Hierarchical timescales in the neocortex: Mathematical mechanism and biological insights

Songting Li[a,b,c,1] **and Xiao-Jing Wang**[d,1]

[a]School of Mathematical Sciences, Shanghai Jiao Tong University, Shanghai 200240, China; [b]Institute of Natural Sciences, Shanghai Jiao Tong University, Shanghai 200240, China; [c]Ministry of Education Key Laboratory of Scientific and Engineering Computing, Shanghai Jiao Tong University, Shanghai 200240, China; and [d]Center for Neural Science, New York University, New York, NY 10003

Edited by Terrence Sejnowski, Computational Neurobiology Laboratory, Salk Institute for Biological Studies, La Jolla, CA; received June 3, 2021; accepted December 7, 2021

A cardinal feature of the neocortex is the progressive increase of the spatial receptive fields along the cortical hierarchy. Recently, theoretical and experimental findings have shown that the temporal response windows also gradually enlarge, so that early sensory neural circuits operate on short timescales whereas higher-association areas are capable of integrating information over a long period of time. While an increased receptive field is accounted for by spatial summation of inputs from neurons in an upstream area, the emergence of timescale hierarchy cannot be readily explained, especially given the dense interareal cortical connectivity known in the modern connectome. To uncover the required neurobiological properties, we carried out a rigorous analysis of an anatomically based large-scale cortex model of macaque monkeys. Using a perturbation method, we show that the segregation of disparate timescales is defined in terms of the localization of eigenvectors of the connectivity matrix, which depends on three circuit properties: 1) a macroscopic gradient of synaptic excitation, 2) distinct electrophysiological properties between excitatory and inhibitory neuronal populations, and 3) a detailed balance between long-range excitatory inputs and local inhibitory inputs for each area-to-area pathway. Our work thus provides a quantitative understanding of the mechanism underlying the emergence of timescale hierarchy in large-scale primate cortical networks.

large-scale cortical network | timescale hierarchy | eigenvector localization | interareal heterogeneity | detailed excitation–inhibition balance of long-range cortical connections

The brain is organized with a delicate structure to integrate and process both spatial and temporal information received from the external world. For spatial information processing, neurons along cortical visual pathways possess increasingly large spatial receptive fields, and its underlying mechanism has been understood as neurons in higher-level visual areas receive input from many neurons with smaller receptive fields in lower-level visual areas, thereby aggregating information across space (1). More recently, a computational model (2) revealed that the timescale over which neural integration occurs also gradually increases from area to area along the cortical hierarchy. The model was based on the anatomically measured directed- and weighted-interareal connectivity of the macaque cortex (3) and incorporated heterogeneity of synaptic excitation calibrated by spine count per pyramidal neuron (4). It has been observed that the decay times increased progressively along the cortical hierarchy when signals propagate in the network, and the temporal hierarchy could change dynamically in response to different types of sensory inputs (e.g., different hierarchy of timescales for somatosensory input versus visual input) (2). By manipulating parameters of the model, simulation results further demonstrated that both within and between regions of anatomical properties could affect the hierarchy of timescales in neuronal population activity (2). A hierarchy of temporal receptive windows is functionally desirable, so that the circuit dynamics operate on short timescales in early sensory areas to encode and process rapidly changing external stimuli, whereas parietal and frontal areas canaccumulate information over a relatively long period of time in decision-making and other cognitive processes (5, 6).

Despite the accumulating evidence in support of timescale hierarchy across cortical areas in mice (7, 8), monkeys (9–15), and humans (16–23), its underlying mechanism remains unclear. In particular, since interareal connections are dense, with roughly 65% of all possible connections present in the macaque cortex (3) and even higher connection density in the mouse cortex (24), what circuit properties are required to ensure that dynamical modes with disparate time constants are spatially localized? How do intraareal anatomical properties determine the intrinsic timescale of each area, and how do these intrinsic timescales remain to be segregated rather than mixed up in the presence of dense interareal connections? In this work, we addressed these questions by a mathematical analysis of the model (2). Using a perturbation method, we identified key required conditions, in particular a detailed excitation–inhibition balance for long-distance interareal connections that is experimentally testable.

The Multiareal Model and Hierarchical Timescales Phenomenon

We first review the mathematical form of the multiareal model of the macaque cortex and the hierarchical timescales phenomenon captured by this model (2). The macaque cortical network model

Significance

In the neocortex, while early sensory areas encode and process external inputs rapidly, higher-association areas are endowed with slow dynamics suitable for accumulating information over time. Such a hierarchy of temporal response windows along the cortical hierarchy naturally emerges in a model of multiareal primate cortex. This finding raises the question of why diverse temporal modes are not mixed in roughly the same way across the whole cortex, despite high connection density and an abundance of feedback loops. We investigate this question by mathematically analyzing the anatomically based network model of macaque cortex and theoretically show that three sufficient conditions of synaptic excitation and inhibition give rise to timescale segregation in a hierarchy, a functionally important characteristic of the cortex.

Author contributions: S.L. and X.-J.W. designed research; S.L. performed research; and S.L. and X.-J.W. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1] To whom correspondence may be addressed. Email: songting@sjtu.edu.cn or xjwang@nyu.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2110274119/-/DCSupplemental.

Published Febuary 2, 2022.

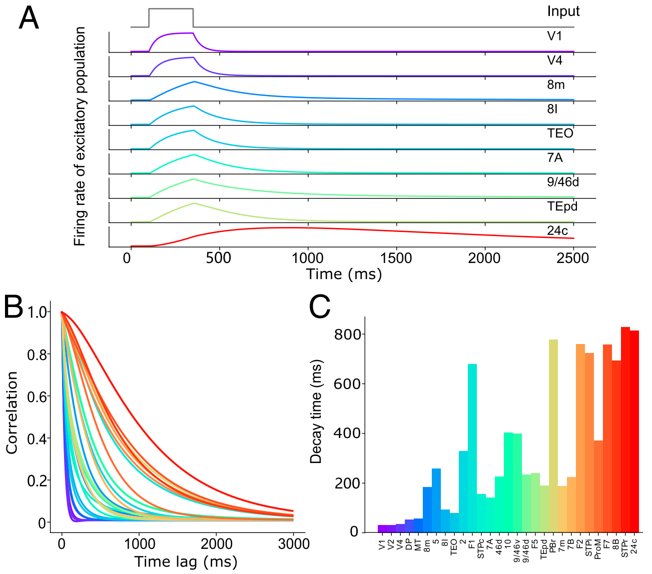

Fig. 1. The hierarchical timescales phenomenon simulated in the macaque multiareal model. (*A*) A pulse of input to area V1 is propagated along the hierarchy, displaying increasing decay times as it proceeds. (*B*) Autocorrelation of area activity in response to white-noise input to V1. (*C*) The dominant time constants in all areas, extracted by fitting single or double exponentials to the autocorrelation curves (2). In *A*–*C*, areas are arranged and colored by position in the anatomical hierarchy.

contains a subnet of 29 areas widely distributed from sensory to association areas in the macaque cortex, and each area includes both excitatory and inhibitory neuronal populations. The neuronal population dynamics in the *i*th area are described as

$$\tau_E \frac{d}{dt}\nu_E^i = -\nu_E^i + \beta_E\left[I_{syn,E}^i + I_{ext,E}^i\right]_+, \qquad \textbf{[1]}$$

$$\tau_I \frac{d}{dt}\nu_I^i = -\nu_I^i + \beta_I\left[I_{syn,I}^i + I_{ext,I}^i\right]_+, \qquad \textbf{[2]}$$

where ν_E^i and ν_I^i are the firing rate of the excitatory and inhibitory populations in the *i*th area, respectively; τ_E and τ_I are their time constants, respectively; and β_E and β_I are the slope of the frequency–current (f-I) curve for the excitatory and inhibitory populations, respectively. The f-I curve takes the form of a rectified linear function with $[I]_+ = max(I, 0)$. In addition, $I_{ext,E}^i$ and $I_{ext,I}^i$ are the external currents, and $I_{syn,E}^i$ and $I_{syn,I}^i$ are the synaptic currents that follow

$$I_{syn,E}^i = (1+\eta h_i)\Big(w_{EE}\nu_E^i + \mu_{EE}\sum_{j=1}^{N} FLN_{ij}\nu_E^j\Big) - w_{EI}\nu_I^i,$$

$$I_{syn,I}^i = (1+\eta h_i)\Big(w_{IE}\nu_E^i + \mu_{IE}\sum_{j=1}^{N} FLN_{ij}\nu_E^j\Big) - w_{II}\nu_I^i,$$

where w_{pq}, $p, q \in \{E, I\}$ is the local coupling strength from the q population to the p population within each area. FLN_{ij} is the fraction of labeled neurons (FLN) from area *j* to area *i* reflecting the strengths of long-range input (3), and μ_{EE} and μ_{IE} are scaling parameters that control the strengths of long-range input to the excitatory and inhibitory populations, respectively. Both local and long-range excitatory inputs to an area are scaled by its position in the hierarchy quantified by h_i (a value normalized between 0 and 1), based on the observation that the hierarchical position of an area highly correlates with the number of spines on pyramidal neurons in that area (2, 4). A constant η maps the hierarchy h_i into excitatory connection strengths. Note that both local and long-range projections are scaled by hierarchy, rather than just local projections, following the observation that the proportion of local to long-range connections is approximately conserved across areas (25). The values of all the model parameters are specified in *Materials and Methods*.

By simulating the model, it has been observed in ref. 2 that the decay time of neuronal response in each area increases progressively along the visual cortical hierarchy when a pulse input is given to area V1, as shown here in Fig. 1*A*. Early visual areas show fast and transient responses while prefrontal areas show slower responses and longer integration times with traces lasting for several seconds after the stimulation. In addition, white-noise input to V1 is also integrated with a hierarchy of timescales by computing the autocorrelation of neuronal activity at each area (2). As shown in Fig. 1*B*, the activity of early sensory areas shows rapid decay of autocorrelation with time lag while that of association areas shows slow decay. In Fig. 1*C*, by fitting single or double exponentials to the decay of the autocorrelation curves (2), the dominant timescale of each area tends to increase along the hierarchy approximately, and thus a hierarchy of widely disparate timescales emerges from this model. It is worth noting, however, that the timescale does not change monotonically with the anatomically defined hierarchy (*x* axis); the precise pattern is sculpted by the measured interareal wiring properties.

Note that, although the multiareal model (Eqs. **1** and **2**) is nonlinear by taking into account a rectified linear f-I curve, the stimuli in our simulations drive all neuronal population activities above the firing threshold with positive input currents to all areas. Therefore, the stimuli essentially drive the network dynamics into the linear regime. Before we perform mathematical analysis to understand the mechanism underlying the emergence of hierarchical timescales in the simulations, to simplify the notation, we rewrite the network dynamics Eqs. **1** and **2** in the linear regime in the form

$$\frac{d}{dt}\boldsymbol{\nu} = W\boldsymbol{\nu} + \boldsymbol{I}_{ext}, \qquad \textbf{[3]}$$

where

$$\boldsymbol{\nu} = \left[\nu_E^1, \ldots, \nu_E^n, \nu_I^1, \ldots, \nu_I^n\right]^T, \quad W = \begin{bmatrix} D_{EE} + F_{EE} & D_{EI} \\ D_{IE} + F_{IE} & D_{II} \end{bmatrix},$$

with $n = 29$, and D_{EE}, D_{EI}, D_{IE}, D_{II} being four diagonal matrices whose *i*th element on their diagonal line is

$$d_{EE}^i = \frac{\beta_E}{\tau_E}\left[(1+\eta h_i)w_{EE} - \frac{1}{\beta_E}\right], \; d_{EI}^i = -\frac{\beta_E}{\tau_E}\left[w_{EI}\right],$$

$$d_{IE}^i = \frac{\beta_I}{\tau_I}\left[(1+\eta h_i)w_{IE}\right], \; d_{II}^i = -\frac{\beta_I}{\tau_I}\left[w_{II} + \frac{1}{\beta_I}\right],$$

respectively, and matrices F_{EE} and F_{IE} being two nondiagonal matrices whose *i*th-row–*j*th-column element is

$$f_{EE}^{ij} = \frac{\beta_E}{\tau_E}\left[(1+\eta h_i)\mu_{EE}FLN_{ij}\right], f_{IE}^{ij} = \frac{\beta_I}{\tau_I}\left[(1+\eta h_i)\mu_{IE}FLN_{ij}\right],$$

respectively. Note that matrices D_{EE}, D_{EI}, D_{IE}, D_{II} reflect local intraareal interactions, while matrices F_{EE} and F_{IE} reflect long-range interareal interactions. In addition, elements in D_{EE}, D_{IE}, F_{EE}, and F_{IE} depend on area hierarchy h_i while elements in D_{EI} and D_{II} are constant. Finally, the external input vector is

$$\boldsymbol{I}_{ext} = \left[\frac{\beta_E}{\tau_E}I_{ext,E}^1, \ldots, \frac{\beta_E}{\tau_E}I_{ext,E}^n, \frac{\beta_I}{\tau_I}I_{ext,I}^1, \ldots, \frac{\beta_I}{\tau_I}I_{ext,I}^n\right]^T.$$

Denoting the eigenvalues and eigenvectors of the connectivity matrix W as λ_i and $\boldsymbol{V}_i$ $(i = 1, 2, \ldots, 2n)$, respectively, i.e., $W\boldsymbol{V}_i = \lambda_i\boldsymbol{V}_i$, the analytical solution of Eq. **3** can be obtained as

$$\boldsymbol{\nu}^i(t) = \sum_{j=1}^{2n}\left(\tilde{a}_j e^{\lambda_j t} + \int_0^t e^{\lambda_j(t-t')}\tilde{I}_j(t')dt'\right)\boldsymbol{V}_j^i, \qquad \textbf{[4]}$$

where $\boldsymbol{\nu}^i$ and $\boldsymbol{V}_j^i$ are the ith element in $\boldsymbol{\nu}$ and $\boldsymbol{V}_j$, respectively, and $\tilde{a}_j$ and $\tilde{I}_j$ are the coefficients for the initial condition and the external input, respectively, represented in the coordinate system of the eigenvectors $\{\boldsymbol{V}_j\}$. Note that, from Eq. **4**, each area integrates input current with the same set of time constants $\{\tau_i\}$ determined by the real part of the eigenvalues, i.e., $\tau_i = -1/Re\{\lambda_i\}$. Therefore, the characteristic timescale of each area across the network is expected to be similar in the general case. To obtain distinct timescales at each area, it requires 1) the localization of eigenvectors $\boldsymbol{V}_j$, i.e., most of the elements in $\boldsymbol{V}_j$ are close to zero, and 2) the orthogonality of all pairs of eigenvectors, i.e., the nonzero elements nearly nonoverlap for different $\boldsymbol{V}_j$.

By computing the eigenvalues and eigenvectors of matrix W, as shown in Fig. 2*A*, the timescale pool $\{\tau_i\}$ derived from eigenvalues can be classified into two groups; one group shows a quite fast timescale of about 2 ms, and the other group includes relatively slow timescales ranging from tens to hundreds of milliseconds. In addition, we are particularly interested in the excitatory population because the majority of neurons in the cortex are excitatory neurons. We observe that the magnitude of the eigenvectors corresponding to the fast timescale is nearly zero for the excitatory population at each area, while that corresponding to the slow timescale is weakly localized and weakly orthogonal, i.e., each eigenvector has a few nonzero elements that almost do not overlap with other eigenvectors' nonzero elements. According to Eq. **4**, the pattern of eigenvectors gives rise to the disparate timescales for the excitatory neuronal population at each cortical area. We next perform mathematical analysis to investigate the sufficient conditions for 1) vanishing magnitude of the excitatory component of fast-eigenmode eigenvectors and 2) weak localization and orthogonality of the excitatory component of slow-eigenmode eigenvectors in this network system.

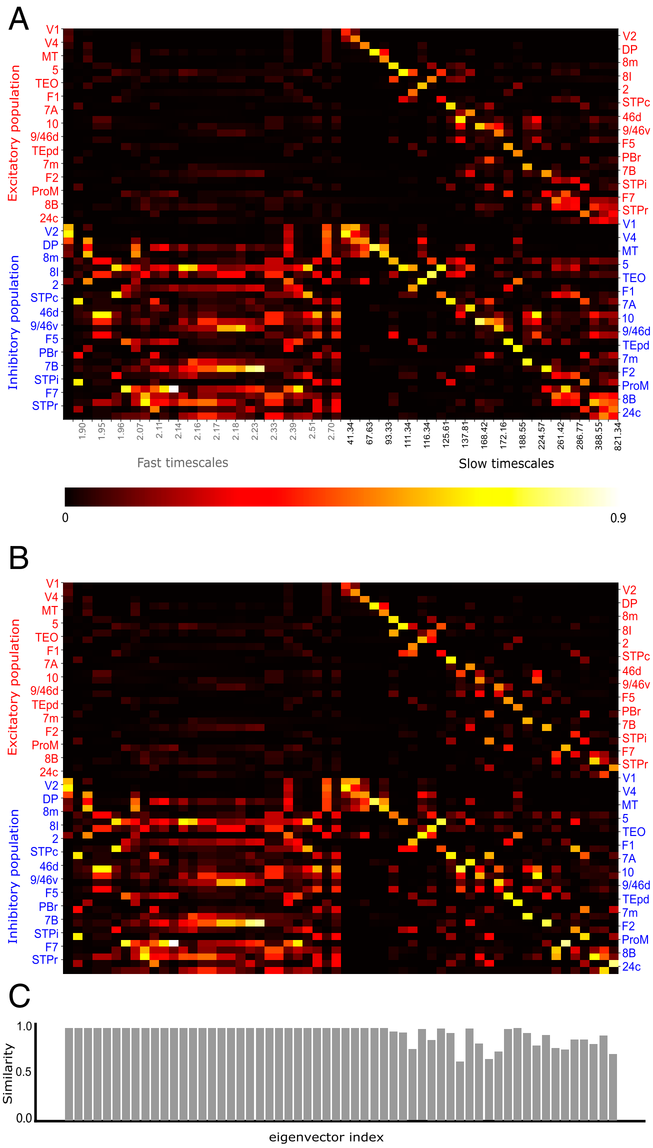

Fig. 2. Eigenvectors of the network connectivity matrix and their approximations from the perturbation analysis. (*A*) Eigenvectors of the network connectivity matrix W. Each column shows the amplitude of an eigenvector at the 29 areas, with corresponding timescale labeled below. (*B*) Eigenvectors of W calculated from the first-order perturbation analysis. (*C*) Similarity measure defined as the inner product of the corresponding eigenvectors in *A* and *B*.

Perturbation Analysis of the Model

We note that the parameters of the model (specified in *Materials and Methods*) give

$$\epsilon = \frac{\beta_E}{\tau_E} \Big/ \frac{\beta_I}{\tau_I} \approx 0.094, \quad \delta = \frac{\mu_{EE}}{\mu_{IE}} - \frac{w_{EI}}{w_{II} + 1/\beta_I} \approx 0.038,$$

which can be viewed as two small parameters to allow us to perform perturbation analysis below.

We first study the network in the absence of the long-range interactions among areas. In this scenario, we study the 2 × 2 block matrix $D = \begin{bmatrix} D_{EE} & D_{EI} \\ D_{IE} & D_{II} \end{bmatrix}$ in which each block is a diagonal matrix defined above. By viewing $\epsilon = \frac{\beta_E}{\tau_E} / \frac{\beta_I}{\tau_I}$ as a small parameter, we have

$$D_{II}, D_{IE} \sim \mathcal{O}(1); \;\; D_{EI}, D_{EE} \sim \mathcal{O}(\epsilon) \qquad [5]$$

from their definitions. Accordingly, we can prove the following proposition:

Proposition 1. *If* $D_{II} \sim \mathcal{O}(1)$, $D_{IE} \sim \mathcal{O}(1)$, $D_{EI} \sim \mathcal{O}(\epsilon)$, $D_{EE} \sim \mathcal{O}(\epsilon)$, *then D has n eigenvalues being* $\mathcal{O}(\epsilon)$ *and n eigenvalues being* $\mathcal{O}(1)$.

Proof. It is straightforward to prove that matrix D can be diagonalized by matrix P; i.e.,

$$\Lambda = P^{-1}DP = \begin{bmatrix} \Lambda_U & O \\ O & \Lambda_L \end{bmatrix} = \begin{bmatrix} D_{EE} + AD_{IE} & O \\ O & -D_{IE}A + D_{II} \end{bmatrix},$$

where

$$P = \begin{bmatrix} I + AB & -A \\ -B & I \end{bmatrix},$$

I is the identity matrix, and diagonal matrix A satisfies $-(D_{EE} + AD_{IE})A + D_{EI} + AD_{II} = 0$ and diagonal matrix B satisfies $D_{IE} + B(D_{EE} + AD_{IE}) + (D_{IE}A - D_{II})B = 0$.

We solve the equation of A and choose one of the two solutions of A as

$$A = \frac{1}{2} D_{IE}^{-1} \left[D_{II} - D_{EE} + \sqrt{(D_{II} - D_{EE})^2 + 4D_{IE}D_{EI}} \right],$$

where the square root of a diagonal matrix is defined as taking the square root of its elements. Due to the fact that $D_{II} \sim \mathcal{O}(1)$, $D_{IE} \sim \mathcal{O}(1)$, $D_{EI} \sim \mathcal{O}(\epsilon)$, $D_{EE} \sim \mathcal{O}(\epsilon)$, we have $A = -D_{EI}D_{II}^{-1} + \mathcal{O}(\epsilon^2) \sim \mathcal{O}(\epsilon)$, and $B = D_{IE}D_{II}^{-1} + \mathcal{O}(\epsilon) \sim \mathcal{O}(1)$. Accordingly, we have

$$\Lambda_U = D_{EE} - D_{EI} D_{II}^{-1} D_{IE} + \mathcal{O}(\epsilon^2) \sim \mathcal{O}(\epsilon),$$

$$\Lambda_L = D_{II} + D_{EI} D_{II}^{-1} D_{IE} + \mathcal{O}(\epsilon^2) \sim \mathcal{O}(1). \qquad \blacksquare$$

From *Proposition 1*, the eigenvalues of D have two separated scales belonging to Λ_U and Λ_L, respectively. As the timescales of the network system are given by $\tau_i = -1/Re\{\lambda_i\}$ (λ_i is the ith diagonal element in matrix Λ), the separation of scales for eigenvalues in Λ_U and Λ_L explains that the intrinsic timescale pool can be classified into two groups with a separation of scales, which is mainly determined by the distinct electrophysiological properties between excitatory and inhibitory neuronal populations within each area described by ϵ. In addition, from the analysis, the eigenvalues in Λ_L with large magnitude (fast timescale) are less sensitive to the hierarchy level because $\Lambda_L \approx D_{II}$, and the elements in D_{II} do not depend on h_i. Therefore, the gradient of h_i across areas barely affects the fast-timescale pool. In contrast, the eigenvalues in Λ_U with small magnitude (slow timescale) are more sensitive to the hierarchy level because $\Lambda_U \approx D_{EE} - D_{EI} D_{II}^{-1} D_{IE}$, and both the elements in D_{EE} and D_{IE} depend on h_i. Therefore, the gradient of h_i across areas increases the range of the slow-timescale pool. Further, the slow timescales of each area in this disconnected network are segregated and follow the hierarchical order h_i as the corresponding eigenvectors are perfectly localized and orthogonal to each other.

Now we consider the multiareal network in the presence of long-range interactions. Adding long-range connectivity to local connectivity matrix D changes the eigenvalues and eigenvectors of matrix D, which can be analyzed in the following.

By multiplying P and P^{-1} (given in the *Proof* of *Proposition 1*) on both sides of W, we have

$$\Gamma = P^{-1} W P = P^{-1} \left(\begin{bmatrix} D_{EE} & D_{EI} \\ D_{IE} & D_{II} \end{bmatrix} + \begin{bmatrix} F_{EE} & O \\ F_{IE} & O \end{bmatrix} \right) P = \Lambda + \Sigma,$$

where

$$\Sigma = P^{-1} \begin{bmatrix} F_{EE} & O \\ F_{IE} & O \end{bmatrix} P = \begin{bmatrix} \Sigma_{UL} & \Sigma_{UR} \\ \Sigma_{LL} & \Sigma_{LR} \end{bmatrix},$$

with $\Sigma_{UL} = (F_{EE} + AF_{IE})(I + AB), \Sigma_{UR} = -(F_{EE} + AF_{IE})A$, $\Sigma_{LL} = \left[BF_{EE} + (I + BA)F_{IE}\right](I + AB)$, $\Sigma_{LR} = -\left[BF_{EE} + (I + BA)F_{IE}\right]A$ and with Λ, A, and B defined in the *Proof* of *Proposition 1*.

Denoting one of the eigenvectors of matrix Γ as $[\boldsymbol{u}, \boldsymbol{v}]^T$ and the corresponding eigenvalue as λ, we have

$$\left(\begin{bmatrix} \Lambda_U & 0 \\ 0 & \Lambda_L \end{bmatrix} + \begin{bmatrix} \Sigma_{UL} & \Sigma_{UR} \\ \Sigma_{LL} & \Sigma_{LR} \end{bmatrix} \right) \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix} = \lambda \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix}. \qquad \textbf{[6]}$$

According to the definition of $\epsilon = \frac{\beta_E}{\tau_E} / \frac{\beta_I}{\tau_I}$ and $\delta = \frac{\mu_{EE}}{\mu_{IE}} - \frac{w_{EI}}{w_{II} + 1/\beta_I}$, $A \sim \mathcal{O}(\epsilon)$, $B \sim \mathcal{O}(1)$ and $F_{EE} \sim \mathcal{O}(\epsilon)$, $F_{IE} \sim \mathcal{O}(1)$, we have $F_{EE} + AF_{IE} \sim \mathcal{O}(\epsilon\delta)$, $I + AB \sim \mathcal{O}(1)$, $BF_{EE} + (I + BA)F_{IE} \sim \mathcal{O}(1)$, and accordingly,

$$\Sigma_{UL} \sim \mathcal{O}(\epsilon\delta), \quad \Sigma_{UR} \sim \mathcal{O}(\epsilon^2\delta), \; \Sigma_{LL} \sim \mathcal{O}(1), \; \Sigma_{LR} \sim \mathcal{O}(\epsilon).$$

As $\Sigma_{UR} \sim \mathcal{O}(\epsilon^2\delta)$ is a higher-order term compared with Λ_U, Λ_L, Σ_{UL}, Σ_{LL}, and Σ_{LR}, it can be dropped out in Eq. **6** and the error of eigenvalue and eigenvector is at most $\mathcal{O}(\epsilon^2\delta)$ (see *SI Appendix, Proposition S1* for a detailed proof). Consequently, we can obtain two equations from Eq. **6** in the vector form,

$$(\Lambda_U + \Sigma_{UL})\boldsymbol{u} = \lambda \boldsymbol{u}, \qquad \textbf{[7]}$$

$$\Sigma_{LL}\boldsymbol{u} + (\Lambda_L + \Sigma_{LR})\boldsymbol{v} = \lambda \boldsymbol{v}. \qquad \textbf{[8]}$$

To describe the eigenvector property of Eqs. **7** and **8**, we first introduce the definitions of weak localization and weak orthogonality as follows:

Definition 1. *A vector $\boldsymbol{u}(\delta)$ is weakly localized if it can be represented as $\boldsymbol{u} = a\boldsymbol{e}_k + \delta\boldsymbol{b} + \mathcal{O}(\delta^2)$ for some k, where $a \sim \mathcal{O}(1)$ is a constant number, $\boldsymbol{b} \sim \mathcal{O}(1)$ is a constant vector, δ is a small parameter, and $\boldsymbol{e}_k$ represents the natural basis with only the kth element being 1 and others being zero, i.e., $\boldsymbol{e}_k = [0, \ldots, 1(kth), \ldots, 0]$.*

Definition 2. *Two vectors $\boldsymbol{u}(\delta)$ and $\boldsymbol{v}(\delta)$ are weakly orthogonal to each other if their inner product $< \boldsymbol{u}, \boldsymbol{v} > \sim \mathcal{O}(\delta)$, where δ is a small parameter.*

With the concept of weak localization and weak orthogonality defined above, we introduce the following proposition that describes the property of $\boldsymbol{u}$ in the system of Eqs. **7** and **8**:

Proposition 2. *In the system described by Eqs.* **7** *and* **8**, *if all matrices are analytic with respect to ϵ and δ; and if $\Sigma_{UL} \sim \mathcal{O}(\epsilon\delta)$, $\Sigma_{LL} \sim \mathcal{O}(1)$, $\Sigma_{LR} \sim \mathcal{O}(\epsilon)$, $\Lambda_U \sim \mathcal{O}(\epsilon)$, $\Lambda_L \sim \mathcal{O}(1)$; and if Λ_U has n simple eigenvalues; then*

1) *there exist n eigenvectors $[\boldsymbol{u}, \boldsymbol{v}]^T$ in which $\boldsymbol{u} = \boldsymbol{0}$, with $\lambda \sim \mathcal{O}(1)$ correspondingly, and*
2) *there exist n eigenvectors $[\boldsymbol{u}, \boldsymbol{v}]^T$ in which $\boldsymbol{u}$ is weakly localized and weakly orthogonal to each other, with $\lambda \sim \mathcal{O}(\epsilon)$ correspondingly.*

Proof. 1) It is noted that $\boldsymbol{u} = \boldsymbol{0}$ is a trivial solution of Eq. **7**. By defining $\bar{\Sigma}_{LR} = \Sigma_{LR}/\epsilon \sim \mathcal{O}(1)$, Eq. **8** becomes

$$(\Lambda_L + \epsilon\bar{\Sigma}_{LR})\boldsymbol{v} = \lambda\boldsymbol{v},$$

in which $\boldsymbol{v}$ is the eigenvector of matrix $\Lambda_L + \epsilon\bar{\Sigma}_{LR}$. By viewing $\epsilon\bar{\Sigma}_{LR}$ as a perturbation matrix to Λ_L, then the leading order of λ shall be the same as that of n elements in the diagonal line of Λ_L, which takes the order of $\mathcal{O}(1)$.

2) In Eq. **7**, if $\boldsymbol{u} \neq \boldsymbol{0}$, by defining $\bar{\Lambda}_U = \Lambda_U/\epsilon$, $\bar{\Sigma}_{UL} = \Sigma_{UL}/\epsilon\delta$, and $\bar{\lambda} = \lambda/\epsilon$, we have

$$(\bar{\Lambda}_U + \delta\bar{\Sigma}_{UL})\boldsymbol{u} = \bar{\lambda}\boldsymbol{u}. \qquad \textbf{[9]}$$

Therefore, $\boldsymbol{u}$ is also the eigenvector of matrix $\bar{\Lambda}_U + \delta\bar{\Sigma}_{UL}$, and $\bar{\lambda}$ is the corresponding eigenvalue. As Λ_U has n simple eigenvalues, so does $\bar{\Lambda}_U$, and then $\boldsymbol{u}$ and $\bar{\lambda}$ are analytic with respect to the perturbation parameter δ (26), i.e., $\boldsymbol{u} = \sum_{i=0}^{\infty} \delta^i \boldsymbol{u}_i$, and $\bar{\lambda} = \sum_{j=0}^{\infty} \delta^j \mu_i$ for δ near zero. Therefore, to the leading order, we have

$$\bar{\Lambda}_U \boldsymbol{u}_0 = \mu_0 \boldsymbol{u}_0,$$

in which μ_0 is the eigenvalue of the diagonal matrix $\bar{\Lambda}_U$, and $\boldsymbol{u}_0$ is the corresponding eigenvector. Accordingly, $\boldsymbol{u}_0 \in \{\boldsymbol{e}_k\}$, and thereafter

$$\boldsymbol{u} = \boldsymbol{e}_k + \delta\boldsymbol{u}_1 + \mathcal{O}(\delta^2), \qquad k = 1, 2, \ldots, n,$$

where $\boldsymbol{e}_k$ represents the kth natural basis, and the leading order of $\lambda = \epsilon\bar{\lambda}$ is $\epsilon\mu_0 \sim \mathcal{O}(\epsilon)$. It is straightforward to verify that $\boldsymbol{u}$ are weakly localized and weakly orthogonal to each other. $\blacksquare$

If we denote the unit-length eigenvector of the connectivity matrix W as $[\boldsymbol{r}_E, \boldsymbol{r}_I]^T$, and denote the corresponding eigenvalue as λ (the same as that of matrix Γ after the similarity transform), then from *Propositions 1* and *2* we have the following:

Proposition 3. *Under the same conditions in Propositions 1 and 2, the unit-length eigenvector $[\boldsymbol{r}_E, \boldsymbol{r}_I]^T$ of the connectivity matrix W has the following properties:*

1) *For eigenvalue $\lambda \sim \mathcal{O}(\epsilon)$, the corresponding $\boldsymbol{r}_E$ is weakly localized and weakly orthogonal to each other, and*
2) *for eigenvalue $\lambda \sim \mathcal{O}(1)$, the corresponding $\boldsymbol{r}_E \sim \mathcal{O}(\epsilon)$.*

Proof. We first consider $[\boldsymbol{r}_E, \boldsymbol{r}_I]^T$ with nonunit length. According to the similarity transform, we have the following linear relation between $[\boldsymbol{r}_E, \boldsymbol{r}_I]^T$ and $[\boldsymbol{u}, \boldsymbol{v}]^T$:

$$\begin{bmatrix} \boldsymbol{r}_E \\ \boldsymbol{r}_I \end{bmatrix} = P \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix} = \begin{bmatrix} I + AB & -A \\ -B & I \end{bmatrix} \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix},$$

i.e., $\boldsymbol{r}_E = (I + AB)\boldsymbol{u} - A\boldsymbol{v}$, and $\boldsymbol{r}_I = -B\boldsymbol{u} + \boldsymbol{v}$. From *Proposition 1*, we have $A \sim \mathcal{O}(\epsilon)$, and $B \sim \mathcal{O}(1)$.

1) From *Proposition 2*, we have $\boldsymbol{u} = \boldsymbol{e}_k + \delta\boldsymbol{u}_1 + \mathcal{O}(\delta^2)$ for $\lambda \sim \mathcal{O}(\epsilon)$ $(k = 1, 2, \ldots, n)$. Accordingly, $\boldsymbol{v}$ can be solved as $\boldsymbol{v} = (\lambda I - \Lambda_L - \Sigma_{LR})^{-1}\Sigma_{LL}\boldsymbol{u} \sim \mathcal{O}(1)$, which gives $\boldsymbol{r}_E \sim \mathcal{O}(1)$, and $\boldsymbol{r}_I \sim \mathcal{O}(1)$. Therefore, the length of $[\boldsymbol{r}_E, \boldsymbol{r}_I]^T$ denoted by c is order $\mathcal{O}(1)$. By normalizing the length of $[\boldsymbol{r}_E, \boldsymbol{r}_I]^T$ to be unity, we have $\boldsymbol{r}_E = c^{-1}(I + AB)\boldsymbol{e}_k + \mathcal{O}(\epsilon) + \mathcal{O}(\delta)$ being weakly localized and weakly orthogonal to each other.
2) From *Proposition 2*, we have $\boldsymbol{u} = \boldsymbol{0}$ for $\lambda \sim \mathcal{O}(1)$. Accordingly, $\boldsymbol{v}$ can be solved as the eigenvector of matrix $(\Lambda_L + \Sigma_{LR})$ with unit length. Therefore, $[\boldsymbol{r}_E, \boldsymbol{r}_I]^T = [-A\boldsymbol{v}, \boldsymbol{v}]^T$, and the length of $[\boldsymbol{r}_E, \boldsymbol{r}_I]^T$ denoted by c is order $\mathcal{O}(1)$. By normalizing the length of $[\boldsymbol{r}_E, \boldsymbol{r}_I]^T$ to be unity, we have $\boldsymbol{r}_E = -c^{-1}A\boldsymbol{v} \sim \mathcal{O}(\epsilon)$. ■

Note that *Propositions 2* and *3* hold for sufficiently small ϵ and δ near zero. However, the convergence radius of the power series of $\boldsymbol{u}$ in *Proposition* 2 is not specified yet. Although difficult to calculate the convergence radius, we can compute the analytical expression of $\boldsymbol{u}_1$ in the power series $\boldsymbol{u} = \sum_{i=0}^{\infty} \delta^i \boldsymbol{u}_i$ in *Proposition 2* to obtain the first-order perturbation solution of $\boldsymbol{u}$ and thereby $\boldsymbol{r}_E$, which could help us gain insight about when weak localization and orthogonality of $\boldsymbol{u}$ and $\boldsymbol{r}_E$ will break down approximately.

To the order of δ in Eq. **9**, we have

$$\bar{\Lambda}_U \boldsymbol{u}_1 + \bar{\Sigma}_{UL}\boldsymbol{u}_0 = \mu_0\boldsymbol{u}_1 + \mu_1\boldsymbol{u}_0. \qquad \textbf{[10]}$$

Without loss of generality, we assume $\boldsymbol{u}_0 = \boldsymbol{e}_k$, and accordingly, $\mu_0 = \bar{\lambda}_k$ is the kth element in the diagonal line of matrix $\bar{\Lambda}_U$. In addition, we normalize $\boldsymbol{u}$ to make the kth element in $\boldsymbol{u}$ denoted by $\boldsymbol{u}^k$ to be unity, and correspondingly $\boldsymbol{u}_i^k = 0$ for $i \geq 1$. By left multiplying $\boldsymbol{e}_j^T$ $(j \neq k)$ to Eq. **10**, we have

$$\boldsymbol{u}_1^j = \frac{\bar{\Sigma}_{UL}^{jk}}{\bar{\lambda}_k - \bar{\lambda}_j}, \qquad \textbf{[11]}$$

where $\bar{\Sigma}_{UL}^{jk}$ is the element in the jth row and kth column in matrix $\bar{\Sigma}_{UL}$. To make the first-order perturbation solution valid, we expect that $\boldsymbol{u}_1^j$ is small compared with δ^{-1}; otherwise the separation of orders will no longer hold in the power series (i.e., first-order term $\delta\boldsymbol{u}_1$ becomes larger than the zeroth-order term $\boldsymbol{e}_k$). In such a case, the spectral gap $\bar{\lambda}_k - \bar{\lambda}_j$ shall be large enough compared with elements in $\bar{\Sigma}_{UL}$. The spectral gap of matrix $\bar{\Lambda}_U$ attributes to the gradient of excitation across areas or simply h_i. In Fig. 2 *B* and *C*, we show that the eigenvector $[\boldsymbol{r}_E, \boldsymbol{r}_I]^T$, which is solved using the perturbation theory in *Propositions 2* and *3* to the first-order accuracy (Eq. **11**), agrees well with the original eigenvector in most cases. However, some eigenvectors show less similarity to the original eigenvectors when the first-order perturbation theory breaks down for the reason discussed above.

Biological Interpretations of the Three Requirements

From the above analysis, three conditions are required to obtain weakly localized and orthogonal eigenvectors to maintain the hierarchy of timescales: 1) small ϵ, 2) small δ, and 3) the gradient of h_i across areas. We briefly summarize the important roles of the three conditions in proving eigenvector localization and orthogonality in the perturbation analysis illustrated in Fig. 3. As shown in Fig. 3 *A* and *B*, to remove intraareal interactions between the excitatory and inhibitory populations within each area, we first change the coordinate system from $(\boldsymbol{r}_E, \boldsymbol{r}_I)$ to $(\boldsymbol{u}, \boldsymbol{v})$ with a transform matrix P given in *Proposition 1*. In the new coordinate system, there is no local interaction between the dynamical variables $\boldsymbol{u}$ and $\boldsymbol{v}$. Furthermore, considering a directed long-range projection from area j to area i in Fig. 3*C*, it has been shown that small δ leads to weak interaction from u^j to u^i, and small ϵ additionally leads to even weaker interaction from v^j to u^i that can be removed with an ignorable error (*SI Appendix*, *Proposition S1*). And a gradient of h_i leads to a nonzero spectral gap between area i and area j. All three conditions result in the weak localization and orthogonality of the $\boldsymbol{u}$ component in the $(\boldsymbol{u}, \boldsymbol{v})$ coordinate system, as proved by *Proposition 2* and Eq. **11**. Finally, as shown in Fig. 3*D*, small ϵ gives rise to the fact that the leading orders of $\boldsymbol{u}$ and $\boldsymbol{r}_E$ are the same, and thus $\boldsymbol{r}_E$ is also weakly localized and orthogonalized similar to $\boldsymbol{u}$, as proved by *Proposition 3*.

We next discuss the biological interpretation of the three conditions. First, according to the definition of $\epsilon = \frac{\beta_E}{\tau_E} / \frac{\beta_I}{\tau_I}$, small ϵ indicates that the electrophysiological properties of excitatory and inhibitory neurons are different, in particular, their

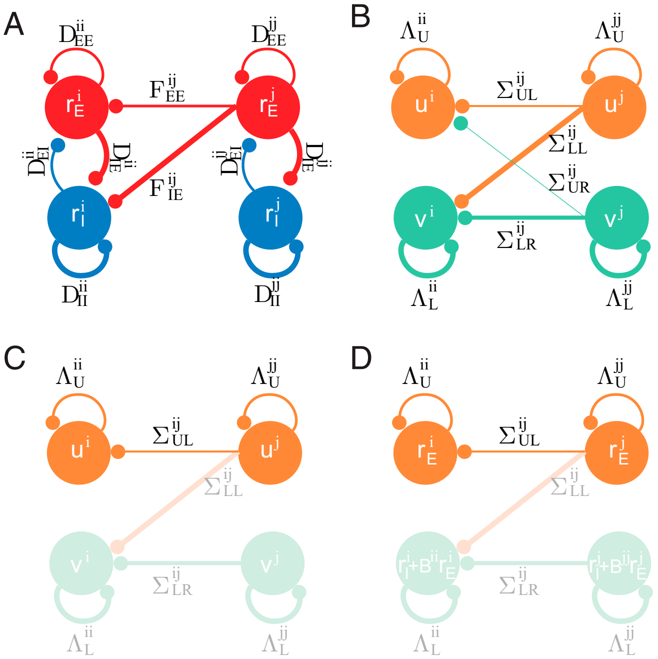

Fig. 3. Schematic illustration for the steps to prove weakly localized and orthogonal eigenvectors of the connectivity matrix *W*. (*A*) Directed interaction from area *j* to area *i* in the original model (Eqs. **1** and **2**). (*B*) One-way interaction from area *j* to area *i* after changing the coordinate system from (r_E, r_I) to (u, v). (*C*) Small δ leads to weak interaction from u^j to u^i, small ϵ additionally leads to even weaker interaction from v^j to u^i that is ignorable (proved in *SI Appendix*, *Proposition S1*), and a gradient of h_i leads to a nonzero spectral gap between area *i* and area *j*. Accordingly, they together lead to the weak localization and orthogonality of the $\boldsymbol{u}$ component in the $(\boldsymbol{u}, \boldsymbol{v})$ coordinate system (proved by *Proposition 2* and Eq. **11**). (*D*) One-way interaction from area *j* to area *i* after changing the coordinate system from $(\boldsymbol{u}, \boldsymbol{v})$ back to (r_E, r_I). To the leading order, one has $u^i \approx r^i_E$, $v \approx r^i_I + B^{ii}r^i_E$. In this step, small ϵ ensures that the leading orders of $\boldsymbol{u}$ and $\boldsymbol{r}_E$ are identical, and so are their localization and orthogonality properties (proved by *Proposition 3*). In *A–D*, the width of lines codes the interaction strength, and light-colored lines and nodes are not important in the proofs.

membrane time constant and the slope of the gain function. The substantial difference of electrophysiological properties between the excitatory and inhibitory neurons has been supported by experimental evidence, i.e., inhibitory neurons have larger slope of the gain function and smaller membrane time constant (27–30).

Second, small $\delta = \frac{\mu_{EE}}{\mu_{IE}} - \frac{w_{EI}}{w_{II}+1/\beta_I}$ indicates the balanced condition between the interareal excitatory and intraareal inhibitory inputs. When the presynaptic excitatory input from the jth area is increased by Δr_E^j, its influence on the excitatory population activity in the ith area in the steady state can be calculated in a straightforward way as $\Delta r_E^i = C^{ij} \Delta r_E^j$, in which $C^{ij} \sim \mathcal{O}(\delta)$. This indicates that the signal from the jth area has a small influence on the activity of the excitatory population in the ith area, because the global long-range excitatory input is balanced with and canceled by the local inhibitory synaptic input, leading to small net inputs in each signal pathway, as shown in Fig. 4. This condition corresponds to a detailed balance of excitation and inhibition that may benefit signal control and gating, as proposed in previous studies (31). The importance of excitation–inhibition balance on timescale hierarchy is supported by a recent study showing that the imbalance of excitation and inhibition could have a substantial effect on the change of intrinsic timescales across brain areas, which is a manifestation of psychosis such as hallucination and delusion (32).

Third, the gradient of h_i parameterizes the gradient of synaptic excitation across areas in the model, supported by the fact that h_i is proportional to the spine count per pyramidal neuron across areas (2, 4) in the form of a macroscopic gradient (33). The gradient of synaptic excitation leads to two consequences: 1) It gives rise to the hierarchy of intrinsic timescale for each area while being disconnected to other parts of the cortex and 2) it stabilizes the localization of intrinsic timescale for each area in the presence of long-range connections. From the perturbation analysis and Eq. **11**, the degree of eigenvector localization is determined by the competition between the strength of long-range connections encoded in matrix $\bar{\Sigma}_{UL}$ and the spectral gap of matrix $\bar{\Lambda}_U$. Therefore, the long-range connections tend to delocalize eigenvectors and thus break the timescale hierarchy, but the heterogeneity of local recurrent excitation level weakens its effect on eigenvector delocalization in a divisive fashion. In fact, the heterogeneity or randomness in local node properties has been shown to give localized eigenvectors in models of a physical system, for instance, a phenomenon known as Anderson localization (34) that describes the transition from a conducting medium (corresponding to delocalized eigenvectors) to an insulating medium (corresponding to localized eigenvectors). A similar mechanism has been identified in studying the eigenvector localization of an idealized neural network with simple nodes in each cortical area (35).

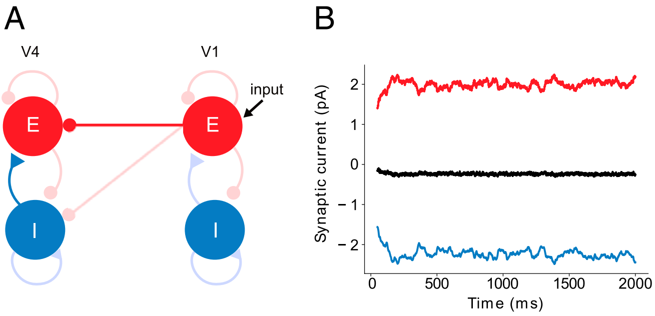

Fig. 4. The illustration of detailed balance between interareal excitation and intraareal inhibition. The projection from V1 to V4 is shown as an example. (*A*) One-way interaction from V1 to V4. V1 receives external Gaussian input. The excitatory population in V4 receives balanced excitatory interareal inputs from V1 (dark red) and intraareal inhibitory inputs from the inhibitory population in V4 (dark blue). Other excitatory and inhibitory interactions in this circuit are colored by light red and blue, respectively. (*B*) Simulation of the synaptic currents received by the V4 excitatory population induced by V1 activity. The interareal excitatory inputs (red) are balanced with the intraareal inhibitory inputs (blue), leading to small net inputs (black).

Discussion

In this work, we investigated the requirements for the emergence of a hierarchy of temporal response windows in a multiareal model of the macaque cortex (2). The original model is a nonlinear dynamical system by including a rectified linear f-I curve, and it becomes essentially linear when neural population activities are all above the firing threshold, as happened in our simulations of a hierarchical timescale phenomenon. This fact enabled us to define the time constants precisely from the eigenmodes of the connectivity matrix and carry out a detailed mathematical analysis to identify biologically interpretable conditions. (Rectified) linear models have been broadly used in theoretical and experimental neuroscience studies (36–38). Although microscopic neural activity is nonlinear in general, it has been shown in a recent study (39) that linear models can capture macroscopic cortical dynamics in the resting state more accurately than nonlinear model families, including neural field models for describing the spatiotemporal average of individual neuronal activities. Nonlinear models are more general for capturing neural circuits. However, for a nonlinear model, the time constants of the system are not uniquely defined. A linear model can be understood as a linearization of a nonlinear dynamical system around an internal state such as the resting state of the brain.

In contrast to previous computational models studying the emergence of timescales (35), the model we studied is anatomically more realistic as it incorporates 1) experimental measurements of directed and weighted anatomical connectivity, 2) a gradient of synaptic excitation reflected by spine counts in pyramidal neurons across areas, and 3) both excitatory and inhibitory neuronal populations. By performing rigorous perturbation analysis, we show that the segregation of timescales is attributable to the localization of eigenvectors of the connectivity matrix, and the parameter regime that makes this happen has three crucial properties: 1) a macroscopic gradient of synaptic excitation, 2) distinct electrophysiological properties between excitatory and inhibitory neuronal populations, and 3) a detailed balance between long-range excitatory inputs and local inhibitory inputs for each area-to-area pathway.

The theoretically identified biological conditions for the segregation of timescales enable us to make experimentally testable predictions. First, the condition of the macroscopic gradient of synaptic excitation suggests that a shallower gradient of synaptic excitation shall lead to less localized eigenvectors. Consequently, the difference of time constants for a pair of areas shall be larger if their synaptic excitation or hierarchical levels are less similar, which can be directly tested in experiments. Second, the condition of distinct electrophysiological properties between excitatory and inhibitory neuronal populations suggests that the change of neuronal physiology will affect the segregation of timescales. This condition can be tested experimentally by using genetic tools to knock down or knock out specific genes to change the firing properties of neurons (40). Third, the condition of detailed balance of excitation and inhibition suggests that areas with unbalanced excitation and inhibition could also alter hierarchical time constants. With growing evidence for excitation/inhibition imbalance in schizophrenia (41, 42), this condition is supported by recent experiments showing that the intrinsic time constants of schizophrenia patients have been substantially changed (32). And this condition can be further tested in animal models with genetic tools to disturb the excitation–inhibition balance (43).

It is worth mentioning that, although the specific pattern of interareal connectivity does not affect the eigenvector localization substantially based on the perturbation analysis, it shapes the timescale hierarchy qualitatively. In particular, the timescale hierarchy does not exactly follow monotonically the areal anatomical hierarchy in the presence of long-range connections, as shown in Fig. 1*C*. Furthermore, within a brain region time constants are heterogeneous across individual neurons (44, 45). To better relate the model with experimentally observed timescales in various specific cortical areas, the roles of long-range connections, cell types, and other circuit properties require further elucidation.

It has been noted that the neuronal activity propagates along the hierarchy with significant attenuation in the model in ref. 2. The attenuation can be alleviated by tuning the model parameters to the regime of strong global balanced amplification (GBA) (46) (parameters in *Materials and Methods*). Balanced amplification was originally introduced for a local neural network, associated with strong nonnormality of the system where eigenmodes are far from being orthogonal with each other (47). A quantity called κ measures the degree of nonnormality of a matrix (48) ($\kappa = 1$ for a normal matrix; the larger the κ value, the more nonnormal the system). We have $\kappa = 4.35$ for the original model (2), which is thus only slightly nonnormal. By contrast, $\kappa = 96.58$ for the model in the strong GBA regime. Therefore, the enhancement of signal propagation in the model correlates with the increase of the nonorthogonality of the eigenvectors or the nonnormality of the connectivity matrix. In the strong GBA regime, $\delta \approx 0.38$, which is 10 times larger than its original value, suggesting that the detailed balance condition is less well satisfied. In such a case, the localization of timescales may no longer exist in this linear model. The situation is different in nonlinear models (2, 49, 50), where inputs may be amplified by strongly recurrent circuit dynamics to enhance signal propagation or routing of information is selectively gated (for a subset of connection pathways in a goal-directed manner) (31, 51), while the conditions for a timescale hierarchy are satisfied. For a nonlinear system, however, eigenmodes can be defined only with respect to a particular network state. Consequently, the time constants observed in single neurons are no longer unique and may differ, for instance, when the brain is at rest or during a cognitive process. It remains to be seen to what extent the conditions identified here hold in the brain's various internal states, while the precise pattern of timescales can be flexibly varied to meet behavioral demands.

Materials and Methods

Model Parameters. In the macaque cortical network model, we set $\tau_E = 20$ ms, $\tau_I = 10$ ms, $\beta_E = 0.066$ Hz/pA, $\beta_I = 0.351$ Hz/pA, $w_{EE} = 24.4$ pA/Hz, $w_{IE} = 12.2$ pA/Hz, $w_{EI} = 19.7$ pA/Hz, $w_{II} = 12.5$ pA/Hz, $\mu_{EE} = 33.7$ pA/Hz, $\mu_{IE} = 25.5$ pA/Hz, and $\eta = 0.68$. We set $w_{EI} = 25.2$ pA/Hz and $\mu_{EE} = 51.5$ pA/Hz for the strong balanced amplification regime (46) introduced in *Discussion*. Some of the parameters are derived from experimental measurements of primary visual cortex (37). The FLN values are obtained from the experimental measurements of macaque cortical connectivity (3). The hierarchy values h_i of each cortical area are obtained by fitting a generalized linear model that assigns hierarchical values to areas (2) such that the differences in hierarchical values predict the supragranular layer neurons (SLNs) measured in experiment (52).

Data Availability. The python code for model simulation and perturbation analysis is publicly available on Github (https://github.com/songting858/mechanism-of-hierarchical-time-constants) (53). Previously published data were used for this work (3).

ACKNOWLEDGMENTS. We thank Y. Xiao and H. Wang for helpful discussions. This work was supported by the National Key R&D Program of China Grant 2019YFA0709503, Shanghai Municipal Science and Technology Major Project 2021SHZDZX0102, National Natural Science Foundation of China Grant 11901388, Shanghai Sailing Program 19YF1421400 (to S.L.), and James Simons Foundation Grant 543057SPI (to X.-J.W.).

1. D. H. Hubel, *Eye, Brain, and Vision* (Scientific American Library/Scientific American Books, 1995).
2. R. Chaudhuri, K. Knoblauch, M. A. Gariel, H. Kennedy, X. J. Wang, A large-scale circuit mechanism for hierarchical dynamical processing in the primate cortex. *Neuron* **88**, 419–431 (2015).
3. N. T. Markov *et al.*, A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cereb. Cortex* **24**, 17–36 (2014).
4. G. Elston, "Specialization of the neocortical pyramidal cell during primate evolution" in *Evolution of Nervous Systems: A Comprehensive Reference*, J. H. Kaass, T. M. Preuss, Eds. (Elsevier, Amsterdam, The Netherlands, 2007), **vol. 4**, pp. 191–242.
5. J. I. Gold, M. N. Shadlen, The neural basis of decision making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007).
6. X. J. Wang, Decision making in recurrent neuronal circuits. *Neuron* **60**, 215–234 (2008).
7. C. A. Runyan, E. Piasini, S. Panzeri, C. D. Harvey, Distinct timescales of population coding across cortex. *Nature* **548**, 92–96 (2017).
8. J. H. Siegle *et al.*, Survey of spiking in the mouse visual system reveals functional hierarchy. *Nature* **592**, 86–92 (2021).
9. T. Ogawa, H. Komatsu, Differential temporal storage capacity in the baseline activity of neurons in macaque frontal eye field and area V4. *J. Neurophysiol.* **103**, 2433–2445 (2010).
10. J. D. Murray *et al.*, A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci.* **17**, 1661–1663 (2014).
11. S. E. Cavanagh, J. D. Wallis, S. W. Kennerley, L. T. Hunt, Autocorrelation structure at rest predicts value correlates of single neurons during reward-guided choice. *eLife* **5**, e18937 (2016).
12. V. Fascianelli, S. Tsujimoto, E. Marcos, A. Genovesio, Autocorrelation structure in the macaque dorsolateral, but not orbital or polar, prefrontal cortex predicts response-coding strength in a visually cued strategy task. *Cereb. Cortex* **29**, 230–241 (2019).
13. M. Spitmaan, H. Seo, D. Lee, A. Soltani, Multiple timescales of neural dynamics and integration of task-relevant signals across cortex. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 22522–22531 (2020).
14. D. J. N. Maisson *et al.*, Choice-relevant information transformation along a ventrodorsal axis in the medial prefrontal cortex. *Nat. Commun.* **12**, 4830 (2021).
15. A. M. G. Manea, A. Zilverstand, K. Uğurbil, S. R. Heilbronner, J. Zimmermann, Intrinsic timescales as an organizational principle of neural processing across the whole rhesus macaque brain. bioRxiv [Preprint] (2021). https://doi.org/10.1101/2021.10.05.463277. Accessed 7 October 2021.
16. U. Hasson, E. Yang, I. Vallines, D. J. Heeger, N. Rubin, A hierarchy of temporal receptive windows in human cortex. *J. Neurosci.* **28**, 2539–2550 (2008).
17. C. J. Honey *et al.*, Slow cortical dynamics and the accumulation of information over long timescales. *Neuron* **76**, 423–434 (2012).
18. Y. Lerner, C. J. Honey, L. J. Silbert, U. Hasson, Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J. Neurosci.* **31**, 2906–2915 (2011).
19. G. J. Stephens, C. J. Honey, U. Hasson, A place for time: The spatiotemporal structure of neural dynamics during natural audition. *J. Neurophysiol.* **110**, 2019–2026 (2013).
20. Y. Yeshurun, M. Nguyen, U. Hasson, Amplification of local changes along the timescale processing hierarchy. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 9475–9480 (2017).
21. R. V. Raut, A. Z. Snyder, M. E. Raichle, Hierarchical dynamics as a macroscopic organizing principle of the human brain. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 20890–20897 (2020).
22. G. Shafiei *et al.*, Topographic gradients of intrinsic dynamics across neocortex. *eLife* **9**, e62116 (2020).
23. R. Gao, R. L. van den Brink, T. Pfeffer, B. Voytek, Neuronal timescales are functionally dynamic and shaped by cortical microarchitecture. *eLife* **9**, e61277 (2020).
24. R. Gămănuţ *et al.*, The mouse cortical connectome, characterized by an ultra-dense cortical graph, maintains specificity by distinct connectivity profiles. *Neuron* **97**, 698–715.e10 (2018).
25. N. T. Markov *et al.*, Weight consistency specifies regularities of macaque cortical networks. *Cereb. Cortex* **21**, 1254–1272 (2011).
26. T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, Berlin, Germany, 1966).
27. B. Ahmed, J. C. Anderson, R. J. Douglas, K. A. Martin, D. Whitteridge, Estimates of the net excitatory currents evoked by visual stimulation of identified neurons in cat visual cortex. *Cereb. Cortex* **8**, 462–476 (1998).
28. L. G. Nowak, R. Azouz, M. V. Sanchez-Vives, C. M. Gray, D. A. McCormick, Electrophysiological classes of cat primary visual cortical neurons in vivo as revealed by quantitative analyses. *J. Neurophysiol.* **89**, 1541–1566 (2003).
29. N. V. Povysheva *et al.*, Parvalbumin-positive basket interneurons in monkey and rat prefrontal cortex. *J. Neurophysiol.* **100**, 2348–2360 (2008).
30. A. V. Zaitsev, N. V. Povysheva, G. Gonzalez-Burgos, D. A. Lewis, Electrophysiological classes of layer 2/3 pyramidal cells in monkey prefrontal cortex. *J. Neurophysiol.* **108**, 595–609 (2012).
31. T. P. Vogels, L. F. Abbott, Gating multiple signals through detailed balance of excitation and inhibition in spiking networks. *Nat. Neurosci.* **12**, 483–491 (2009).
32. K. Wengler, A. T. Goldberg, G. Chahine, G. Horga, Distinct hierarchical alterations of intrinsic neural timescales account for different manifestations of psychosis. *eLife* **9**, e56151 (2020).
33. X. J. Wang, Macroscopic gradients of synaptic excitation and inhibition in the neocortex. *Nat. Rev. Neurosci.* **21**, 169–178 (2020).
34. P. Anderson, Absence of diffusion in certain random lattices. *Phys. Rev.* **109**, 1492–1505 (1958).
35. R. Chaudhuri, A. Bernacchia, X. J. Wang, A diversity of localized timescales in network activity. *eLife* **3**, e01239 (2014).

36. A. Roxin, N. Brunel, D. Hansel, Role of delays in shaping spatiotemporal dynamics of neuronal activity in large networks. *Phys. Rev. Lett.* **94**, 238103 (2005).
37. T. Binzegger, R. J. Douglas, K. A. Martin, Topology and dynamics of the canonical circuit of cat V1. *Neural Netw.* **22**, 1071–1078 (2009).
38. D. Jercog *et al.*, UP-DOWN cortical dynamics reflect state transitions in a bistable network. *eLife* **6**, e22425 (2017).
39. E. Nozari *et al.*, Is the brain macroscopically linear? A system identification of resting state dynamics. bioRxiv [Preprint] (2020). https://doi.org/10.1101/2020.12.21.423856. Accessed 11 August 2021.
40. J. Gingras *et al.*, Global Nav1.7 knockout mice recapitulate the phenotype of human congenital indifference to pain. *PLoS One* **9**, e105895 (2014).
41. J. H. Foss-Feig *et al.*, Searching for cross-diagnostic convergence: Neural mechanisms governing excitation and inhibition balance in schizophrenia and autism spectrum disorders. *Biol. Psychiatry* **81**, 848–861 (2017).
42. R. Jardri *et al.*, Are hallucinations due to an imbalance between excitatory and inhibitory influences on the brain? *Schizophr. Bull.* **42**, 1124–1134 (2016).
43. C. L. Gatto, K. Broadie, Genetic controls balancing excitatory and inhibitory synaptogenesis in neurodevelopmental disorder models. *Front. Synaptic Neurosci.* **2**, 4 (2010).
44. A. Bernacchia, H. Seo, D. Lee, X. J. Wang, A reservoir of time constants for memory traces in cortical neurons. *Nat. Neurosci.* **14**, 366–372 (2011).
45. S. E. Cavanagh, L. T. Hunt, S. W. Kennerley, A diversity of intrinsic timescales underlie neural computations. *Front. Neural Circuits* **14**, 615626 (2020).
46. M. R. Joglekar, J. F. Mejías, G. R. Yang, X. J. Wang, Inter-areal balanced amplification enhances signal propagation in a large-scale circuit model of the primate cortex. *Neuron* **98**, 222–234.e8 (2018).
47. B. K. Murphy, K. D. Miller, Balanced amplification: A new mechanism of selective amplification of neural activity patterns. *Neuron* **61**, 635–648 (2009).
48. L. N. Trefethen, M. Embree, *Spectra and Pseudospectra* (Princeton University Press, 2020).
49. J. F. Mejias, X. J. Wang, Mechanisms of distributed working memory in a large-scale model of the macaque neocortex. bioRxiv [Preprint] (2020). https://doi.org/10.1101/760231. Accessed 2 April 2021.
50. H. S. Chien, C. J. Honey, Constructing and forgetting temporal context in the human cerebral cortex. *Neuron* **106**, 675–686.e11 (2020).
51. X. J. Wang, G. R. Yang, A disinhibitory circuit motif and flexible information routing in the brain. *Curr. Opin. Neurobiol.* **49**, 75–83 (2018).
52. N. T. Markov *et al.*, Anatomy of hierarchy: Feedforward and feedback pathways in macaque visual cortex. *J. Comp. Neurol.* **522**, 225–259 (2014).
53. S. Li, Simulation and analysis code. GitHub. https://github.com/songting858/mechanism-of-hierarchical-time-constants. Deposited 9 September 2021.