlations in cortical circuits, Trends in neurosciences **41**, 689 (2018).
[6] J. S. Isaacson and M. Scanziani, How inhibition shapes cortical activity, Neuron **72**, 231 (2011).
[7] K. A. Ferguson and J. A. Cardin, Mechanisms underlying gain modulation in the cortex, Nature Reviews Neuroscience **21**, 80 (2020).
[8] C. Van Vreeswijk and H. Sompolinsky, Chaos in neuronal networks with balanced excitatory and inhibitory activity, Science **274**, 1724 (1996).
[9] M. Ercsey-Ravasz, N. T. Markov, C. Lamy, D. C. Van Essen, K. Knoblauch, Z. Toroczkai, and H. Kennedy, A predictive network model of cerebral cortical connectivity based on a distance rule, Neuron **80**, 184 (2013).
[10] Petilla terminology: nomenclature of features of gabaergic interneurons of the cerebral cortex, Nature Reviews Neuroscience **9**, 557 (2008).
[11] R. Tremblay, S. Lee, and B. Rudy, Gabaergic interneurons in the neocortex: from cellular properties to circuits,

Neuron **91**, 260 (2016).

[12] N. T. Markov, M. Ercsey-Ravasz, C. Lamy, A. R. Ribeiro Gomes, L. Magrou, P. Misery, P. Giroud, P. Barone, C. Dehay, Z. Toroczkai, *et al.*, The role of long-range connections on the specificity of the macaque interareal cortical network, Proceedings of the National Academy of Sciences **110**, 5187 (2013).

[13] R. Rosenbaum and B. Doiron, Balanced networks of spiking neurons with spatially dependent recurrent connections, Physical Review X **4**, 021039 (2014).

[14] R. Pyle and R. Rosenbaum, Spatiotemporal dynamics and reliable computations in recurrent spiking neural networks, Physical review letters **118**, 018103 (2017).

[15] R. Rosenbaum, M. A. Smith, A. Kohn, J. E. Rubin, and B. Doiron, The spatial structure of correlated neuronal variability, Nature neuroscience **20**, 107 (2017).

[16] C. K. Pfeffer, M. Xue, M. He, Z. J. Huang, and M. Scanziani, Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons, Nature neuroscience **16**, 1068 (2013).

[17] L. Campagnola, S. C. Seeman, T. Chartrand, L. Kim, A. Hoggarth, C. Gamlin, S. Ito, J. Trinh, P. Davoudian, C. Radaelli, *et al.*, Local connectivity and synaptic dynamics in mouse and human neocortex, Science **375**, eabj5861 (2022).

[18] A. Litwin-Kumar, R. Rosenbaum, and B. Doiron, Inhibitory stabilization and visual coding in cortical circuits with multiple interneuron subtypes, Journal of neurophysiology **115**, 1399 (2016).

[19] J. H. Lee, C. Koch, and S. Mihalas, A computational analysis of the function of three inhibitory cell types in contextual visual processing, Frontiers in Computational Neuroscience **11**, 28 (2017).

[20] K. A. Wilmes, M. A. Petrovici, S. Sachidhanandam, and W. Senn, Uncertainty-modulated prediction errors in cortical microcircuits, bioRxiv , 2023 (2023).

[21] R. A. Silver, Neuronal arithmetic, Nature Reviews Neuroscience **11**, 474 (2010).

[22] H. Adesnik, W. Bruns, H. Taniguchi, Z. J. Huang, and M. Scanziani, A neural circuit for spatial summation in visual cortex, Nature **490**, 226 (2012).

[23] N. R. Wilson, C. A. Runyan, F. L. Wang, and M. Sur, Division and subtraction by distinct cortical inhibitory networks in vivo, Nature **488**, 343 (2012).

[24] B. V. Atallah, W. Bruns, M. Carandini, and M. Scanziani, Parvalbumin-expressing interneurons linearly transform cortical responses to visual stimuli, Neuron **73**, 159 (2012).

[25] S. El-Boustani, N. R. Wilson, C. A. Runyan, and M. Sur, El-boustani et al. reply, Nature **508**, E3 (2014).

[26] S.-H. Lee, A. C. Kwan, and Y. Dan, Interneuron subtypes and orientation tuning, Nature **508**, E1 (2014).

[27] J. Del Rosario, S. Coletta, S. H. Kim, Z. Mobille, K. Peelman, B. Williams, A. J. Otsuki, A. D. C. Valerio, K. Worden, L. T. Blanpain, *et al.*, Lateral inhibition in v1 controls neural & perceptual contrast sensitivity, Nature Neuroscience (2025).

[28] F. Scala, D. Kobak, S. Shan, Y. Bernaerts, S. Laturnus, C. R. Cadwell, L. Hartmanis, E. Froudarakis, J. R. Castro, Z. H. Tan, *et al.*, Layer 4 of mouse neocortex differs in cell types and circuit organization between sensory areas, Nature communications **10**, 4174 (2019).

[29] N. W. Gouwens, S. A. Sorensen, F. Baftizadeh, A. Budzillo, B. R. Lee, T. Jarsky, L. Alfiler, K. Baker, E. Barkan, K. Berry, *et al.*, Integrated morphoelectric and transcriptomic classification of cortical gabaergic cells, Cell **183**, 935 (2020).

[30] J. Veit, G. Handy, D. P. Mossing, B. Doiron, and H. Adesnik, Cortical vip neurons locally control the gain but globally control the coherence of gamma band rhythms, Neuron **111**, 405 (2023).

[31] M. M. Edwards, J. E. Rubin, and C. Huang, State modulation in spatial networks with three interneuron subtypes, bioRxiv (2024).

[32] H. K. Kato, S. K. Asinof, and J. S. Isaacson, Network-level control of frequency tuning in auditory cortex, Neuron **95**, 412 (2017).

[33] R. Khajeh, F. Fumarola, and L. Abbott, Sparse balance: Excitatory-inhibitory networks with small bias currents and broadly distributed synaptic weights, PLOS Computational Biology **18**, e1008836 (2022).

[34] C. van Vreeswijk and H. Sompolinsky, Chaotic balanced state in a model of cortical circuits, Neural computation **10**, 1321 (1998).

[35] M. J. Richardson, Firing-rate response of linear and nonlinear integrate-and-fire neurons to modulated current-based and conductance-based synaptic drive, Physical Review E **76**, 021919 (2007).

[36] R. Rosenbaum, A diffusion approximation and numerical methods for adaptive neuron models with stochastic inputs, Frontiers in computational neuroscience **10**, 39 (2016).

[37] C. Kehrer, N. Maziashvili, T. Dugladze, and T. Gloveli, Altered excitatory-inhibitory balance in the nmda-hypofunction model of schizophrenia, Frontiers in molecular neuroscience **1**, 226 (2008).

[38] J.-M. Fritschy, Epilepsy, e/i balance and gabaa receptor plasticity, Frontiers in molecular neuroscience **1**, 201 (2008).

[39] S. B. Nelson and V. Valakh, Excitatory/inhibitory balance and circuit homeostasis in autism spectrum disorders, Neuron **87**, 684 (2015).

[40] H. S. Seung and U. Sümbül, Neuronal cell types and connectivity: lessons from the retina, Neuron **83**, 1262 (2014).

[41] C. Teeter, R. Iyer, V. Menon, N. Gouwens, D. Feng, J. Berg, A. Szafer, N. Cain, H. Zeng, M. Hawrylycz, *et al.*, Generalized leaky integrate-and-fire models classify multiple neuron types, Nature communications **9**, 709 (2018).

[42] C. N. Winston, D. Mastrovito, E. Shea-Brown, and S. Mihalas, Heterogeneity in neuronal dynamics is learned by gradient descent for temporal processing tasks, Neural Computation **35**, 555 (2023).

[43] R. Gast, S. A. Solla, and A. Kennedy, Macroscopic dynamics of neural networks with heterogeneous spiking thresholds, Physical Review E **107**, 024306 (2023).

[44] R. Gast, S. A. Solla, and A. Kennedy, Neural heterogeneity controls computations in spiking neural networks, Proceedings of the National Academy of Sciences **121**, e2311885121 (2024).

Supplementary Information: Inhibitory cell type heterogeneity in a spatially structured mean-field model of V1

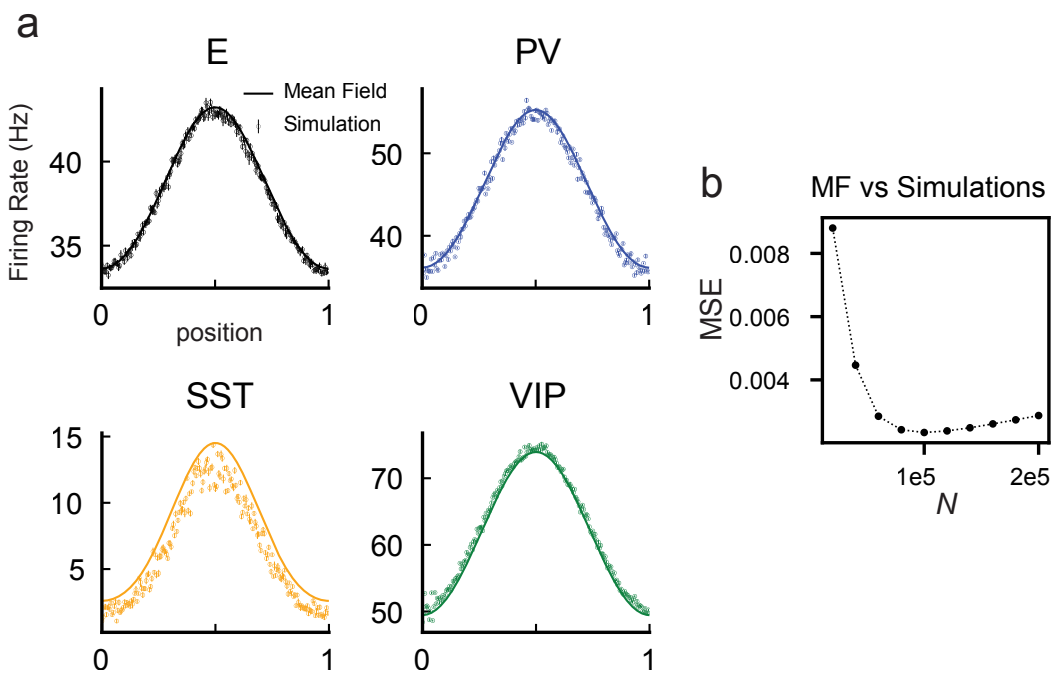

Figure 1: Firing rates in the four-population model with conductance-based synapses. (a) Stationary firing rate from mean field theory (solid lines) and firing rates calculated from simulations (error bars corresponding to mean ± sem) with spiking networks of $N = 5 \times 10^5$. (b) Mean square error (MSE) between mean field theory and simulation firing rates as a function of network size N.

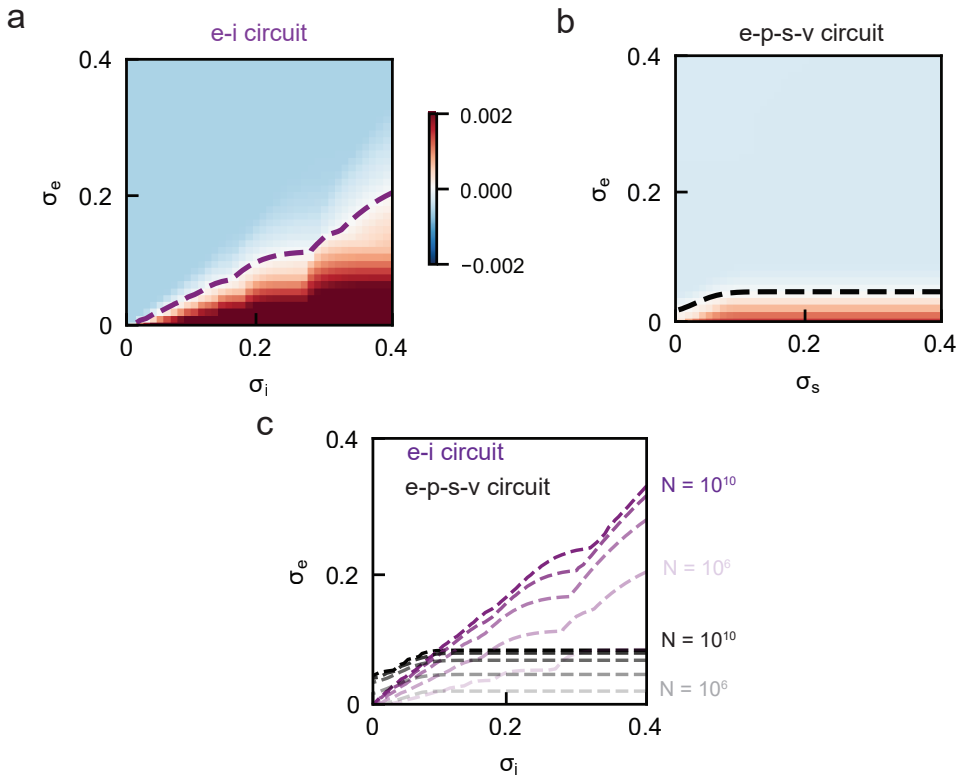

Figure 2: Stability with conductance-based synapses. (a) E and homogeneous I stability diagram with $N = 10^7$ (b) Stability with the E-PV-SST-VIP circuit (c) Stability boundaries with varying network size N for the 2-population (black) and 4-population (purple) models.