PNAS RESEARCH ARTICLE | APPLIED MATHEMATICS
APPLIED BIOLOGICAL SCIENCES

Efficient models of cortical activity via local dynamic equilibria and coarse-grained interactions

Zhuo-Cheng Xiao[a,b,c], Kevin K. Lin[d], and Lai-Sang Young[e,1]

Affiliations are included on p. 12.

Contributed by Lai-Sang Young; received November 21, 2023; accepted May 14, 2024; reviewed by David Hansel, Gregor Kovacic, and Jonathan E. Rubin

Biologically detailed models of brain circuitry are challenging to build and simulate due to the large number of neurons, their complex interactions, and the many unknown physiological parameters. Simplified mathematical models are more tractable, but harder to evaluate when too far removed from neuroanatomy/physiology. We propose that a multiscale model, coarse-grained (CG) while preserving local biological details, offers the best balance between biological realism and computability. This paper presents such a model. Generally, CG models focus on the interaction between groups of neurons—here termed "pixels"—rather than individual cells. In our case, dynamics are alternately updated at intra- and interpixel scales, with one informing the other, until convergence to equilibrium is achieved on both scales. An innovation is how we exploit the underlying biology: Taking advantage of the similarity in local anatomical structures across large regions of the cortex, we model intrapixel dynamics as a single dynamical system driven by "external" inputs. These inputs vary with events external to the pixel, but their ranges can be estimated *a priori*. Precomputing and tabulating all potential local responses speed up the updating procedure significantly compared to direct multiscale simulation. We illustrate our methodology using a model of the primate visual cortex. Except for local neuron-to-neuron variability (necessarily lost in any CG approximation) our model reproduces various features of large-scale network models at a tiny fraction of the computational cost. These include neuronal responses as a consequence of their orientation selectivity, a primary function of visual neurons.

neuronal networks | coarse-graining | visual cortex | multiscale modeling

Significance

Understanding the human brain is a major scientific endeavor of the 21st century. Computational models offer invaluable insight, but the complexity of the cortex poses challenges: Large-scale, biologically detailed models are difficult to build and expensive to simulate, while oversimplification can compromise a model's biological relevance. Here, we propose a modeling strategy that is computationally efficient and enables a high degree of fidelity to neurobiology. Motivated by nonequilibrium statistical mechanics, our multiscale modeling approach respects cortical structures and exploits them to eliminate redundant computation of local responses. We demonstrate the strategy's effectiveness on a model of the visual cortex. Our method is generalizable and has the potential to advance neuroscience by enabling more comprehensive models of the neocortex.

1. Introduction

Biology-based models of brain circuitry, i.e., models that incorporate relevant neuroanatomy and physiology, have the potential to capture emergent behaviors and offer insights into mechanisms; they also have predictive power, an ultimate goal of neurobiological modeling (1–6). These benefits, however, come at considerable costs: Network models that involve hundreds of thousands of neurons and large numbers of parameters are costly—sometimes prohibitively so—to build, to tune, and to run, limiting their scalability and scope. Idealized models, e.g., phenomenological, reduced, or mean-field models, are more tractable. But inferences from such models can be hard to evaluate: The farther removed a model is from detailed anatomical structures, the less directly connected its outputs will be to the underlying biological substrate and to empirical observations. There is a trade-off between computational tractability and biological realism.

Imagine the following two-step approach to neuronal network modeling: First, design a "premodel" focusing on the relevant neurobiology without regard for computational cost. Then, in a second step, deal with practical issues of implementation by simplifying as needed but striving to retain as much of the realism of the premodel as is computationally feasible. This paper aims to systematize the second step for modeling cortical networks in mammals: We propose a general coarse-graining strategy aimed at significantly reducing the computational cost of large-scale cortical models and illustrate the ideas proposed using a previously constructed model of the primate visual cortex (7).

In the rest of the Section 1, we present an overview of our proposed strategy, postponing discussion of related work to Section 3.

1.1. Proposed Coarse-Graining (CG) Strategy. Assuming that a premodel of cortical dynamics in the form of a large-scale, biologically detailed network has been designed, our aim is a feasible and efficient implementation that minimizes information loss.

Reviewers: D.H., CNRS; G.K., Rensselaer Polytechnic Institute; and J.E.R., University of Pittsburgh.

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: lsy@cims.nyu.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2320454121/-/DCSupplemental.

Published June 26, 2024.

We propose to do this by tracking cortical dynamics on two spatial scales, and to take advantage of the underlying biology to gain computational speed.

1.1.1. Proposed procedure. In the context of neuronal networks, coarse-graining typically means subdividing the cortical region of interest into blocks of nearby neurons and treating each block as a point in an abstract "neural space." At issue is how to define the dynamics of such a reduced model. Instead of relying on best guesses or favoring explicit interaction kernels that are amenable to mathematical analysis (as is often done in mathematical neuroscience), our CG dynamics will be informed by the premodel.

To track simultaneously the dynamics of the CG system and dynamical interactions within blocks, a general approach well-accepted in multiscale modeling is to alternate between "fine" and "coarse" scales (8, 9): (a) compute local responses of each block to its surroundings, and (b) integrate outputs from the previous step(s) to form new input signals for each block, thus propagating signals across the network. Both (a) and (b) follow interaction rules deduced from anatomical and physiological measurements in the premodel.

A crucial observation—one that allows us to exploit the underlying biology to speed up the computation significantly—is that in many cortical systems such as the visual cortex, the set of all potential local responses can be computed *in advance*—without knowledge of the specific stimulus or CG dynamics. Precomputing potential local responses and storing them in a library, we reduce step (a) to lookup and interpolation.

1.1.2. Biological justification. While the mammalian brain as a whole is highly heterogeneous, network architecture (including cell types, local circuits, and longer-range connections) tends to be structurally similar within a specific layer of a specific cortical region. These anatomical facts are well documented in the visual cortex, especially in V1 (primary visual cortex); see e.g., refs. 10–15. Though network architectures differ, there are many examples showing that a similar picture holds in other regions of cortex (16–19).

Mathematically, layers of V1 are described by 2D networks with translationally invariant structures across the cortical surface. But because long-range connections between brain regions have specific targets, structural similarity does not imply that V1 responses to visual stimuli will be similar across the surface. For example, the optic nerve conveys signals from specific locations of the retina to specific locations of the lateral geniculate nucleus, and from there signals are transmitted to specific cortical locations. When a visual scene is presented to the eye, it stimulates different regions of the retina differently, eliciting different responses at different locations in the visual cortex (20).

With this picture in mind, and focusing on a single layer, we now explain why potential local responses can be precomputed.

(i) As local neuronal populations are anatomically similar and their differing responses are due to the different inputs they receive, we posit that local dynamics can be described by a single dynamical system driven by variable "external" inputs; "external" here means from sources outside of the local system.

(ii) It is a basic fact that cortical neurons primarily communicate by electrical spikes, and the impact of a spike depends largely on the presynaptic cell (the cell sending the signal), such as whether it is excitatory or inhibitory, and on the signal transduction properties of the synapse. The exact identity or physical location of the presynaptic neuron is not, in and of itself, relevant to neuronal computation. Thus if there are n different input sources, then external inputs to a local population—from all presynaptic cells combined—can be described by an n-tuple $(X_1(t), X_2(t), \cdots, X_n(t))$, where $X_i(t)$ is the incoming spike rate from source i at time t.

(iii) The $X_i(t)$ vary with time and with events in and outside of the local population, but one can estimate a viable range, i.e., a finite interval containing all biologically reasonable rates. The set of viable input rates forms a bounded set in $\mathbb{R}^n$, and "potential local responses" are responses of the local dynamical system in (i) to all n-tuples of input rates in this bounded set. See Section 2.2.

1.2. Relation to Nonequilibrium Statistical Physics, Balanced State Ideas in Neuroscience, and Multiscale Simulations. Consider the following very rough analogy with the physical phenomenon of heat conduction: Take a large, spatially homogeneous interacting particle system and couple it to heat baths. If the baths are at different temperatures, the system will be driven out of thermal equilibrium and settle into a nonequilibrium steady state (NESS). As discussed above, when suitably restricted, e.g., to a layer of V1, neuronal networks are spatially homogeneous in their anatomical structure, driven out of equilibrium by visual stimuli that vary with spatial location and with time. When the stimulus is time-stationary, such a system also tends to a NESS.

It has been proposed in statistical mechanics (21) that a NESS can be viewed as a continuum of mesoscopic-sized local systems in local thermal equilibrium, with well-defined local temperatures that reflect the balance of energy and other fluxes across the system. In the same spirit, we propose that for a neuronal system in a NESS, local populations are in local dynamic equilibrium (LDE), their firing rates determined by excitation and inhibition both within the local circuit and among nearby circuits (22). It can be shown that internally, local systems that are in LDE are in approximate balanced states, a concept developed originally for systems with homogeneous connection probabilities and inputs (23, 24). For a network subjected to spatially varying external input, a NESS can thus be seen as a collection of LDEs, or approximate balanced states, connected to one another through excitatory and inhibitory currents.

In terms of computational strategy, the NESS point of view leads naturally to a multiscale approach (8, 9). However, most neuronal systems do not exhibit the sharp timescale separation between local dynamics and larger-scale interactions upon which most multiscale methods depend for speedup (Section 2.5). The application of such methods at the level of biological realism considered in this paper is thus untested and not without challenges. A modification that gives satisfactory results is proposed.

1.3. Primary Visual Cortex As an Illustrative Example. In this paper, we demonstrate our proposed computational strategy using a previously constructed premodel. In this proof-of-concept paper, using an existing network allows us to focus on the proposed CG ideas and evaluate how well the algorithm works by comparing CG model outputs with those from the premodel.

The biologically detailed network that will serve as premodel for us is that of an input layer of the primary visual cortex (V1) of monkeys (7). One of the most important functions of V1 is the identification of orientations of edges at specific spatial locations, edges that the cortex eventually puts together to form contours of the objects we see. This visual capability of neurons is known as orientation selectivity (OS). Hubel and Wiesel made some very insightful observations more than half a century ago (25), but neural mechanisms for OS are subtle and not fully understood.

They pose a nontrivial challenge to mean-field or coarse-graining approaches, and we believe it is an appropriate first test for our CG strategy to correctly reproduce this phenomenon.

2. Results

2.1. Premodel: A Large-Scale Network of the Monkey Primary V1. As explained in Section 1, this paper introduces a method for deriving an efficient multiscale model from a premodel that can be as detailed biologically as one sees fit. To illustrate the procedure, we will use as premodel a previously constructed network of the visual cortex.

This section introduces the premodel. The material is taken from ref. 7, with modifications from refs. 26 and 27. The description below and in *SI Appendix* are not intended to be complete: Our aim here is to provide enough information so that readers who are not experts in visual neuroscience could—without delving into the references above—appreciate the major differences between the biologically detailed model described in this section and the representation we propose in the next.

2.1.1. The LGN-L4-L6 circuit. In the first stages of visual information processing, signals received by the retina are passed by the optic nerve to the Lateral Geniculate Nuclei (LGN) in the thalamus, which relay the signals to the primary V1, and from there to higher visual cortical areas. This basic information is discussed in standard neuroscience texts. For a general, nontechnical intro to the primate visual system, see ref. 28; for a concise summary of V1 properties, see ref. 29. V1, the largest and most complex of all the visual cortical areas, is further subdivided into layers (11, 30). Of primary interest to us is layer 4Cα: This is the input layer in the Magnocellular ("Magno") pathway, one of the two major pathways from the retina to V1.

There are three components in the premodel: layer 4Cα (henceforth abbreviated as "L4"), LGN, and layer 6 ("L6") of V1. L4 receives feedforward input from LGN, and feedback from L6. A schematic diagram of this LGN-L4-L6 circuit is shown in Fig. 1*A*.

There is a continuous point-to-point map, called the retinotopic map, from one's (2D) visual field to the retina. As cells from LGN, L4, and L6 receive feedforward signals from specific, highly localized regions of the retina (31), we can coordinatize these components according to their points of origin in the retina, i.e., the retinotopic map can be extended to LGN, L4, and L6. Our premodel is about a small patch of LGN and cortex targeted by projections from the retina in one eye, a few degrees from the fovea or "center of gaze."

Below, we describe the network architecture and a few of the most salient features in the three components of the premodel. The exposition here, together with what we provide in *SI Appendix*, sections 1A and B, should suffice for understanding the present work. For further details and for references to anatomical and physiological papers that support the premodel, we refer the reader to the original publication (7).

2.1.1.1. LGN (feedforward input to L4). For our purposes, LGN cells relay information from the retina to the cortex: They receive input from retinal ganglion cells and output to cells in L4 (and to a lesser extent L6); LGN relay cells do not interact among themselves. They come in two kinds, ON and OFF, forming two fairly regular mosaics on an essentially 2D surface. A striking feature of LGN is that it is very sparse. In the Magno layers at the eccentricity modeled, there are on the order of 10 cells (5 ON and 5 OFF) in a region of area $0.25° \times 0.25°$, compared to about 4,000 cells in a similar size region in L4.

In the Magno pathway, LGN cells are detectors of changes in luminance in their relatively small receptive fields (RF). In the absence of visual input, they fire about 20 spikes/s. ON-LGN cells are excited, firing vigorously up to over 100 spikes/s, when the luminance in their RF changes from dark to light, and are silent when it goes from light to dark. OFF cells do the reverse, spiking vigorously when their RF change from light to dark.

2.1.1.2. L4 (the principal component in the premodel). About 80% of the neurons are Excitatory (E); the rest are Inhibitory (I). (These percentages differ slightly from ref. 7 but are quite standard.) Cells communicate by firing action potentials, or "spikes," and sending them along axons. The sending cell is called presynaptic; the receiving cell postsynaptic. Signals from excitatory presynaptic cells bring a postsynaptic cell closer to its spiking threshold; signals from inhibitory presynaptic cells bring a postsynaptic cell farther from its spiking threshold. See Section 2.1.2 below.

As a mathematical idealization, it is assumed in ref. 7 that E and I-neurons are located on two square lattices on a 2D surface. Anatomical facts of the following type are incorporated. The probability of connection among model cells is dependent on cell types (i.e. E or I) and drops off with distance following a truncated Gaussian. E-neurons have longer axons, hence longer reach, compared to I-cells. Peak connection probability among E-cells is ∼15%, while peak connection probabilities for E→I, I→E, and I→I are much higher, at about 60%. The connection strength also depends on cell type but is independent of the distance between pre- and postsynaptic cells.

There are two somewhat distinct populations of E-cells: simple and complex. Simple cells follow detailed spiking patterns of

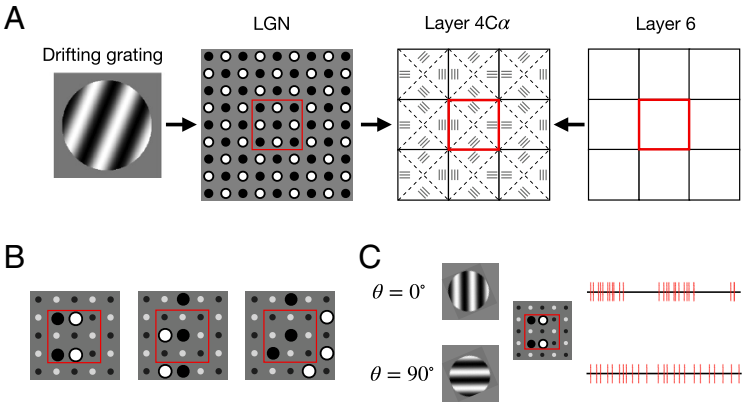

Fig. 1. (*A*) LGN-L4-L6 network in the premodel, with modeling of L6 following ref. 7. Signals mostly (but not necessarily) in the form of drifting gratings are sent to LGN, then passed to L4 which also receives feedback from L6. LGN is represented schematically by a lattice of ON (white) and OFF (black) cells. L4 is subdivided into squares called hypercolumns (HCs). Each HC consists of neurons sharing similar receptive fields. Roughly, cells in the red, center square in the LGN panel project to the red, center square in L4. Each HC is further subdivided into orientation domains (ODs); wiring to LGN is determined by the orientation of the three bars indicating the intended preferred orientation. (*B*) Sample LGN configurations to which cells in vertical-preferring ODs (those marked by vertical bars in panel *A*) are wired. (*C*) Pooled LGN spike trains received by a V1 cell wired to the given LGN configuration. Two spike trains, elicited by two different grating orientations, are shown. The alignment of the top grating with LGN cells causes the pooled LGN spike train (formed from the four LGN spike trains) to concentrate in half-cycles. This firing pattern produces higher firing in the recipient L4 cell than the more evenly spaced LGN spikes produced by the orthogonal grating.

LGN cells; complex cells do not. In ref. 7, simple cells typically receive 4 to 6 LGN inputs, while the corresponding numbers for complex cells are 0 to 2. Complex cells, on the other hand, have significantly higher firing rates than simple cells. In ref. 7, this is achieved by giving them a larger number of presynaptic E-cells.

2.1.1.3. L6 (feedback to L4). In the real cortex, L6 receives some input from LGN and interacts with multiple other layers of V1, including L4. As there is no way to systematically model such feedback dynamics without modeling all the layers involved, as well as the regions with which these layers interact, the authors of ref. 7 made the simplifying assumption that L6 activity followed LGN alone; its neurons were modeled as spike trains providing excitatory input to L4.

2.1.2. Equations of neuronal dynamics.

2.1.2.1. LGN dynamics. Inputs to LGN cells are time-dependent light intensity maps $L(x, t)$ giving the luminance of the visual stimulus at location x at time t, filtered through certain spatial and temporal kernels. Since our main focus is L4 and we will not be working directly with LGN dynamics, we will omit discussion of this topic, referring the interested reader to ref. 7 or to Section 4 of ref. 27.

2.1.2.2. Dynamics of L4 neurons. L4 neurons are modeled as conductance-based, leaky integrate-and-fire (LIF) point neurons, the dynamics of which can be summarized as follows: The membrane potential V of a neuron of type $Q \in \{E, I\}$ is described by

$$\frac{dV}{dt} = -\frac{1}{\tau_L^Q} V - g_E^Q(t)(V - V^E) - g_I^Q(t)(V - V^I)\,. \quad [1]$$

The three terms on the right side of [**1**] are the Leak, Excitatory, and Inhibitory currents, respectively. In rescaled units (following ref. 32), the spiking threshold is $V = 1$; upon spiking, V is immediately reset to 0 and remains there for a refractory period of 2 ms for E-cells and 1 ms for I-cells. When the refractory period is over, the evolution of V resumes according to Eq. **1**. The time constants are $\tau_L^Q = 20$ ms for $Q = E$ and 15 ms for $Q = I$; $V^E = 14/3$ and $V^I = -2/3$ are the E and I "reversal potentials," and $g_E^Q(t)$ and $g_I^Q(t)$ are E and I-conductances, which we now describe.

The I-conductance of a neuron of type-Q, $g_I^Q(t)$, is driven by the spiking activity of presynaptic I-cells from within L4 and has the form

$$g_I^Q(t) = S^{QI} \sum_{i=1}^{\infty} G_I(t - t_i^I),$$

where $G_I : [0, \infty) \to [0, \infty)$ describes the time course of the conductance: It rises quickly from 0, decays exponentially after 5 to 10 ms, and satisfies $\int G_I\, dt = 1$. The t_i^I are the arrival times of incoming I-spikes, and S^{QI} is the synaptic coupling strength from I-neurons to neurons of type Q. That is, each spike fired by a presynaptic I-cell elevates the inhibitory conductance of a postsynaptic cell of type Q for ∼10 ms by an amount proportional to S^{QI}. See *SI Appendix*, section 1E.

E-conductance is defined similarly, but there are four sources of excitatory input, so the right side of the equation for $g_E^Q(t)$ is the sum of 4 terms. Three of the sources are synaptic: They are LGN, L4, and L6, with coupling strengths $S^{Q\mathrm{LGN}}$, S^{QE}, and $S^{Q\mathrm{L6}}$, respectively. These inputs have different time courses: E-spikes from L4 and L6 have an AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid) component that is fast-acting, decaying exponentially after a few ms, and a much longer-acting NMDA (N-methyl-D-aspartic acid) part that lasts for ∼80 ms. LGN spikes have AMPA only. The fourth source, called "*ambient*," represents modulatory influences not explicitly modeled; it is depicted by Poisson spike trains. See *SI Appendix*, section 1D for more detail.

We remark that while Eq. **1** governing single-neuron dynamics is exceedingly simple, the coupling among neurons leads to very complex interactions. When stimulated, a neuron receives barrages of E and I-spikes, producing a great deal of fluctuations in its membrane potential. Even as the mean drift is upward, it is generally hard to predict when spiking threshold is crossed. Exact firing rates of LIF neurons cannot be expressed in closed form in terms of the parameters above; they depend on input patterns (as we will see momentarily) as well as recurrent interactions. See ref. 33.

2.1.3. Orientation selectivity. Nearly all V1 neurons have a preferred orientation: it is most excited when an edge in this orientation is detected in its receptive field, less so as the angle between the edge and the neuron's preferred orientation is increased, and shows little to no increase in firing when the edge is orthogonal to its preferred orientation (25, 34). This response property is known as orientation selectivity, abbreviated OS in the rest of this paper. OS is of central importance as a visual function, and it comes about in a rather subtle way as we will explain, making it a challenge for CG strategies.

Below, we outline some facts and conjectures related to the origin of OS. These ideas were made precise and implemented in ref. 7, which confirmed that they produced model outputs consistent with experimental data.

2.1.3.1. LGN-to-V1 wiring: From empirical observations to model. Of relevance are the following two empirical facts:

(a) Based on their electrophysiology experiments half a century ago, Hubel and Wiesel proposed that OS is derived from feedforward convergence of LGN cells. Specifically, they hypothesized that V1 simple cells are connected to two or three spatially aligned rows of LGN cells that alternate in polarity, i.e., a row of ON and a parallel row of OFF, or three rows in parallel alternating in ON/OFF (25). Receptive field mappings from many authors (35, 36) have since provided evidence in favor of this hypothesis.

(b) Imaging results show that V1 cells with like orientation preference tend to be grouped together in what are called orientation domains and that these domains appear to form "pinwheel" structures around certain "singularities" (37).

Following (b), the authors of ref. 7 divided L4 into hypercolumns (HCs) representing cells with nearly identical receptive fields; each HC is subdivided into OD, and each OD is assigned an intended orientation preference. See Fig. 1*A*. Following the suggestion that the alignment of a V1 cell's LGN afferents may be connected to its preferred orientation, E-cells in each OD are wired to a group of LGN cells with a configuration as proposed in (a) and an orientation following its assigned preference. As noted earlier, LGN cells in the Magno layers are very sparse, and there are constraints on how far LGN cells can project. Based on these and other restrictions, it was concluded in ref. 7 that the number of LGN cells connected to a V1 cell is likely quite small (Section 2.1.1). Examples of LGN configurations presynaptic to a V1 cell are shown in Fig. 1*B*. I-cells are given similar numbers of LGN inputs but they are chosen randomly.

2.1.3.2. Model validation. Drifting gratings, visual stimuli that are periodic in space and in time, have been widely used to calibrate properties of visual neurons in experiments. In real V1,

cellular responses vary with grating orientation. A V1 neuron is said to be optimally driven when the grating is aligned with its preferred orientation, orthogonally driven when the grating is orthogonal to the preferred orientation. Tuning curves, graphs of firing rates plotted against the orientation of the grating, summarize the OS of neurons; notably, they peak at preferred orientations. See, e.g., ref. 34.

Model outputs in ref. 7 are in strong agreement with data: Nearly all V1 cells have OS with realistic tuning curves. Moreover, most cells prefer orientations not far from those of their assigned intended orientations, with most preferred orientations varying continuously as a result of recurrent interactions.

2.1.3.3. Mechanism behind OS. How OS is acquired by V1 cells requires an explanation because it is a well-established fact that LGN cells, which supply the sole feedforward input to V1 in monkeys, are not orientation-selective (38–41). We give below an explanation for how OS is conferred on V1 cells. The modeling in ref. 7 is based on this explanation, various versions of which are known within the neuroscience community.

Consider an LGN configuration as proposed by Hubel and Wiesel projecting to a V1 cell. When optimally driven, all the LGN cells are simultaneously excited during roughly half of the cycle, and none are excited during the other half. For an orthogonal grating, most of the time about half of the LGN cells are excited, and they send to V1 spike trains that are more evenly distributed in time. See Fig. 1*C*; for further illustration, see *SI Appendix*, section 1B. One way to understand why a concentrated LGN spike volley produces a stronger V1 response is that a strong, persistent, net-positive current leaves the leak term in Eq. **1** with less time to act before the threshold is reached.

The discussion above explains how *simple* E-cells acquire their OS, but complex E-cells (which have at most one or two LGN inputs) and I-cells (which have random LGN configurations) enjoy OS as well. In ref. 7, complex cells and I-cells inherit the orientation preference of nearby simple E-cells.

2.1.3.4. Why OS is a challenge to theorists attempting model reduction. OS means a cell fires preferentially in response to certain grating orientations. One might naively assume that this has to do with differentials in input firing rates, i.e., more LGN spikes are received when the stimulus is aligned with the cell's preferred orientation. But that is not the case: Because the set of LGN afferents to a V1 cell is fixed and LGN cells are not OS, the number of incoming LGN spikes per sec is independent of grating orientation. For a simple cell wired to LGN as hypothesized (see above), the pooled spike trains of its LGN afferents exhibit certain temporal patterns (Fig. 1*C*), patterns that depend on the wiring and vary with grating orientation. The challenge in performing CG is to capture how such a cell integrates these input patterns with recurrent cortical interaction to produce a suitable firing rate response.

2.1.4. Outlook. To motivate the need for a different computational strategy, we examine some of the numbers in the biologically detailed models (7, 26, 27): These models covered 9 HCs; each HC has ∼4,000 cells in L4, is driven by 10 LGN cells and has feedback from ∼300 cells from L6. This adds up to a total of about 40,000 cells in the entire model. L4 cells were modeled with integrate-and-fire neurons, as were LGN cells, and L6 cells were modeled as spike trains. The dynamics were updated at 0.1 ms intervals to capture the order in which the interactions take place. Simulation time was about 2 min for each second of neuronal activity. Parameter determination was challenging and labor-intensive, though manageable.

Note, however, that 9 HCs represent a rather small visual area of 0.75° × 0.75°, and this LGN-L4-L6 circuit is only a small part of V1. A rough estimate suggests that a similarly detailed depiction of a multilayer model of V1 covering say a 10° × 10° area in parafovea will involve scaling up the number of neurons by factors upward of 10^3, and simulation cost scales at least linearly with network size. These challenges caused us to consider CG models as an alternative.

2.2. A Coarse Graining Strategy. Recall that we view a cortical network presented with a stationary visual stimulus as being in a NESS if the system is in local dynamic equilibrium and excitatory and inhibitory forces across the cortical surface are in balance. Motivated by these ideas, we propose in this section an iterative scheme that, if and when it converges, would lead to a NESS of the CG system.

As is typical in CG strategies, we divide cortical space into blocks of local populations which we will refer to as "pixels," associate to each pixel a finite set of relevant statistics, and track the time evolution of these statistics. In the example we use for illustration, pixels are local populations of L4 in V1 (Section 2.1), and the statistics we track are firing rates. Two distinguishing features of our proposal are

(i) interaction kernels and local responses of our CG system are derived carefully from a biologically detailed premodel;
(ii) potential local responses are precomputed.

With (i) we seek to retain biological realism from the premodel; (ii) speeds up the computation significantly.

2.2.1. Setup. We consider a patch of 4 × 4 HCs of layer L4 in V1. Following ref. 27, we divide each HC into four ODs with intended orientation preferences 0, 45, 90, and 135°. Keeping in mind that pixels should be small enough to justify the assumption of homogeneity in neuronal activity (see "working hypothesis" below), and at the same time large enough to reduce the computational cost, we subdivide each HC into 10 × 10 pixels (Fig. 2*A*). With this subdivision, each OD in a HC corresponds to ∼25 pixels, and each pixel contains, following cell densities in the premodel, ∼32 E-cells and ∼8 I-cells on average. The E-population is further subdivided into simple (S) and complex (C) cells, with ∼70% of the E-population being simple. It is necessary to treat S- and C-cells separately as they receive different LGN and cortical inputs.

Our main working hypothesis is this: we assume in our CG model that (a) cells of the same type (S, C, or I) within each pixel have the same firing rates, and (b) the inputs they receive all have identical firing rates. This is obviously a simplification of what goes on in the premodel, where there is much neuron-to-neuron variability, but a necessary assumption in any coarse-graining.

Accepting this hypothesis, the problem can now be framed mathematically as follows: For each $p \in \mathcal{P}$, where $\mathcal{P}$ is the set of all 40 × 40 pixels in the patch, let (f_p^S, f_p^C, f_p^I) denote the mean firing rates of simple-E, complex-E, and I-cells in pixel p. We refer to $f = \{(f_p^S, f_p^C, f_p^I), p \in \mathcal{P}\}$ as a firing rate configuration, and let $\mathcal{F}$ denote the set of all such configurations. In our illustrative example, visual stimuli are drifting gratings at high contrast, at fixed spatial and temporal frequencies. We vary only their orientation θ, where $\theta = 0$ is equated with vertical. Given θ, we seek a mapping

$$\Phi_\theta : \mathcal{F} \to \mathcal{F}$$

that describes how each pixel p equilibrates to the stimulus and to its surroundings, i.e., $\Phi_\theta(f)$ tells us how each pixel responds to LGN, L6, and ambient inputs when L4 inputs are computed

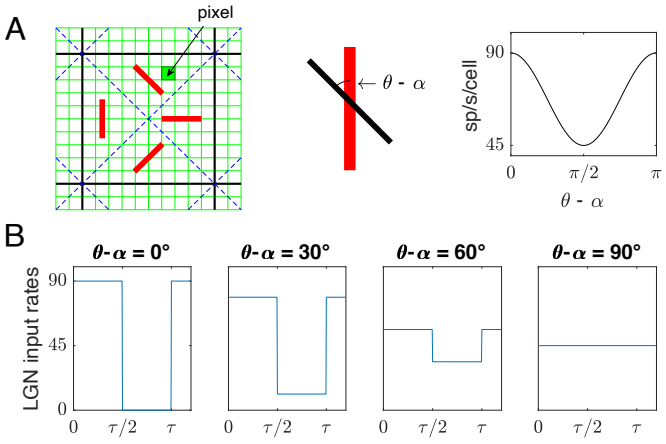

Fig. 2. (*A*) *Left*: A single hypercolumn (center square) divided into four ODs with solid red bars indicating intended orientation preferences; finer, green mesh shows subdivision into pixels. *Middle*: black bar represents grating orientation. *Right*: the peak LGN input rate to simple cells, i.e., $n_\alpha^{SLGN+}(\theta)/N^{SLGN}$, as a function of $\theta-\alpha$. (*B*) Poisson rates for LGN feedforward input received by simple cells, when the intended preferred angle α and angle of grating θ differ by 0, 30, 60, and 90°. Here τ denotes the temporal period of the grating.

from a configuration frozen at f. A fixed point of Φ_θ, if it exists, would be a NESS of the CG system.

2.2.2. Local responses: Main computation. There are two steps in the computation of $\Phi_\theta(f)$. First, we identify all the inputs from lateral and external sources received by cells of type $Q \in \{S, C, I\}$ in each pixel p. Then, we propose new firing rates in response to these inputs. The goal of this subsection is to explain these two steps under the following

"No-mixture" assumptions:

1. each pixel is contained entirely in a single OD, and
2. inputs from LGN and from L6 are identical for all cells of type-Q on each OD.

These assumptions are only partially valid in the premodel (LGN and L6 inputs do cross OD boundaries), but they allow us to present the CG procedure in its entirety with few distractions. In Section 2.3, we will modify the procedure to allow "mixtures."

2.2.2.1. Inputs to pixels. Let θ and (f^S, f^C, f^I) be given. We fix a pixel p, assumed to be in an OD with intended preference α for some $\alpha \in \{0°, 45°, 90°, 135°\}$. We state below how the spike rate from each input source is computed from (f^S, f^C, f^I) and θ; see *SI Appendix*, sections 1A and B for details on parameters and how various relevant quantities are extracted from the premodel.

From within L4: Let n_p^{QE} and n_p^{QI} denote the number of E and I spikes per second received by each cell of type Q in pixel p. To ensure that interactions among pixels follow closely those between the corresponding groups of neurons in the premodel, we set

$$n_p^{QE} = \sum_{p' \in \mathcal{P}} \left[c_{p \leftarrow p'}^{QS} \cdot f_{p'}^S + c_{p \leftarrow p'}^{QC} \cdot f_{p'}^C \right], \; n_p^{QI} = \sum_{p' \in \mathcal{P}} c_{p \leftarrow p'}^{QI} \cdot f_{p'}^I, \quad [2]$$

where $c_{p \leftarrow p'}^{QR}$ is the mean number of type-R cells in pixel p' in the premodel presynaptic to a cell of type-Q in pixel p, averaged over all cells of type-Q in pixel p.

From LGN: For simplicity, we omit LGN dynamics from the CG model, modeling only LGN feedforward inputs to V1 cells. Our proposed LGN inputs are idealizations of LGN firing patterns observed experimentally as well as those produced by model LGN cells in ref. 7. See Fig. 1*C*.

Following average values in the premodel, we assume that N^{QLGN}, the numbers of LGN inputs to neurons of type Q, are 4.5, 1.5, and 4 for Q = S, C, and I, respectively; these numbers do not depend on the pixel index p. Each LGN cell is assumed to fire 45 spikes/s when (strongly) driven, so the mean number of LGN spikes/s received by each type-Q neuron is

$$n^{QLGN} = 45 \times N^{QLGN}. \quad [3]$$

However, knowing the mean rate n^{QLGN} is not enough: As explained in Section 2.1.3, simple E-cells in L4 derive their OS not from mean LGN firing rates (averaged over time), but from the temporal firing patterns of their LGN afferents. This means we need to know not just the mean LGN-to-L4 rate, but how it varies in time.

When optimally driven, i.e., when $\alpha = \theta$, the LGN spikes received by an S-cell are concentrated on half of the grating cycle; they become less concentrated as $|\alpha - \theta|$ increases from 0°, and are roughly evenly distributed when $\theta - \alpha = \pm 90°$ (see Fig. 1*C*). To simulate the spike patterns above, we assume as an idealization that the summed LGN input to each S-cell has the following form: As a function of time, it is a square wave, with rates n_α^{SLGN+} and n_α^{SLGN-} sp/s occurring at the peaks and troughs, which alternate at half-cycles. These rates satisfy

$$n_\alpha^{SLGN+} + n_\alpha^{SLGN-} = 90 \times N^{SLGN},$$
$$n_\alpha^{SLGN+}(\theta) = \left\{67.5 + 22.5 \cdot \cos[2(\theta - \alpha)]\right\} \times N^{SLGN}. \quad [4]$$

See Fig. 2*B*; more details are given in *SI Appendix*, section 1B.

From L6: Following ref. 7, we assume a phenomenological L6-to-L4 feedback term proportional to the mean LGN input received by L4, with

$$n_\alpha^{QL6}(\theta) = \left\{32 + 28 \cdot \cos[2(\theta - \alpha)]\right\} \times N^{QL6}. \quad [5]$$

Following the premodel, the constants in Eq. **5** (including the number N^{QL6} of presynaptic cells L6) are chosen to ensure the firing rate of a single L6 cell lies in the 4-60 Hz range as the grating angle varies.

2.2.2.2. Local response functions. We now define Φ_θ under the no-mixture assumption. Since all cells of type Q in a pixel p are assumed to be identical and to receive identical inputs (see Working Hypothesis), it suffices to compute the response of a single cell. To reduce computational cost, we work with a single, stand-alone neuron (i.e., not embedded in a network) and supply it with inputs from the various sources as estimated above, thereby simulating the inputs it receives in the premodel.

Let $(\tilde{f}^S, \tilde{f}^C, \tilde{f}^I) = \Phi_\theta(f^S, f^C, f^I)$. For each pixel p and each cell type Q, we compute $\tilde{f}_p^Q$ as follows: We assume our representative neuron is governed by an LIF model (Section 2.1.2) with the same time constants and synaptic weights as in the premodel. Incoming spike rates for the various input sources are estimated as above, and except for $\tilde{f}_p^S$ (where LGN input rates are oscillatory, as explained above), spike arrival times are assumed to be homogeneous Poisson processes. The mean firing rate of the LIF neuron in response to these inputs is $\tilde{f}_p^Q$.

With all the emphasis on the role of oscillations in LGN inputs to S-cells, a natural question is why $\tilde{f}_p^S$ is treated as a single

number. Indeed, in the premodel, spike patterns of individual S-cells can have pronounced phases, but phases among nearby neurons are mixed as a result of different spatial offsets in their LGN afferents. When averaged across local populations, such patterns become mild to nonexistent; see ref. 7.

Having explained how $\tilde{f}_p^Q$ can be computed in terms of θ and (f^S, f^C, f^I), we now propose a further simplification: Since pixels interact only through the terms n_p^{QE} and n_p^{QI} defined in Eq. **2**, and inputs from LGN and from L6 can be deduced from α and θ, our working hypothesis implies there are well-defined response functions Ψ^S, Ψ^C, and Ψ^I with the property that

$$\tilde{f}_p^Q = \Psi^Q(n_p^{QE}, n_p^{QI}, \alpha(p), \theta)\,, \quad Q \in \{S, C, I\}. \qquad [6]$$

The significance of Eq. **6** is this: Whereas $\tilde{f}_p^Q$ depends, a priori, on the previous firing rate configuration (f^S, f^C, f^I), this reformulation shows $\tilde{f}_p^Q$ is really a function of $\alpha(p)$, n_p^{QE}, and n_p^{QI}. This suggests we can precompute potential values of $\tilde{f}_p^Q$ without simulating the full premodel.

2.2.3. The Full Iterative Scheme. We now remove the "no-mixture" assumptions, proposing a simple approximation that we have found to be reasonably accurate in practice.

2.2.3.1. Averaging of local responses. In the premodel, it is assumed that alignments of LGN afferents for L4 simple cells come in a finite number of angles denoted by α, but not all simple cells in each pixel share the same alignment. For pixels that lie on the boundary between two ODs, half of the cells have one alignment and half the other; though strongly biased, angles are still mixed in pixels adjacent to the boundary. If we want to make LGN inputs identical for all cells in a pixel—as one would in CG models—then these inputs should be weighted averages of those in Section 2.2.2. Such a strategy would lead to a family of input rates parameterized by grating angle θ and pixel position p, i.e., $\{n_p^{\mathrm{SLGN+}}(\theta), p \in \mathcal{P}\}$, instead of $\{n_\alpha^{\mathrm{SLGN+}}(\theta), \alpha = 0, 45, 90, 135°\}$. Similar considerations are required for L6 inputs, for which there is more substantial averaging across ODs as L6 cells have a sizable reach.

Allowing the LGN and L6 inputs to depend on p and not just on α increases dimension of the domain of local response functions, an increase in computation cost we can afford in this relatively simple illustrative example. However, anticipating greater complexity in future uses of the CG strategy, we propose to average outputs instead of inputs. More precisely, for each p, we fix weights

$$\{(w_p^{LGN}(\alpha_1), w_p^{L6}(\alpha_2)) \mid \alpha_1, \alpha_2 = 0, 45, 90, 135°\}$$

with $w_p^R(\cdot) \geq 0$ and $\sum_{\alpha_1} w_p^{LGN}(\alpha_1) = \sum_{\alpha_2} w_p^{L6}(\alpha_2) = 1$ for each p. Roughly—though not literally—these weights represent the proportion of influence from the different ODs for each p, for LGN and for L6, as deduced from the premodel. (See *SI Appendix*, section 1C for details.) We then modify the computation of Ψ^S in the last subsection to

$$\Psi^S(n_p^{SE}, n_p^{SI}, \alpha_1, \alpha_2, \theta) \qquad [7]$$

defined to be the response when the LGN input to the LIF neuron oscillates between $n_{\alpha_1}^{\mathrm{SLGN\pm}}(\theta)$ and the L6 input is $n_{\alpha_2}^{\mathrm{QL6}}(\theta)$, and then set

$$\tilde{f}_p^S = \sum_{\alpha_1,\alpha_2} w_p^{LGN}(\alpha_1) w_p^{L6}(\alpha_2) \Psi^S(n_p^{SE}, n_p^{SI}, \alpha_1, \alpha_2, \theta)\,. \qquad [8]$$

For $Q \in \{C, I\}$, LGN inputs are assumed to be identical for all C and I-cells (Section 2.2.2), and $\tilde{f}^Q$ is averaged over α_2 only.

Hereafter, we will refer to the mapping $(f^S, f^C, f^I) \mapsto (\tilde{f}^S, \tilde{f}^C, \tilde{f}^I)$, defined via the averaging above, as Φ_θ.

2.2.3.2. Precomputation of local response functions. Our proposal is to tabulate the values of the local response functions Ψ^S, Ψ^C, and Ψ^I in advance for use in the iterative scheme below. Once this library of local responses is built, it can be used any number of times, for simulations with a wide range of initial conditions and grating stimuli. The arguments of the Ψ^Q are as in Eq. **7**.

We remark briefly on the domains of these functions: For n_p^{QE} and n_p^{QI}, we estimate—using the premodel as a guide—a large enough region of E and I-inputs from L4 that contains all values one may encounter under reasonable conditions. Notice that this region should include not just values from steady-state responses to drifting gratings, but also a wide range of initial conditions and visual stimuli (*SI Appendix*, section 2B). As for θ, recall that LGN and L6 inputs depend only on $|\alpha - \theta|$, and consecutive values of α are 45° apart. Taking advantage of the symmetries in the OD organization, we note that it suffices to consider $\theta \in [0, 22.5°]$ (more details in *SI Appendix*, section 1C). In fact, we consider only $\theta \in \{0, 7.5, 15, 22.5°\}$, interpolating linearly for intermediate values of θ.

Technical details of the precomputation of Ψ^Q, including the set of input values, grid densities, how each data point is produced, interpolation methods, etc., are discussed in *SI Appendix*, section 2.

2.2.3.3. Iterative scheme. We now describe a scheme that, given a grating orientation θ, stepsize $h \in (0, 1)$, and initial firing rate configuration $f(0)$, produces a sequence of firing rate configurations $f(1), f(2), \cdots$. First, we carry out Steps 1 to 3 below for each pixel p:

(1) compute n_p^{QE} and n_p^{QI}, the lateral L4E/I inputs, from $f(k)$ following Eq. **2**;
(2) look up the precomputed values of Ψ^Q, for n_p^{QE}, n_p^{QI} and all pairs (α_1, α_2), interpolating as needed; and
(3) use Eq. **8** to estimate the new value of $\Phi_\theta(f(k))$.

Then

(4) update $f(k)$ by the stepping scheme

$$f(k+1) = (1-h)f(k) + h\Phi_\theta(f(k))\,, \qquad [9]$$

(5) repeat Steps 1 to 4 until convergence.

Step 4, including the need to consider stepsizes <1, is discussed in Section 2.5, where we demonstrate numerically that given a drifting grating, there is a unique configuration $f_* = (f_*^S, f_*^C, f_*^I)$ to which all initial conditions tested converge. This configuration is therefore the sought-after NESS for the stimulus presented.

2.4. Demonstration of Algorithm. We now present some simulations to illustrate the proposed iterative scheme. In the examples shown, $h = 1/3$ in Eq. **9**; see Section 2.5 for more detailed analysis. We continue to use $f = (f^S, f^C, f^I)$ to denote firing rate configurations in pixel space, using $f(k)$, etc., when referring specifically to the state of the iterative scheme.

2.4.1. First examples. To evaluate the performance of the iterative scheme, we start with different initial firing rate configurations $f(0)$ and observe the convergence of the $f(k)$ to a NESS in response to drifting gratings. The NESS with respect to a stimulus is defined operationally to be $f_* = (f_*^S, f_*^C, f_*^I)$, where

$$f_* = \frac{1}{100}\sum_{k=251}^{350} f(k)\,, \qquad [10]$$

and $f(k)$ is the firing rate configuration after k iterations. (As we will see, convergence happens long before $k = 250$, to a NESS that does not depend on $f(0)$; see Section 2.5). For simplicity, here we focus only on the E configurations and measure convergence by the L^2 norms of E-firing rates, i.e.,

$$|f - f_*|_E := \left(\frac{1}{|\mathcal{P}|}\sum_{p\in\mathcal{P}} |f_p^E - (f_*^E)_p|^2\right)^{1/2},$$

where f is an arbitrary firing rate configuration and f_p^E is the average of S- and C-rates in pixel p (weighted by the proportions of S- and C-cells present); $(f_*^E)_p$ is defined similarly.

Results for two examples are presented in Fig. 3. In the first example (Fig. 3*A*), we started from the background and drove the system with a vertical ($\theta = 0°$) grating. The algorithm produced abrupt and drastic changes, with firing rates rising rapidly in the first two iterations from a few spikes/s to ~70 spikes/s in certain locations. From there, it began to relax, settling into an activity pattern with $|f - f_*|_E < 1$ by step 9. Observe that the activity pattern in the last panel shown is consistent with theoretical expectations of NESS responses to vertical gratings: ODs designated as vertical-preferring (Fig. 2*A*) have the highest firing rates, and ODs designated as horizontal-preferring have the lowest firing rates.

In Fig. 3*B*, we started from the (computed) NESS for $\theta = 0°$, then presented the $\theta = 60°$ grating to demonstrate the type of transition that occurs in local circuits when the eye makes saccadic movements. Here, we see an abrupt shift in firing patterns in the very first iteration; convergence was achieved quickly as before.

2.4.2. Orientation selectivity in the CG model. As the firing rate response of a simple cell is determined not by the magnitude of its feedforward input current but by the temporal patterns of LGN spikes it receives (see Section 2.1), we view OS as a major proof of concept for the proposed algorithm. Below, we offer two views of the model's computed NESS in response to drifting gratings in the full range of orientations: activity level across the cortical surface, and tuning curves for individual pixels.

Fig. 4*A* shows NESS profiles of E-activity in our 16-HC coarse-grained model. As the grating orientation varies from 0° to 90°, the "hot spots" (where spiking is strongest) morph from one cortical location to another. Recall that there are only four ODs corresponding to four distinct sets of LGN inputs. That the NESSes vary essentially continuously with grating orientation is largely the result of recurrent interaction among pixels. Observe that the hot spots are consistent with the designated ODs in Fig. 2*A*.

The ultimate test of the model's ability to capture OS is whether or not individual pixels in the CG model have reasonable tuning curves, i.e., $f_p^E(\theta)$ as a function of grating orientation θ for a fixed pixel p. This is a tough test because no attempt was made in the algorithm proposed to ensure OS for individual pixels. Fig. 4*B* shows tuning curves for several pixels marked in the panel on the *Left*. Viewed as averaged measures of orientation tuning, these graphs possess the right quantitative characteristics, though they lack the variability of tuning curves for real neurons (see, e.g., ref. 34). The preferred orientation, i.e., the θ at which f_p^E peaks, is consistent with the location of the pixel.

For comparison with the premodel, we include in Fig. 4*C* a set of simulation results from ref. 26. This panel, which shows E-firing rates in response to the 0° grating averaged over a 1-s period over groups of neurons the same size as our pixels, is analogous to the *Upper Left* panel in Fig. 4*A* (except that the premodel had only 9 HC). The overall firing patterns are entirely similar, confirming that our algorithm reproduces premodel activity reasonably accurately. It is also evident that E-activity of the premodel shows a great deal more variability from pixel to pixel: The premodel is a network of 40,000 neurons with connectivity, LGN, L6 inputs, coupling strengths, and spiking thresholds varying from neuron to neuron—all these quantities

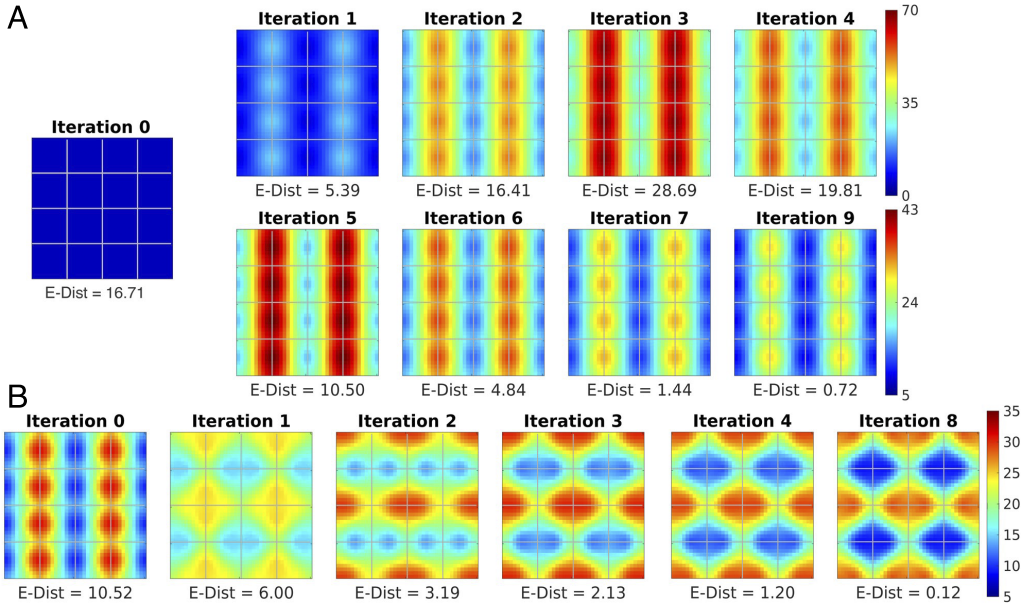

Fig. 3. Computing NESS using the proposed iterative scheme. Shown are E firing rates, i.e., a weighted sum of S and C rates; see Section 2.4.1. In (*A*), the initial condition is $f(0) = (f^S(0), f^C(0), f^I(0)) \equiv (2.5, 8, 18)$, uniform for all pixels, and the stimulus is the $\theta = 0°$ (vertical) grating. Distance to NESS (in $|\cdot|_E$ norm; see text) is indicated below each panel. Note the change in color scale in the second row. In (*B*), the initial condition is the computed NESS in response to the $\theta = 0°$ grating, and the stimulus is the $\theta = 60°$ grating.

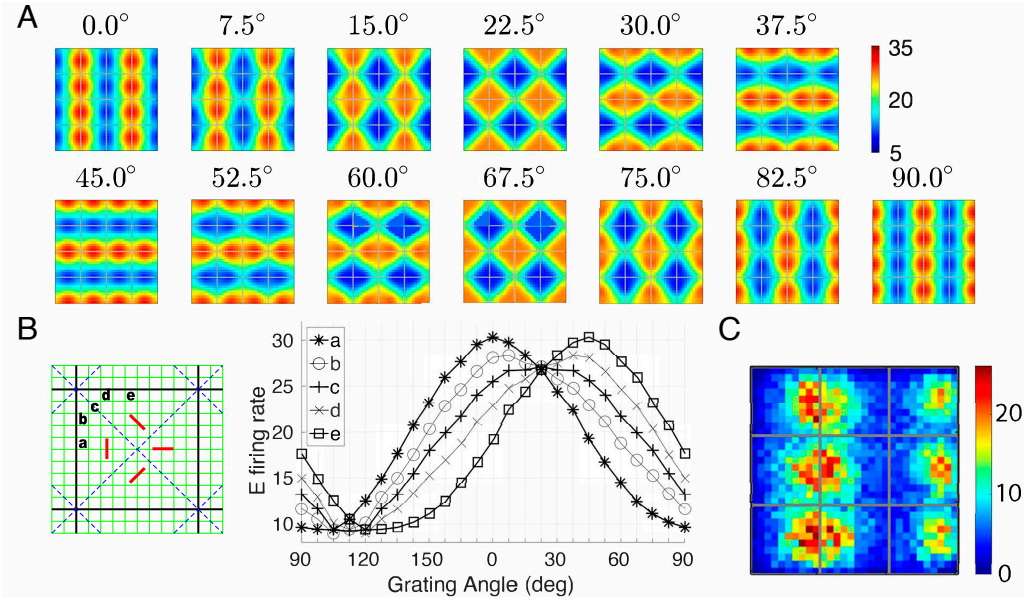

Fig. 4. Demonstration of OS in the CG model. (*A*) NESS in response to gratings from vertical ($\theta = 0°$) to horizontal ($\theta = 90°$); shown are E-firing rates. (*B*) Tuning curves for five pixels a to e corresponding to crosses counterclockwise in the *Left* panel (which shows the firing rates for $\theta = 0°$). (*C*) E-firing rates in the premodel in response to vertical grating; this is Fig. 7*A* from ref. 26. Shown are E firing rates, i.e., a weighted sum of S and C rates (Section 2.4.1). More details are in Section 2.4.2.

are replaced in the CG model by single numbers that correspond to statistical means. On the other hand, producing the data for Fig. 4*A* using the CG model (not counting precomputation time) was $\gtrsim 10^3\times$ faster than running the equivalent premodel for Fig. 4*C*, taking only a fraction of a second on our computer using our (not optimized) implementation. These differences highlight the expected trade-off between fine-scale variability in detailed network models versus the computation speed of coarse-grained systems.

2.4.3. Stress tests. Since neuronal responses to stimuli are known to be complex (see, e.g., ref. 42), we performed an additional ~3,500 tests to challenge the algorithm using a variety of initial configurations designed to identify situations where the algorithm may fail (*SI Appendix*, section 4). The tests included normal conditions, random noise, as well as stress tests, by which we mean initial conditions that occur rarely but that may potentially pose a threat to convergence. The algorithm converged in almost all cases. In a small number (4 out of 3,500) of runs, transients led to arguments (n_p^{SE}, n_p^{SI}) (Eq. **7**) that fell outside the precomputation domain. This could (presumably) be handled by enlarging the domain, but for simplicity, we discarded those runs. The results show that initial conditions for which the E-I balance is significantly disrupted tended to place great stress on the algorithm. An example is shown in Fig. 5; see also *SI Appendix*, section 3.

2.5. Analysis of Proposed Algorithm. The numerical results of the previous section suggest that the iteration scheme [**9**] has an attractive fixed point for stepsize $h = 1/3$. Here we examine convergence behavior, then compare computed local responses to firing rates in the model.

First, we introduce a slight simplification.

2.5.1. Symmetrization. Observe the following organization of hypercolumns (HC) and OD in our model: Between any two HCs, there is an isometry that carries each OD to one with matching intended preferred orientation. Let us refer to a firing rate configuration as "symmetric" when it is invariant under these isometries, and call a stimulus "symmetric" when it is represented by an LGN input that is invariant under these isometries. Drifting gratings are symmetric, and when a stimulus is symmetric, we expect the NESS to also be symmetric (Figs. 3–5). Indeed, the subspace of all symmetric firing rate configurations is invariant under the map Φ_θ (Section 2.2). Numerical evidence (see, e.g., *SI Appendix*, section 3) suggests that this subspace is attracting: Trajectories converge quickly to it before moving more slowly toward the fixed point. For this reason, we focus on dynamics within this subspace. Throughout this section, we use symmetric initial conditions, symmetrize the firing rate configuration at each step, and measure distances using an L^2 norm of the E and I firing rates in one HC, denoted $|\cdot|_{HC}$ (*SI Appendix*, section 4A).

2.5.2. Stepsize analysis. We first explain the need to consider $h < 1$, i.e., the reason for averaging $f(k)$ and $\Phi_\theta(f(k))$ in Eq. **9**. In precomputations of local responses as functions of inputs to a pixel, it is the pixel's steady-state firing rates that are computed (otherwise the result would depend on initial conditions). Thus $h = 1$ is equivalent to allowing the local system representing dynamics within pixels to equilibrate fully to (fixed) external inputs before updating the set of external inputs. In dynamical systems with sharp scale separation, this is often a good approximation, but neuronal dynamics do not have such timescale separation: In the real cortex as in the premodel,

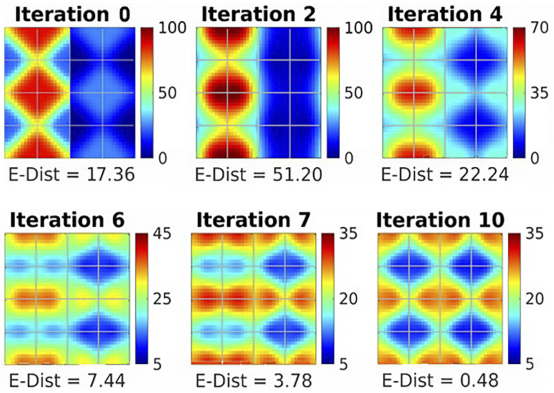

Fig. 5. Example stress test. The initial configuration was constructed by taking the NESS for the 22.5° grating and increasing E-firing rate on the *Left* half of the model by 60% while decreasing it by 60% on the *Right* half; I-firing rates are untouched. The stimulus is the $\theta = 60°$ grating. Note different color scales for iterations 2 to 7.

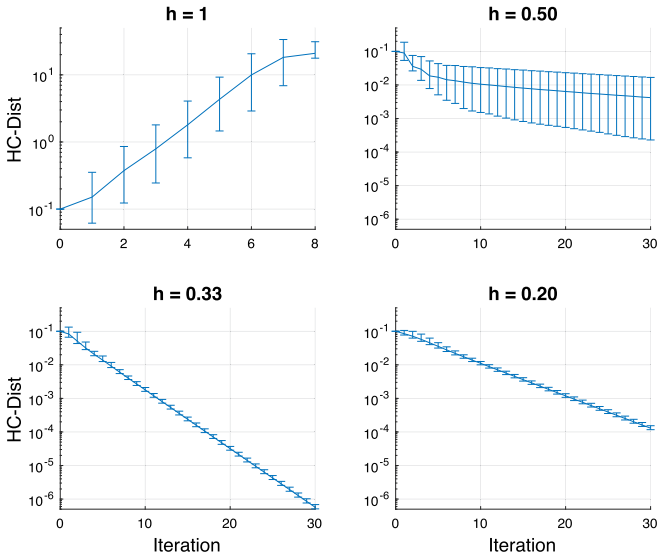

Fig. 6. Local convergence with different stepsizes h in the iteration scheme [**9**]. To compute a putative fixed point f_*, we fix the grating to $\theta = 0°$ (vertical) and sample 500 random initial conditions, evolving each until they are within $\varepsilon = 10^{-12}$ of each other in the $|\cdot|_{HC}$ norm. These are then averaged to form f_*. We then choose 500 new random initial conditions $f^{(i)}(0)$ with $|f^{(i)}(0) - f_*|_{HC} = 0.1$. Each $f^{(i)}(0)$ is run for 350 steps with $h \in \{1, 1/2, 1/3, 1/5\}$ and $|f^{(i)}(n) - f_*|_{HC}$ are computed. Shown are the max, min, and median distances. For $h = 1$, the scheme is unstable, and all orbits "blow up" in nine steps, meaning for at least one pixel p the L4 rates (n^E_p, n^I_p) fall out of the domain of the precomputed local responses. The remaining plots have slow convergence for $h = 1/2$ and exponential convergence for $h = 1/3$ and $1/5$.

timescales of interaction among neurons within a pixel are similar to those with neurons in surrounding pixels. Thus one cannot, a priori, expect $h = 1$ in the iteration scheme to work. Indeed, for initial conditions near f_*, $h = 1$ usually leads to firing rates oscillating with larger and larger amplitudes: Locations that are overly suppressed tend to be overexcited in the next step, and vice versa. See *SI Appendix*, section 4B.

We found numerically that the scheme [**9**] converges for $h \lesssim 1/2$. Moreover, fixed points computed using a range of $h \in (0, 1/2)$ are insensitive to the choice of h, differing from each other by $<10^{-11}$ in the HC norm $|\cdot|_{HC}$. For h away from $1/2$, e.g. $h = 1/3$, convergence appears to be exponential, with a slower rate as h decreases. See Fig. 6.

We hypothesize that the map Φ_θ has a fixed point, though not a stable one. For a configuration f, let $G_h(f) := (1-h)f + h\Phi_\theta(f)$ be the map implied in Eq. **9**. Then $G_1(f) = f$ if and only if $G_h(f) = f$ for all $h > 0$. Moreover, if Φ_θ can be linearized at a fixed point f, then λ is an eigenvalue of $D\Phi_\theta(f)$ if and only if $(1-h) + h\lambda$ is an eigenvalue of $DG_h(f)$. Since exponential convergence for G_h is guaranteed when the spectral radius of DG_h is <1, the results in Fig. 6 give an upper bound on the spectral radius of $D\Phi_\theta$.

2.5.3. Global convergence. We sampled from a group of ~3,500 randomly chosen initial conditions designed to be far from the steady state f_* (see *SI Appendix*, section 4C for details), ran the scheme [**9**] with the $\theta = 0°$ grating and computed distances to the fixed point f_* in the HC-norm as before. As can be seen from Fig. 7, following some initial transients, all trajectories appear to converge at the same exponential rate, consistent with the fixed point hypothesized in the previous paragraph.

2.5.4. Comparison with network dynamics. Next, we compare local responses given by Φ_θ (Section 2.2) with firing rates from the premodel. A priori, the two can be quite different because network interactions typically result in structured spiking with partial synchronization, whereas spike arrival times in the local response computation are Poissonian.

We simulate a spiking network to obtain mean steady-state firing rates, then use these as inputs to compute the local responses given by Φ_θ. If Φ_θ were a perfect representation of local network dynamics, then the predicted responses should be equal to network firing rates. As one can see in Fig. 8, the local response functions give a reasonable estimate though they systematically overestimate a little the firing rates of both E- and I-cells. A phenomenon that may contribute to the discrepancy is that when spiking is partially synchronized, a larger fraction of spikes arrive just before or during refractory making them less effective; see ref. 43.

3. Discussion

3.1. Remarks on Proposed Computational Strategy. We seek an alternative to large-scale simulations of anatomically detailed neuronal networks, one that would allow us to speed up the computation while retaining as much of the biological realism as possible. We propose to track the dynamics on two different levels: a CG level and local dynamics. Upon presentation of a stimulus, we alternately update the CG and local dynamics, viewing the system as having converged to a NESS when i) the pixels in the CG model are in equilibrium with one another, and ii) the local dynamics are in equilibrium with the pixel's external environment. To ensure that our CG model mimics the full network, we draw connectivity data directly from the network; and to ensure that local information is not lost, simulated local dynamics are used to determine interaction among pixels.

An innovation of this paper is the way we take advantage of the underlying biology to represent local systems as small dynamical systems driven by a finite number of external forces, a representation that enables us to precompute potential local responses without reference to the global landscape. The benefits of precomputing such responses once and for all are clear: First, they are relatively easy to compute, typically varying smoothly and monotonically with individual parameters; and second, maintaining a library of precomputed responses obviates the

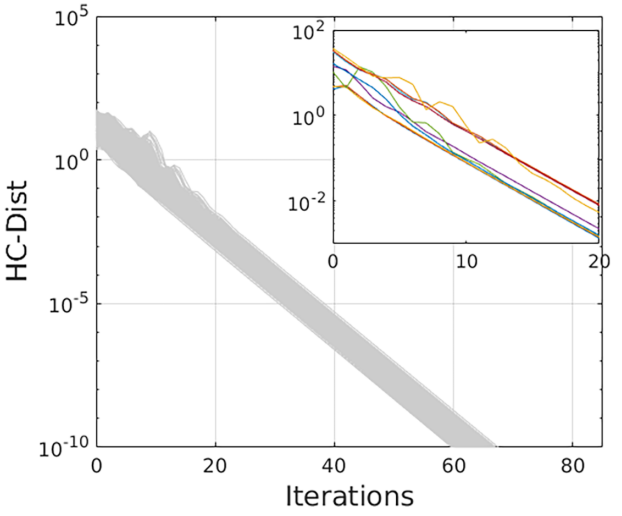

Fig. 7. Testing the global convergence of the iteration scheme for $h = 1/3$. This figure plots the trajectories of distances of ~3,500 tests to the fixed point f_* in the HC-norm. The *Inset* shows the initial segments for five randomly selected trajectories.

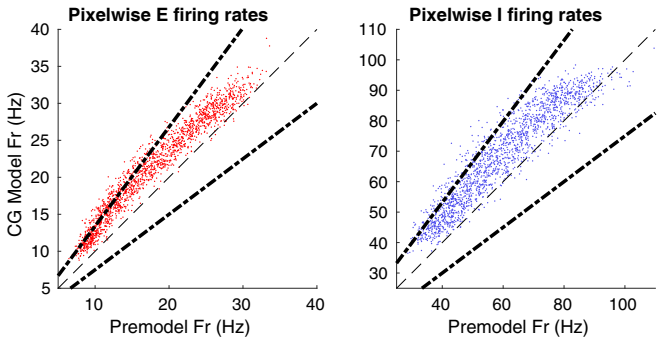

Fig. 8. Comparison between the firing rates from network simulations and local response functions. For $\theta \in \{0, 7.5, 15, 22.5°\}$, we simulate a version of the premodel (timestep = 0.1 ms) and compute its firing rates. Then for each pixel, we compute local responses using the network firing rates as input into the precomputed map Φ_θ, and randomly select 25% to plot against corresponding results from the premodel. Thick dashed lines indicate a relative error of 33%. For E firing rates, 80.4% fall within the 33% bound; for I, 92.7%.

need to simulate the local dynamics at each spatial location in each time step, for each visual stimulus, and for each initial condition, simulations that are quite similar but are duplicated over and over in usual iterative schemes. The reduction in computation costs was evident: e.g., it took <1 s for the examples in Fig. 3 to converge, compared to several minutes for the premodel in ref. 7.

Another notable feature is our use of multiscale ideas in a setting without timescale separation. In this setting, the convergence of our iteration scheme cannot be taken for granted. Analyses of the iteration scheme's convergence properties and its justification on the basis of network dynamics are beyond the scope of the present paper and will be explored in future work.

3.2. Related Work. What we propose is the use of rate models as coarse-grained descriptions of network dynamics. There is an extensive literature on this topic, beginning with the seminal work of Knight on population coding (44) and that of Wilson, Cowan, and coworkers on rate models (45–47), continuing to more recent work on neural fields (see, e.g., refs. 48 and 49), density-based methods (50, 51), diagrammatic expansions (52, 53), and equation-free modeling (54). With some exceptions, much of this body of work has focused on analyzing "universal" or generic phenomena, e.g., bifurcation and pattern formation, with the apparent goal of building general mathematical theories of neuroscience. In contrast, we seek to build a model of the cortex that is simple enough to be understood yet realistic enough to shed light on specific biological mechanisms—to be queried and to offer guidance for future experiments. These differences in aims have resulted in different model choices.

Closest in spirit to our work are (55, 56). To our knowledge, these are the first papers that sought to produce coarse-grained models of large-scale biologically realistic networks. Starting also from a network model of the visual cortex, the authors derived CG equations using analytic approximations valid in a mean-driven, high-conductance regime that is consistent with a weak-coupling, ∞-population limit (51).

We have employed a different computational strategy as local populations—in V1 as in our premodel—are finite in size, neurons respond to both mean and fluctuating stimuli, and they are not weakly coupled. Indeed, aside from E-to-E (for which connection probabilities are lower), each neuron in our premodel is coupled to ~50% of nearby E and I neurons, with synapses strong enough that each incoming spike can drive a neuron's membrane potential a few percent toward or away from its spiking threshold; see *SI Appendix*, section 1. In the absence of good analytic approximations for such regimes, we have chosen to compute our local responses by direct simulation, a procedure made practical by the observation that such responses can be precomputed.

Our scheme bears some similarity to multiscale methods originally developed for computational physics and engineering (8, 9). However, these methods typically rerun local dynamics at each step, requiring timescale separation for speedup, whereas our system does not exhibit sharp scale separation (see above) and we precompute local responses only once.

3.3. Potential Applications and Extensions of CG Methodology. The monkey visual cortex, which is very similar to that of humans, is rich in data and offers a window into the rest of the cerebral cortex. It is an ideal brain region for the development of mathematical modeling techniques and a viable starting point for a quantitative approach to the neurobiology of cortical computation. The primary visual cortex, or V1, consists of multiple interconnected layers, of which one (layer 4Cα) was studied in the present paper. We envision using the principal layers of V1 as our next proving grounds for the modeling strategy proposed, with a longer-term goal of building a semirealistic multilayer model of V1. Such a model could be a major contribution to neuroscience.

Extensions of the techniques developed in this paper will be needed to carry out these next steps. Going beyond the input layer 4Cα, network architecture becomes more complex, but we expect the computation of potential local responses to remain tractable: Local populations are generally driven by no more than a handful of "external" sources, with respect to which firing rates tend to vary in relatively simple ways.

A major challenge in building cortical models is to locate viable physiological parameters, e.g., synaptic coupling weights and connection probabilities. (We avoided parameter tuning in this paper by starting from an existing model.) For parameter exploration, the significantly higher speed of our CG computation is an advantage over direct network simulations. Also, being mathematically cleaner, CG frameworks are better set up for machine learning (ML) techniques.

When complexity is high, the number of unconstrained parameters is large, or experimental guidance is inadequate, we anticipate that ML will be used to fill in the gaps (57). Techniques to "learn" dynamical models using ML methods are being developed; see, e.g., refs. 58–62 for some that have been found promising in physical applications. Such tools can potentially be combined with our approach to improve the predictive power of CG models.

4. Materials and Methods

All algorithms were implemented in MATLAB (2022a) from MathWorks, Inc. The CG and premodel were run on a 16-core machine (Advanced Micro Devices Ryzen 9; Ubuntu 20.04 LTS. Precomputations were done on Lenovo SD650 nodes at the New York University's Greene High Performance Computing Cluster.

Data, Materials, and Software Availability. Software source code has been made publicly available (63).

ACKNOWLEDGMENTS. K.K.L. was supported in part by NSF grant DMS-1821286. L.-S.Y. was supported in part by NSF grant DMS-1901009. Z.-C.X.

acknowledges the support of the Courant Institute of Mathematical Sciences, New York University, where much of this work was performed.

Author affiliations: [a]New York University - East China Normal University Institute of Mathematical Sciences, New York University, Shanghai 200124, China; [b]Institute of Brain and Cognitive Science, New York University - East China Normal University, New York University, Shanghai 200124, China; [c]College of Art and Sciences, New York University, Shanghai 200124, China; [d]Department of Mathematics, University of Arizona, Tucson, AZ 85721; and [e]Department of Mathematics, Courant Institute of Mathematical Sciences, New York University, New York, NY 10012

Author contributions: L.-S.Y. designed research; Z.-C.X., K.K.L., and L.-S.Y. performed research; Z.-C.X. contributed new analytic tools; Z.-C.X., K.K.L., and L.-S.Y. analyzed data; and Z.-C.X., K.K.L., and L.-S.Y. wrote the paper.

1. T. J. Sejnowski, C. Koch, P. S. Churchland, Computational neuroscience. *Science* **241**, 1299–1306 (1988).
2. L. F. Abbott, Theoretical neuroscience rising. *Neuron* **60**, 489–495 (2008).
3. W. Gerstner, H. Sprekeler, G. Deco, Theory and simulation in neuroscience. *Science* **338**, 60–65 (2012).
4. D. S. Bassett, P. Zurn, J. I. Gold, On the nature and use of models in network neuroscience. *Nat. Rev. Neurosci.* **19**, 566–578 (2018).
5. H. Sompolinsky, Computational neuroscience: Beyond the local circuit. *Curr. Opin. Neurobiol.* **25**, xiii–xviii (2014).
6. G. T. Einevoll *et al*., The scientific case for brain simulations. *Neuron* **102**, 735–744 (2019).
7. L. Chariker, R. Shapley, L. S. Young, Orientation selectivity from very sparse LGN inputs in a comprehensive model of macaque V1 cortex. *J. Neurosci.* **36**, 12368–12384 (2016).
8. E. Weinan, B. Engquist, X. Li, W. Ren, E. Vanden-Eijnden, Heterogeneous multiscale methods: A review. *Commun. Comput. Phys.* **2**, 367–450 (2007).
9. I. G. Kevrekidis *et al*., Equation-free, coarse-grained multiscale computation: Enabling microscopic simulators to perform system-level analysis. *Commun. Math. Sci* **1**, 715–762 (2003).
10. C. Beaulieu, Z. Kisvarday, P. Somogyi, M. Cynader, A. Cowey, Quantitative distribution of GABA-immunopositive and-immunonegative neurons and synapses in the monkey striate cortex (area 17). *Cereb. Cortex* **2**, 295–309 (1992).
11. E. M. Callaway, Local circuits in primary visual cortex of the macaque monkey. *Annu. Rev. Neurosci.* **21**, 47–74 (1998).
12. D. Fitzpatrick, J. S. Lund, G. G. Blasdel, Intrinsic connections of macaque striate cortex: Afferent and efferent connections of lamina 4C. *J. Neurosci.* **5**, 3329–3349 (1985).
13. C. Holmgren, T. Harkany, B. Svennenfors, Y. Zilberter, Pyramidal cell communication within local networks in layer 2/3 of rat neocortex. *J. Physiol.* **551**, 139–153 (2003).
14. J. S. Lund, T. Yoshioka, J. B. Levitt, Comparison of intrinsic connectivity in different areas of macaque monkey cerebral cortex. *Cereb. Cortex* **3**, 148–162 (1993).
15. J. S. Lund, A. Angelucci, P. C. Bressloff, Anatomical substrates for functional columns in macaque monkey primary visual cortex. *Cereb. Cortex* **13**, 15–24 (2003).
16. R. B. Tootell, M. S. Silverman, R. L. De Valois, G. H. Jacobs, Functional organization of the second cortical visual area in primates. *Science* **220**, 737–739 (1983).
17. T. D. Albright, R. Desimone, C. G. Gross, Columnar organization of directionally selective cells in visual area MT of the macaque. *J. Neurophysiol.* **51**, 16–31 (1984).
18. H. L. Read, J. A. Winer, C. E. Schreiner, Modular organization of intrinsic connections associated with spectral tuning in cat auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 8042–8047 (2001).
19. J. Lübke, V. Egger, B. Sakmann, D. Feldmeyer, Columnar organization of dendrites and axons of single and synaptically coupled excitatory spiny neurons in layer 4 of the rat barrel cortex. *J. Neurosci.* **20**, 5300–5311 (2000).
20. M. Connolly, D. Van Essen, The representation of the visual field in parvicellular and magnocellular layers of the lateral geniculate nucleus in the macaque monkey. *J. Comp. Neurol.* **226**, 544–564 (1984).
21. S. R. De Groot, P. Mazur, *Non-Equilibrium Thermodynamics* (North-Holland, 1962).
22. L. S. Young, Towards a mathematical model of the brain. *J. Stat. Phys.* **180**, 612–629 (2020).
23. C. Van Vreeswijk, H. Sompolinsky, Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* **274**, 1724–1726 (1996).
24. C. van Vreeswijk, H. Sompolinsky, Chaotic balanced state in a model of cortical circuits. *Neural Comput.* **10**, 1321–1371 (1998).
25. D. H. Hubel, T. N. Wiesel, Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* **160**, 106–154 (1962).
26. L. Chariker, R. Shapley, L. S. Young, Contrast response in a comprehensive network model of macaque V1. *J. Vision* **20**, 16–16 (2020).
27. L. Chariker, R. Shapley, M. Hawken, L. S. Young, A computational model of direction selectivity in macaque V1 cortex based on dynamic differences between on and off pathways. *J. Neurosci.* **42**, 3365–3380 (2022).
28. D. H. Hubel, *Eye, Brain, and Vision* (Scientific American Library/Scientific American Books, 1995).
29. M. Carandini, Area V1. *Scholarpedia* **7**, 12105 (2012).
30. L. C. Sincich, J. C. Horton, The circuitry of V1 and V2: Integration of color, form, and motion. *Annu. Rev. Neurosci.* **28**, 303–326 (2005).
31. D. L. Adams, J. C. Horton, A precise retinotopic map of primate striate cortex generated from the representation of angioscotomas. *J. Neurosci.* **23**, 3771–3789 (2003).
32. D. McLaughlin, A neuronal network model of sharpening and dynamics of orientation tuning in an input layer of macaque primary visual cortex. *Proc. Natl. Acad. Sci. USA* **97**, 8087–8092 (2000).
33. A. V. Rangan, L. S. Young, Dynamics of spiking neurons: Between homogeneity and synchrony. *J. Comput. Neurosci.* **34**, 433–460 (2013).
34. D. L. Ringach, R. M. Shapley, M. J. Hawken, Orientation selectivity in macaque V1: Diversity and laminar dependence. *J. Neurosci.* **22**, 5639–5651 (2002).
35. J. P. Jones, L. A. Palmer, The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *J. Neurophysiol.* **58**, 1187–1211 (1987).
36. G. C. DeAngelis, I. Ohzawa, R. Freeman, Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. I. General characteristics and postnatal development. *J. Neurophysiol.* **69**, 1091–1117 (1993).
37. K. Obermayer, G. G. Blasdel, Geometry of orientation and ocular dominance columns in monkey striate cortex. *J. Neurosci.* **13**, 4114–4129 (1993).
38. T. N. Wiesel, D. H. Hubel, Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. *J. Neurophysiol.* **29**, 1115–1156 (1966).
39. E. Kaplan, R. Shapley, X and Y cells in the lateral geniculate nucleus of macaque monkeys. *J. Physiol.* **330**, 125–143 (1982).
40. T. Hicks, B. Lee, T. Vidyasagar, The responses of cells in macaque lateral geniculate nucleus to sinusoidal gratings. *J. Physiol.* **337**, 183–200 (1983).
41. A. Derrington, P. Lennie, Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. *J. Physiol.* **357**, 219–240 (1984).
42. M. V. Tsodyks, W. E. Skaggs, T. J. Sejnowski, B. L. McNaughton, Paradoxical effects of external modulation of inhibitory interneurons. *J. Neurosci.* **17**, 4382–4388 (1997).
43. Y. Li, L. Chariker, L. S. Young, How well do reduced models capture the dynamics in models of interacting neurons? *J. Math. Biol.* **78**, 83–115 (2019).
44. B. W. Knight, Dynamics of encoding in a population of neurons. *J. Gen. Physiol.* **59**, 734–766 (1972).
45. H. R. Wilson, J. D. Cowan, Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* **12**, 1–24 (1972).
46. H. R. Wilson, J. D. Cowan, A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* **13**, 55–80 (1973).
47. G. B. Ermentrout, J. D. Cowan, A mathematical theory of visual hallucination patterns. *Biol. Cybern.* **34**, 137–150 (1979).
48. P. C. Bressloff, "Lecture notes on mathematical modelling in the life sciences" in *Waves in Neural Media: From Single Neurons to Neural Fields*, A. Stevens, M. C. Mackey, Eds. (Springer, New York, 2014), pp. xx+436 .
49. S. Coombes, P. beim Graben, R. Potthast, J. Wright, *Neural Fields: Theory and Applications* (Springer, 2014).
50. D. Q. Nykamp, D. Tranchina, A population density approach that facilitates large-scale modeling of neural networks: Analysis and an application to orientation tuning. *J. Comput. Neurosci.* **8**, 19–50 (2000).
51. D. Cai, L. Tao, M. Shelley, D. W. McLaughlin, An effective kinetic representation of fluctuation-driven neuronal networks with application to simple and complex cells in visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 7757–7762 (2004).
52. M. A. Buice, J. D. Cowan, Statistical mechanics of the neocortex. *Prog. Biophys. Mol. Biol.* **99**, 53–86 (2009).
53. A. V. Rangan, Diagrammatic expansion of pulse-coupled network dynamics. *Phys. Rev. Lett.* **102**, 158101 (2009).
54. C. R. Laing, T. Frewen, I. G. Kevrekidis, Reduced models for binocular rivalry. *J. Comput. Neurosci.* **28**, 459–476 (2010).
55. D. McLaughlin, R. Shapley, M. Shelley, Large-scale modeling of the primary visual cortex: Influence of cortical architecture upon neuronal response. *J. Physiol. Paris* **97**, 237–252 (2003).
56. M. Shelley, D. McLaughlin, Coarse-grained reduction and analysis of a network model of cortical response: I. Drifting grating stimuli. *J. Comput. Neurosci.* **12**, 97–122 (2002).
57. B. A. Richards *et al*., A deep learning framework for neuroscience. *Nat. Neurosci.* **22**, 1761–1770 (2019).
58. J. Pathak, B. Hunt, M. Girvan, Z. Lu, E. Ott, Model-free prediction of large spatiotemporally chaotic systems from data: A reservoir computing approach. *Phys. Rev. Lett.* **120**, 024102 (2018).
59. L. Lu, P. Jin, G. Pang, Z. Zhang, G. E. Karniadakis, Learning nonlinear operators via deeponet based on the universal approximation theorem of operators. *Nat. Mach. Intell.* **3**, 218–229 (2021).
60. Z. Li *et al*., Fourier neural operator for parametric partial differential equations. arXiv [Preprint] (2020). http://arxiv.org/abs/2010.08895 (Accessed 17 May 2021).
61. T. Bolton, L. Zanna, Applications of deep learning to ocean data inference and subgrid parameterization. *J. Adv. Model. Earth Syst.* **11**, 376–399 (2019).
62. A. J. Chorin, F. Lu, Discrete approach to stochastic parametrization and dimension reduction in nonlinear dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 9804–9809 (2015).
63. Z. C. Xiao, NYU-EffiModel. GitHub. https://github.com/Texense/NYU-EffiModel.git. Accessed 12 June 2024.