Article

Neuron

Inter-areal Balanced Amplification Enhances Signal Propagation in a Large-Scale Circuit Model of the Primate Cortex

Highlights

- Long-range excitatory loops destabilize the system, complicating propagation
- Balanced amplification facilitates robust signal transmission in large-scale models
- Stable transmission in synchronous and asynchronous modes in spiking network models
- Prefrontal activity beyond threshold supports ignition model of conscious processing

Authors

Madhura R. Joglekar, Jorge F. Mejias, Guangyu Robert Yang, Xiao-Jing Wang

Correspondence

xjwang@nyu.edu

In Brief

Joglekar et al. propose a basic circuit motif that allows for stable signal transmission in large-scale cortical-circuit models. The motif contains strong long-range recurrent excitation stabilized by local feedback inhibition, extending the balanced amplification mechanism.

Joglekar et al., 2018, Neuron *98*, 1–13
April 4, 2018
https://doi.org/10.1016/j.neuron.2018.02.031

CellPress

Please cite this article in press as: Joglekar et al., Inter-areal Balanced Amplification Enhances Signal Propagation in a Large-Scale Circuit Model of the Primate Cortex, Neuron (2018), https://doi.org/10.1016/j.neuron.2018.02.031

Neuron
Article

Inter-areal Balanced Amplification Enhances Signal Propagation in a Large-Scale Circuit Model of the Primate Cortex

Madhura R. Joglekar,[1,3] Jorge F. Mejias,[1,4] Guangyu Robert Yang,[1,5] and Xiao-Jing Wang[1,2,6,*]
[1]Center for Neural Science, New York University, 4 Washington Place, New York, NY 10003
[2]Key Laboratory of Brain Functional Genomics and Institute of Cognitive Neuroscience, East China Normal University, Shanghai 200062, China
[3]Present address: Courant Institute of Mathematical Sciences, 251 Mercer St., New York, NY 10012
[4]Present address: Swammerdam Institute for Life Sciences, Center for Neuroscience, Faculty of Science, University of Amsterdam, 1098XH Amsterdam, the Netherlands
[5]Present address: Mortimer B. Zuckerman Mind Brain Behavior Institute, Department of Neuroscience, Columbia University, New York, NY 10027, USA
[6]Lead Contact
*Correspondence: xjwang@nyu.edu
https://doi.org/10.1016/j.neuron.2018.02.031

SUMMARY

Understanding reliable signal transmission represents a notable challenge for cortical systems, which display a wide range of weights of feedforward and feedback connections among heterogeneous areas. We re-examine the question of signal transmission across the cortex in a network model based on mesoscopic directed and weighted inter-areal connectivity data of the macaque cortex. Our findings reveal that, in contrast to purely feedforward propagation models, the presence of long-range excitatory feedback projections could compromise stable signal propagation. Using population rate models as well as a spiking network model, we find that effective signal propagation can be accomplished by balanced amplification across cortical areas while ensuring dynamical stability. Moreover, the activation of prefrontal cortex in our model requires the input strength to exceed a threshold, which is consistent with the ignition model of conscious processing. These findings demonstrate our model as an anatomically realistic platform for investigations of global primate cortex dynamics.

INTRODUCTION

In computational neuroscience, there is a lack of knowledge about multi-regional brain circuits. New questions that are not crucial for understanding local circuits arise when we investigate how a large-scale brain system works. In particular, reliable signal propagation is a prerequisite for information processing in a hierarchically organized cortical system. A number of studies have been devoted to signal propagation from area to area in the mammalian cortex (Perkel and Bullock, 1968; Kumar et al., 2010; van Rossum et al., 2002; Shadlen and Newsome, 1998; Diesmann et al., 1999; Kumar et al., 2008; Marsálek et al., 1997). It was found that a major challenge is to ensure stable transmission, with the signal undergoing neither successive attenuation nor amplification as it travels across multiple areas in a hierarchy. In spite of the insights these studies provided, virtually all previous works did not incorporate data-constrained cortical connectivity and made unrealistic assumptions—for instance, areas are considered identical, network architecture is strictly feedforward, and connection weights are the same at all stages of the hierarchy. Here we argue that achieving stable signal propagation becomes even more challenging with the inclusion of more realistic network architecture and connectivity. Mechanisms that improve signal propagation in simpler models may no longer work for more biological models.

Inter-areal cortical networks are highly recurrent and are abundant with feedback loops (Markov et al., 2014a). These rich feedback connections pose the risk of destabilizing the system through reverberation as the signal is transmitted across areas. Additionally, local cortical circuits are strongly recurrently connected (Markov et al., 2011), further contributing to system instability. Therefore, mechanisms that improve signal propagation in feedforward networks may quickly lead to instability in a brain-like network. Building such a model requires quantitative anatomical data. Recently, detailed mesoscopic connectivity data has become available for both macaque monkey (Markov et al., 2011, 2014a; Wang and Kennedy, 2016) and mouse (Oh et al., 2014; Zingg et al., 2014). This data delineates a complex inter-areal cortical network with connection weights spanning several orders of magnitude. Some areas receive strong inter-areal connections, while some other areas appear more disconnected from the rest of the cortex. It is particularly challenging to facilitate signal propagation to the weakly connected areas while maintaining stability for the more strongly connected core areas (Markov et al., 2013). Recent work employs directed and weighted inter-areal connectivity data to build biologically realistic, large-scale dynamical models of the primate cortex (Chaudhuri et al., 2015; Mejias et al., 2016) with both local and

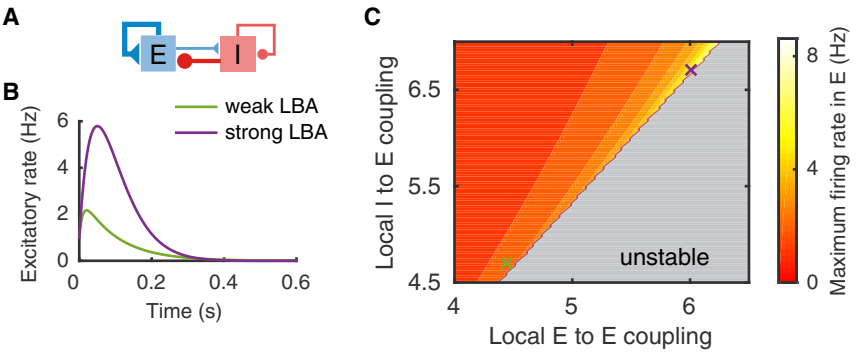

Figure 1. Balanced Amplification in an Inhibition-Stabilized Local Circuit

(A) Model scheme with strong recurrent excitation balanced by strong lateral inhibition.

(B) Local Balanced Amplification (LBA) results in a transient amplification of the excitatory firing rate prior to decay in response to a brief input that sets the initial rate to 1.

(C) Peak response of the excitatory firing rate as a function of recurrent excitation (local E-to-E connection) and lateral inhibition (local I-to-E connection). The gray region on the lower right indicates instability. The green and purple crosses correspond to the parameter values used in (B).

long-range recurrent connections. This connectivity data spans 29 widely distributed cortical areas across the occipital, temporal, parietal, and frontal lobes (Markov et al., 2014a). These anatomically calibrated models thus provide a useful framework for re-examining signal propagation in the cortex.

We propose a novel, biologically plausible mechanism based on transient signal amplification to improve reliable cortical signal transmission. Our mechanism is inspired by the balanced amplification mechanism (Murphy and Miller, 2009), extending its central idea from the local circuit to a large-scale system. Balanced amplification was originally studied in local inhibition-stabilized network models (Ozeki et al., 2009; Murphy and Miller, 2009). These networks are characterized by a strong recurrent excitation, which drives the neural activity toward instability, followed by a strong lateral inhibition that stabilizes neural activity (Figure 1A). These two factors combined result in a transient amplification of the excitatory firing rate in response to a brief input prior to stabilization, a phenomenon referred to as balanced amplification (Murphy and Miller, 2009) (Figure 1B), or Local Balanced Amplification (LBA) for a local network. Increasing LBA can evoke a stronger transient excitatory response prior to decay (Figures 1B and 1C), which leads to a transient amplification of activity in the local circuit (Murphy and Miller, 2009). Our proposed signal transmission mechanism is a form of balanced amplification in which inter-areal excitatory connection strengths are increased together with increased intra-areal lateral inhibition.

We test our mechanism in a range of large-scale models of the primate cortex, including in recent population-rate models with heterogeneity across areas (Chaudhuri et al., 2015) and with a cortical laminar structure (Mejias et al., 2016). The inter-areal connectivity in these models is set according to a connectivity dataset of the macaque cortex, in which the directed and weighted connectivity matrix was obtained using tract-tracing techniques (Markov et al., 2014a; see STAR Methods). Briefly, a retrograde tracer was injected into a given (target) area, labeling presynaptic neurons in source areas that connect to the target area. The relative weight of a directed connection was measured as the number of labeled neurons in a source area divided by the total number of labeled neurons in all source areas, called the Fraction of Labeled Neurons (FLN) (Markov et al., 2014a). These rate models (Chaudhuri et al., 2015; Mejias et al., 2016) also incorporate heterogeneity across cortical areas (Chaudhuri et al., 2015), assuming the number of spines per pyramidal cell as a proxy of the strength of excitatory inputs that varies from area to area (Elston, 2007). Here, we demonstrate our mechanism in these population-rate models (Chaudhuri et al., 2015; Mejias et al., 2016) by changing the excitatory and inhibitory connection strengths based on our extension of balanced amplification, thereby facilitating improved propagation.

In order to examine synchronous (Diesmann et al., 1999; Kumar et al., 2008) and asynchronous (Shadlen and Newsome, 1998; van Rossum et al., 2002; Vogels and Abbott, 2009; Brunel, 2000a) transmission, which cannot be properly captured with firing rate models, we build a large-scale cortical spiking network model. We find that our mechanism always improves signal transmission in all areas for each model—by as much as 100-fold in some areas. Furthermore, our large-scale spiking model displays several key features of subliminal, preconscious, and conscious processing (Dehaene et al., 2006; King et al., 2016). Taken together, the findings demonstrate that our network models offer a valuable platform to study a wide range of dynamical questions that involve long-range interactions between cortical areas.

RESULTS

Transmission in a Realistic Large-Scale Cortical Network

Here we demonstrate the challenge of reliable signal transmission through a large-scale network consisting of population rate models (Chaudhuri et al., 2015). On investigating signal transmission in this dynamical model of the macaque cortex (Chaudhuri et al., 2015), we find that inter-areal excitatory loops between cortical areas make reliable signal transmission especially difficult. In Figure 2A, an input is applied to V1, which is the lowest in the cortical hierarchy (Markov et al., 2014b; Chaudhuri et al., 2015; see STAR Methods), and the maximum firing rate is shown for V1 and area 24c at the top of the hierarchy. Inter-areal connections in the model are governed by two global coupling parameters, μ_{EE} and μ_{IE}, corresponding to the long-range excitatory-to-excitatory and excitatory-to-inhibitory coupling, respectively. A small increase in μ_{EE} can result in the system behavior switching from strong attenuation to instability, as shown in Figure 2B. A more systematic characterization of the model's behavior reveals that a gradual increase in μ_{EE} leads to a sharp transition from a regime characterized by strong signal

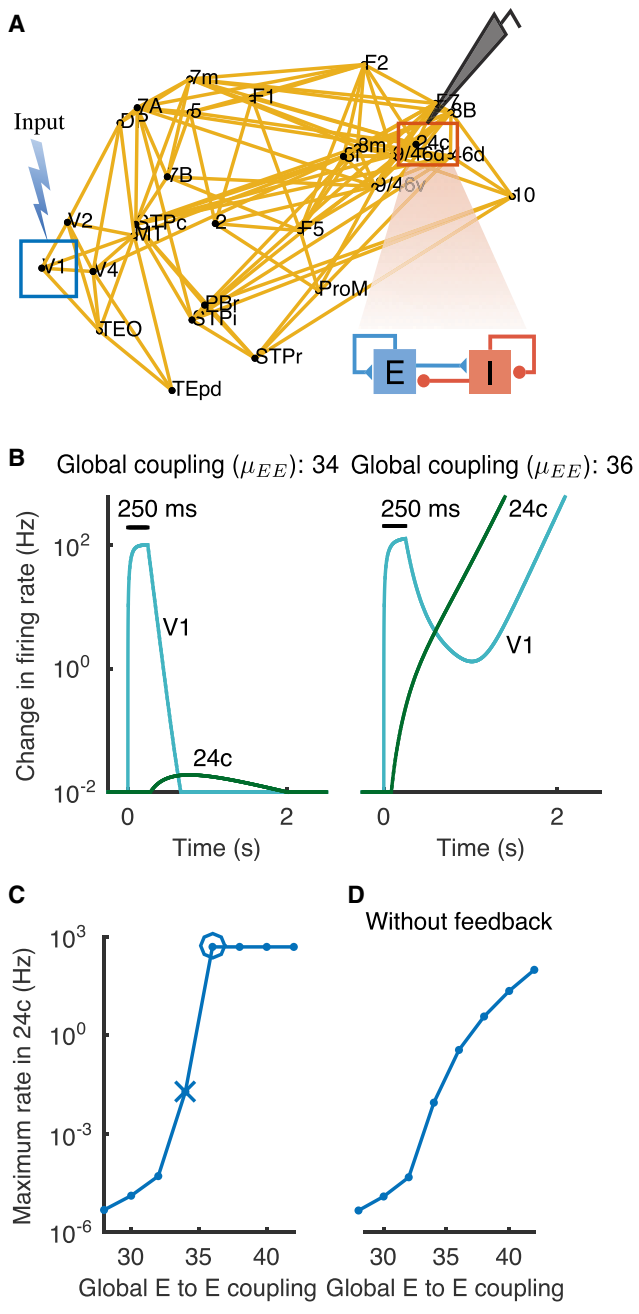

Figure 2. Signal Propagation in a Large-Scale Network with Recurrent Excitatory Loops Leads to Either Strong Attenuation or Instability

(A) Network model of the 29 cortical areas with the strongest inter-areal projection strengths (FLN values >0.005). Input is injected to the excitatory population in area V1 at the lowest position in the cortical hierarchy, and the peak firing rate in area 24c at the top of the hierarchy is recorded.

(B) The response of the excitatory populations in areas V1 and 24c to a current pulse of 250 ms to V1, using two close values of the excitatory global coupling parameter μ_{EE}. Response (left) in 24c shows an attenuation of four orders of magnitude when $\mu_{EE} = 34$. With $\mu_{EE} = 36$ (right), excitatory firing rates in V1 and 24c exponentially grow, and the system becomes unstable.

(C) The peak excitatory population response in 24c as a function of μ_{EE}, dynamical instability corresponds to capped firing rate at 500 Hz. The particular parameter values corresponding to weak propagation and instability in (B) are indicated by a cross and a circle respectively.

(D) Similar to (C), but μ_{EE} is varied in the absence of feedback projections. See also Figure S1.

attenuation to a regime indicating instability (Figure 2C). Neither regime allows for a realistic propagation of the signal across cortical areas.

To answer the question of whether this sharp transition is due to the inter-areal excitatory loops, we examine the model's behavior when feedback projections are removed from the network. This reveals a smooth transition from the regime with weak propagation to a regime with improved propagation (Figure 2D), suggesting that removing feedback projections, and therefore the presence of inter-areal excitatory loops, alleviates the problem of effective transmission. Around half of the inter-areal projections present in the anatomical connectivity data, however, correspond to feedback projections, with strengths comparable to those of feedforward projections (Markov et al., 2014a; Figure S1); this suggests that feedback projections cannot be ignored. The question of signal propagation, therefore, becomes especially pertinent in a biologically realistic cortical model endowed with feedback connections.

Extending Balanced Amplification beyond the Local Circuit

Increasing balanced amplification has been demonstrated to allow for transient amplification in local circuits that are characterized by strong recurrent excitation stabilized by inhibition (Murphy and Miller, 2009; Figures 1A–1C). It can be analytically shown, using a phase diagram of the network activity as a function of connectivity parameters, that moving along the stability boundary in the direction of increasing LBA (Figure 1C) leads to a progressive increase in the steady-state excitatory firing rate (see STAR Methods). This can be used to intuitively understand the transient amplification achieved with stronger LBA. A key idea of the present work is an extension of this mechanism from local circuits to large-scale models (Figure 3A) to boost inter-areal signal transmission while maintaining system stability. To this end, we replace the increase in local recurrent excitation with an increase in the global excitatory coupling μ_{EE}, and we stabilize the system with stronger lateral inhibition, as in the case of LBA. This principle of strong long-range excitation stabilized by strong local inhibition constitutes an extension of the balanced amplification mechanism for large-scale systems, which we name Global Balanced Amplification (GBA). From now on, we refer to increasing global excitatory coupling and stabilizing the system with stronger lateral inhibition—while keeping other model parameters the same—as increasing GBA.

To quantify how increasing GBA affects propagation in the large-scale network model (Chaudhuri et al., 2015), we measure the quality of signal transmission by comparing the peak value of the excitatory firing rate in area 24c with the same peak value in area V1; the ratio between both peaks is defined as the "propagation ratio." The response of the different cortical areas to a pulse input in V1, as the signal propagates along the hierarchy, shows a strong attenuation of ~10,000 fold from area V1 to

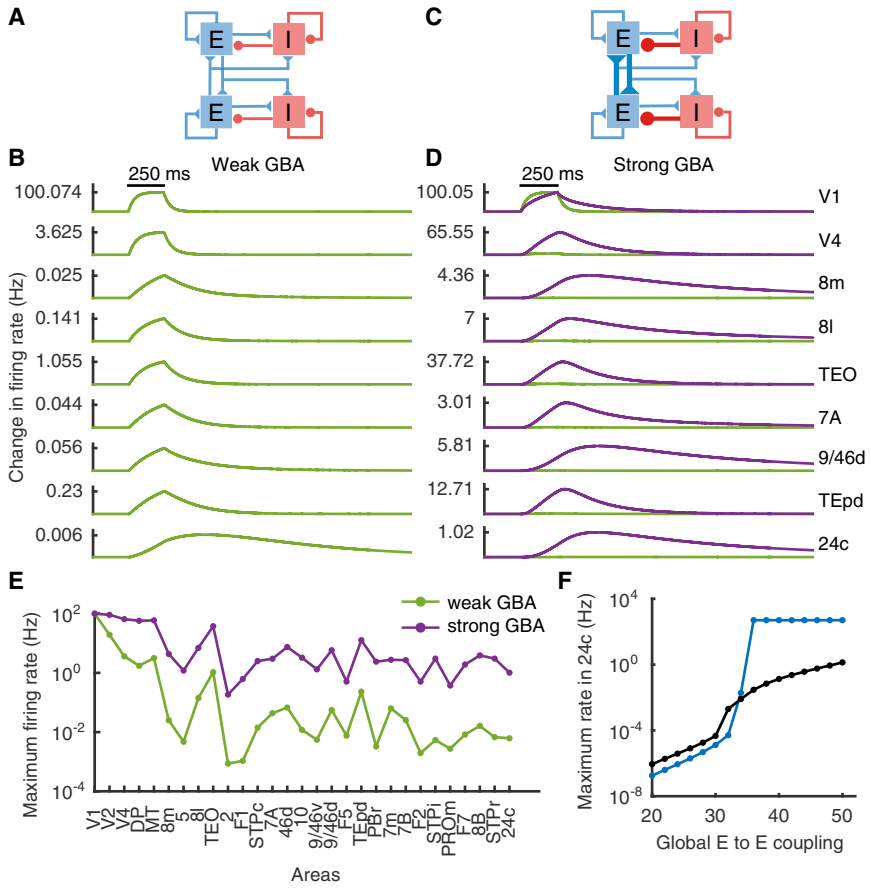

Figure 3. Global Balanced Amplification in the Large-Scale Model Improves Signal Propagation

(A) Model scheme with weak GBA.

(B) With weak GBA, there is 10,000-fold attenuation of signals from V1 to area 24c (top to bottom).

(C) Model scheme with strong GBA.

(D) With strong GBA, signal propagation is enhanced by 100-fold (purple), overlaid on the green curve from (B). With weak GBA, there is 10,000-fold attenuation of signals from V1 to area 24c (top to bottom).

(E) The maximum firing rate across areas as the response to the pulse input to V1 propagates along the hierarchy for weak (green) and strong (purple) GBA.

(F) Peak firing rate response in area 24c with strong GBA (black) overlaid on the curve corresponding to a sole increase in global excitatory coupling (blue curve, Figure 2C), demonstrating that network instability is prevented by GBA. See also Figures S2 and S3.

area 24c (Chaudhuri et al., 2015; Figure 3B). Interestingly, by increasing GBA (Figure 3C), the propagation ratio is improved by around two orders of magnitude (Figure 3D). More precisely, a substantial improvement is observed across most of the cortical areas (Figures 3D and 3E). More systematic simulations show that, as opposed to simply increasing the global excitatory coupling (Figure 2C), increasing GBA leads to a smooth transition from the weak to the improved propagation regime (Figure 3F). A consistent improvement in propagation ratio is revealed through a parametric analysis balancing strong global excitatory coupling with increased lateral inhibition (Figure S2A).

From a mathematical point of view, balanced amplification in inhibition-stabilized networks results from the non-normality of the underlying connectivity matrix (Murphy and Miller, 2009) (a non-normal matrix is one in which its eigenvectors are not mutually orthogonal). Non-normality can be examined through Schur decomposition, by expressing the effective connectivity across network basis patterns through self-connections and feedforward connections. An improvement in signal propagation in the model correlates with an increase in the non-normality measure (Henrici, 1962) of the anatomical connectivity matrix underlying the large-scale dynamics (Figure S2B).

Our mechanism is robust either when feedback projections are removed from the large-scale network or in the absence of heterogeneity across areas (Figures S3A and S3B). Moreover, we test our mechanism after symmetrizing the anatomical connectivity matrix and either partially (Figure S3C) or completely (Figure S3D) removing the hierarchical organization. Removal of the hierarchical organization weakens propagation, but incorporating the mechanism continues to show an improvement in propagation.

Inspired by the "small-world" property characterized by high clustering coefficients and short path lengths, we assess the effect of removing weak anatomical connections with a parametric threshold of connection strength (Ercsey-Ravasz et al., 2013; Figure S3E). The mechanism consistently reveals a significant improvement in propagation with stronger GBA.

We also examine (Figure S3F) the robustness of our mechanism upon incorporating into our model recent data on structural asymmetries in the recruitment of inhibition and excitation (D'Souza et al., 2016), wherein a gradual scaling down of the inhibition:excitation ratio is observed from the most feedforward to the most feedback pathway. Monitoring the propagation ratio on varying the scaling parameter governing the strength of long-range excitatory to inhibitory connections based on hierarchy reveals a consistent improvement in propagation (Figure S3F).

Incorporating a Cortical Laminar Structure

Does our signal propagation mechanism still hold in a cortical system where feedforward and feedback projections are wired in a layer-dependent manner (Felleman and Van Essen, 1991; Barbas and Rempel-Clower, 1997; Markov et al., 2014b)? We further test efficient signal propagation using a recent model which incorporates laminar structure in the cortical areas (Mejias et al., 2016; Figure 4A). Increasing GBA improves signal transmission in this model as well (Figure 4B), suggesting that our mechanism is also valid for models equipped with more detailed laminar-specific projection patterns (Mejias et al., 2016). Improved propagation is accompanied by a small change

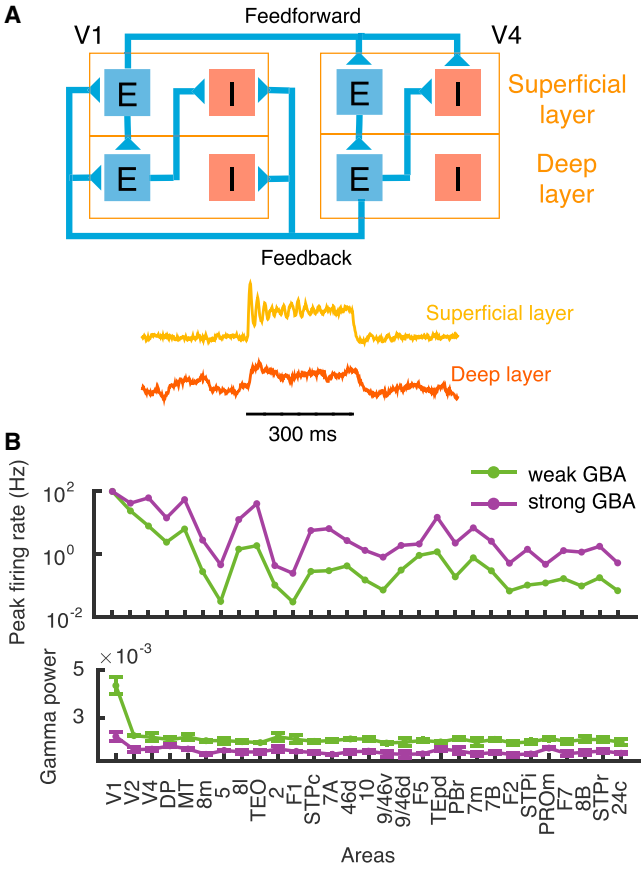

Figure 4. Increasing Global Balanced Amplification Also Improves Signal Propagation in a Large-Scale Model Endowed with a Laminar Structure

(A) Circuit diagram (upper) showing the intra- and inter-areal connectivity in the large-scale model with a superficial layer and a deep layer in each local area. In addition to the connections shown, the E-I circuit in each layer of every area has local connectivity as in Figure 2A. Sample oscillatory activity (lower) for the local circuit in the superficial layer (top) and deep layer (bottom) in response to a brief input of 300 ms given to the excitatory population in the superficial layer of V1.

(B) Peak firing rate (upper) across areas as the response to a pulse input to V1 propagates along successive areas in the hierarchy, for weak (green) and strong (purple) GBA. Gamma power (lower) across cortical areas (except for early areas) is slightly reduced from the weak (green) to strong (purple) GBA regime.

in gamma power across areas (Figure 4B), stemming from the stronger lateral inhibition associated with improved GBA. The moderate gamma power is in good agreement with experimental observations, which show the presence of a weakly coherent gamma rhythm during cortical interactions and its importance in local and long-distance interactions (Kopell et al., 2000; Bastos et al., 2015; Wang, 2010).

Two (Asynchronous and Synchronous) Regimes of Transmission in a Large-Scale Spiking Network Model

Signal transmission in neural systems is mediated through spiking activity; therefore, it is crucial to assess our mechanism in a more realistic spiking neural network model. We thereby extend our investigation by building a large-scale spiking network model (Figure 5A; see STAR Methods). Inter-areal connectivity is based on the same anatomical data used above (Markov et al., 2014a), and the inter-areal delays are introduced by considering the corresponding inter-areal wiring distances (Markov et al., 2014a) and assuming a constant axonal conduction velocity. Propagation of spiking activity has been previously studied in the context of asynchronous firing rates (Shadlen and Newsome, 1998; van Rossum et al., 2002; Vogels and Abbott, 2009) and synchronous activity (Diesmann et al., 1999; Kumar et al., 2008), both of which have been observed experimentally (Romo et al., 1999; Reyes, 2003; DeMarse et al., 2016; Vaadia et al., 1995).

We first evaluate the performance of our mechanism in the asynchronous regime (Figure 5), and later in the synchronous regime (Figure 6). To examine asynchronous propagation, we stimulate V1 with a long (150-ms) pulse input. The corresponding raster plot for weak GBA (Figure 5B) shows a strong response activity only in early visual areas. Weak activity is observed in a part of the frontal eye fields (area 8l), but is conspicuously absent from the dorsolateral prefrontal cortex (dlPFC), which has been associated with working memory and decision making (Histed et al., 2009). Increasing GBA facilitates signal propagation to higher cortical areas (Figure 5C). Stronger response activity is observed in the higher areas, including those in the dLFPC (area 46d, 9/46d), the frontopolar cortex (area 10), the parietal area 7 (7A and 7B, 7 m) in the dorsal stream, and the frontal eye fields (area 8l, 8 m) as indicated by the peak firing rate responses across areas (Figures 5D and 5E). For weak GBA, signals propagate mainly along the ventral visual stream (V1, V2, V4, TEO, and TEpd) (Figure 5B). This is presumably due to the much stronger anatomical projection weights between these areas compared to the overall connectivity (with a significant difference between their average projection strengths, $p = 0.015$) (Figures 5F and 5G). These stronger weights enable propagation along ventral areas through recurrent excitatory loops, while the signal fails to reach higher areas due to relatively weak connections.

Even with strong GBA, many areas do not show noticeable responses to an input to V1. These previously silent areas can, however, be activated when the input is directed to a different sensory area. For example, an input to the primary somatosensory cortex (area 2) uncovers a new set of areas showing propagation (Figure S4) with the same connectivity parameter values used in Figures 5B and 5C. Propagation was observed largely in the somatosensory areas of the parietal lobe, and it extended to prefrontal areas for strong GBA.

Our model displays a second mode of transmission in the synchronous regime. Following previous works on synchronous propagation (Diesmann et al., 1999; Moldakarimov et al., 2015), we stimulated V1 with a brief (10-ms) input pulse rather than a long-lasting stimulus. Connectivity is set stronger as compared to the asynchronous propagation case (see STAR Methods) to allow for a quick buildup of network activity, since stronger connectivity leads to a higher degree of population synchrony (Figure S5). For weak GBA, signal propagates in the visual areas, but it does not reach higher cognitive areas in prefrontal cortex (Figure 6A), as in the case of the asynchronous model

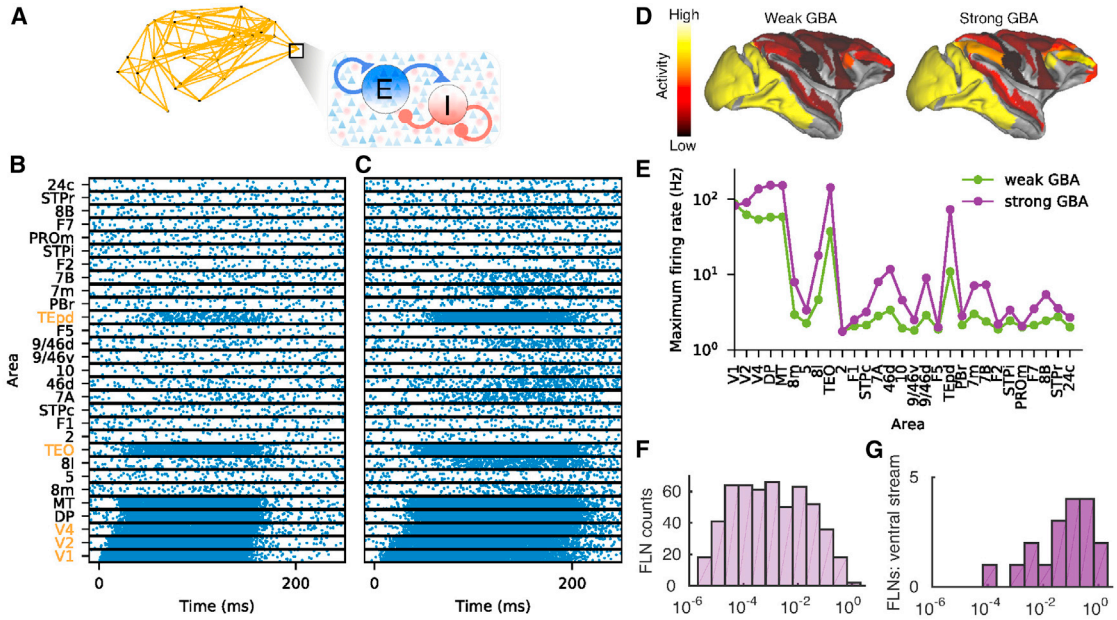

Figure 5. Reliable Signal Propagation in the Asynchronous Regime in a Spiking Network Model

(A) Circuit diagram for local-circuit connectivity for the spiking network model. Each cortical area contains 1,600 excitatory and 400 inhibitory neurons.

(B) Response to a 150 ms pulse input to V1 as it propagates along the hierarchy. The areas along the ventral stream showing strong response activity are indicated with orange labels.

(C) Similar to (B), but with strong GBA.

(D) Spatial activity pattern across the macaque cortical surface corresponding to parameters in (B) and (C).

(E) Peak firing rate across areas as the response to a pulse input to V1 propagates along the hierarchy for weak (green) and strong (purple) GBA.

(F) FLN strengths of long-range connectivity across areas span five orders of magnitude.

(G) FLN strengths along the ventral visual stream areas are stronger than the average, leading to more effective signal propagation along the ventral pathway than the dorsal pathway. See also Figures S4 and S5.

(Figure 5B). Increasing GBA enables successful signal propagation to several higher areas, including those in the dlPFC (areas 46d, 9/46d) and the frontal eye fields (areas 8l, 8 m) (Figures 6B–6D).

Early response onset occurs along the ventral stream (Figure 6B); the signal then propagates to higher cognitive areas and eventually to the superior temporal polysensory (STP) areas involved in multisensory integration, which form part of a cluster that shapes functional connectivity (Chaudhuri et al., 2015). After testing multiple hypotheses (Figure S6), we found that the observed response onset times in our large-scale network are best predicted by a shortest-path toy model (Figure 6E). In this toy model, we first ignore projections whose strength falls below a certain threshold value, and we then assume that the signal, starting from V1, follows the shortest possible path to reach any given area. The shortest path is determined based on the anatomical inter-areal wiring distances (Markov et al., 2014a), and a constant conduction velocity is assumed (see STAR Methods).

Threshold Crossing for Signal Propagation and Conscious Perception

The emergence of activity across several cortical areas in Figures 5C and 6B is reminiscent of the ‘‘global ignition’’ observed during conscious perception (Dehaene et al., 2006; Del Cul et al., 2007; Dehaene and Changeux, 2011; King et al., 2016). Global ignition is characterized by a distributed cerebral activation pattern, contingent on the strong parieto-frontal network activation that emerges when the bottom-up input exceeds a certain threshold (Dehaene et al., 2006). To examine these phenomena, we monitor the activity across different cortical lobes in our asynchronous propagation model (Figure 5C) on successively increasing the input current strength arriving at V1. From our original simulations on the large-scale spiking model with improved propagation (Figures 5C and 5E), we estimate that the signal from V1 reaches around 16 cortical areas; we will refer to these here as ‘‘active areas’’ (see STAR Methods). On varying the input strength, the normalized peak response across these areas (Figure 7A) reveals a clear separation across the four cortical lobes. Low input (Figure 7A) shows increased activity in the occipital lobe compared to the other lobes, similar to subliminal processing (Dehaene et al., 2006) associated with weak stimulus strength. As we increase the input, activity starts to emerge in the temporal lobe, including areas in the ventral stream, and is followed by parietal activation including area 7, which is involved in visuo-motor coordination. Activity across areas in the prefrontal cortex, a necessary requirement for conscious processing (Dehaene et al., 2006), emerges at similarly high values of input current (Dehaene et al., 2006; Figure 7A). Simultaneous emergence (as in Dehaene et al., 2014)

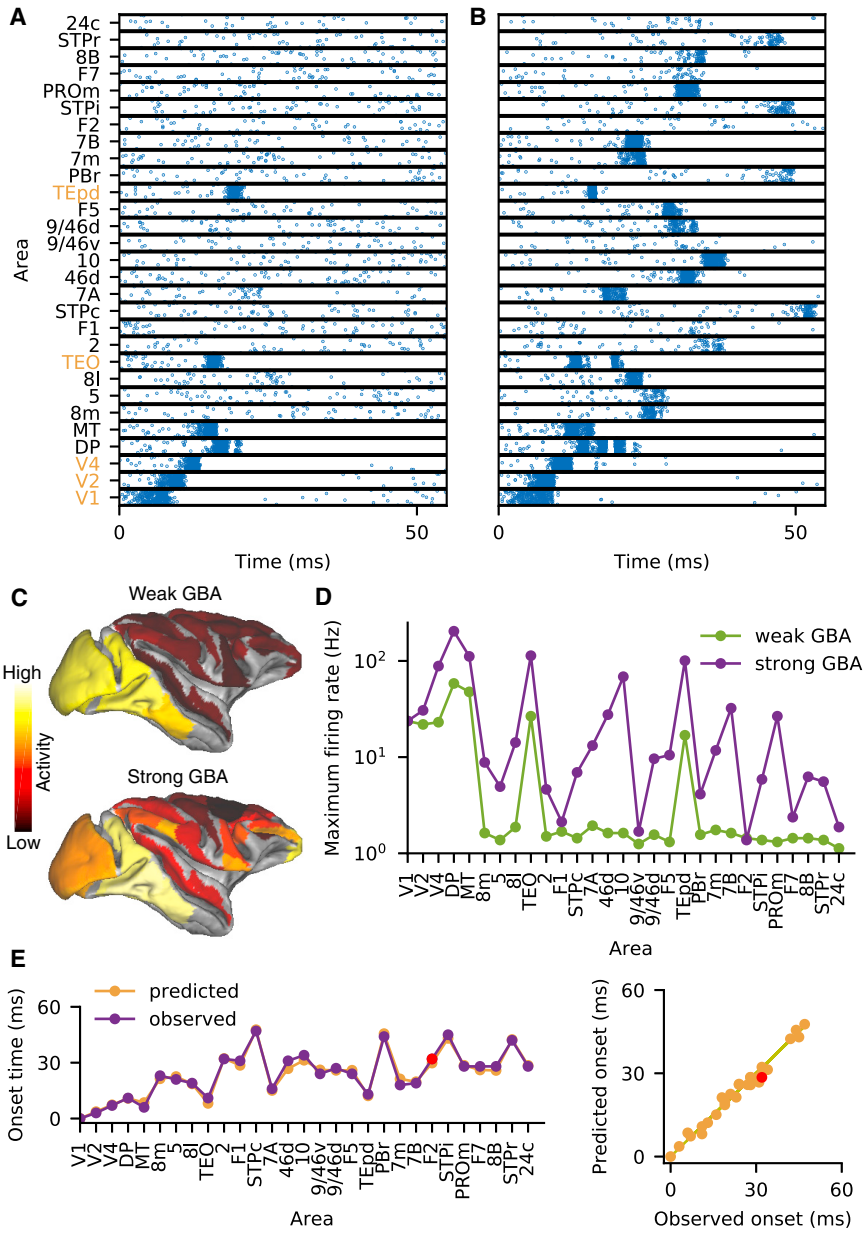

Figure 6. Reliable Signal Propagation in the Synchronous Regime in a Spiking Network Model

(A) Response to a brief input to V1 as it propagates along the hierarchy. The areas along the ventral stream showing strong response activity are indicated with orange labels.

(B) Similar to (A) but with strong GBA.

(C) Spatial activity pattern across the macaque cortical surface corresponding to parameters in (A) and (B).

(D) Peak firing rate across areas as the response to the pulse input to V1 propagates along the hierarchy for weak (green) and strong (purple) GBA.

(E) Predicted and observed response onset times (left) across areas as the signal propagates along the hierarchy. Area F2 which fails to show response activity is indicated in red. Predicted and observed onset times (right) from (B) where each dot indicates an area. The line indicates the diagonal $y = x$. See also Figures S5 and S6.

is observed across several prefrontal areas, including dlPFC, frontopolar cortex, and frontal eye fields (Figure 7A). These areas lie in the set of cliques comprising the densely connected cortical core (Markov et al., 2013). A more careful quantification of the global ignition phenomenon observed in our model (Figure S7) reveals a sudden jump in the number of activated areas on increasing input current strength beyond a certain threshold.

When feedback projections are removed from the large-scale network (Figure 7B), we observe that the signal fails to reach prefrontal areas even for strong input current. Activity becomes restricted largely to the occipito-temporal region in the early visual areas and along the ventral stream. This activity pattern in the absence of feedback projections is similar to that of preconscious processing, which is associated with the absence of top-down attention (Dehaene et al., 2006).

Finally, we tested whether the quantitative anatomical connectivity structure has a key role in determining the activity pattern across the cortical lobes observed in our model (Figure 7A). To do so, we randomly rewired the anatomical projection strengths (only for those projections with non-zero connection weights) while maintaining the network topology. With scrambled quantitative connectivity, the response curve as a function of input strength becomes similar across the temporal, parietal, and frontal lobes (Figure 7C), and the threshold effect disappears. This finding indicates that the quantitative connectivity structure has a critical role in the emergence of lobe-specific activity and the higher input threshold for prefrontal activity; the heterogeneity across areas and the network topology are not enough, by themselves, to explain these phenomena.

DISCUSSION

Transmission of signals is essential for neural coding of external stimuli, and it represents a major topic about the dynamical operation of a multi-regional cortical system (Perkel and Bullock, 1968; Kumar et al., 2010). However, most previous models for signal propagation used a purely feedforward network architecture, in contrast to the recent report that about half of inter-areal cortical connectivity consists of feedback projections (Markov et al., 2014a; Figure S1). We re-examine the long-standing problem of signal transmission in large-scale circuit models that

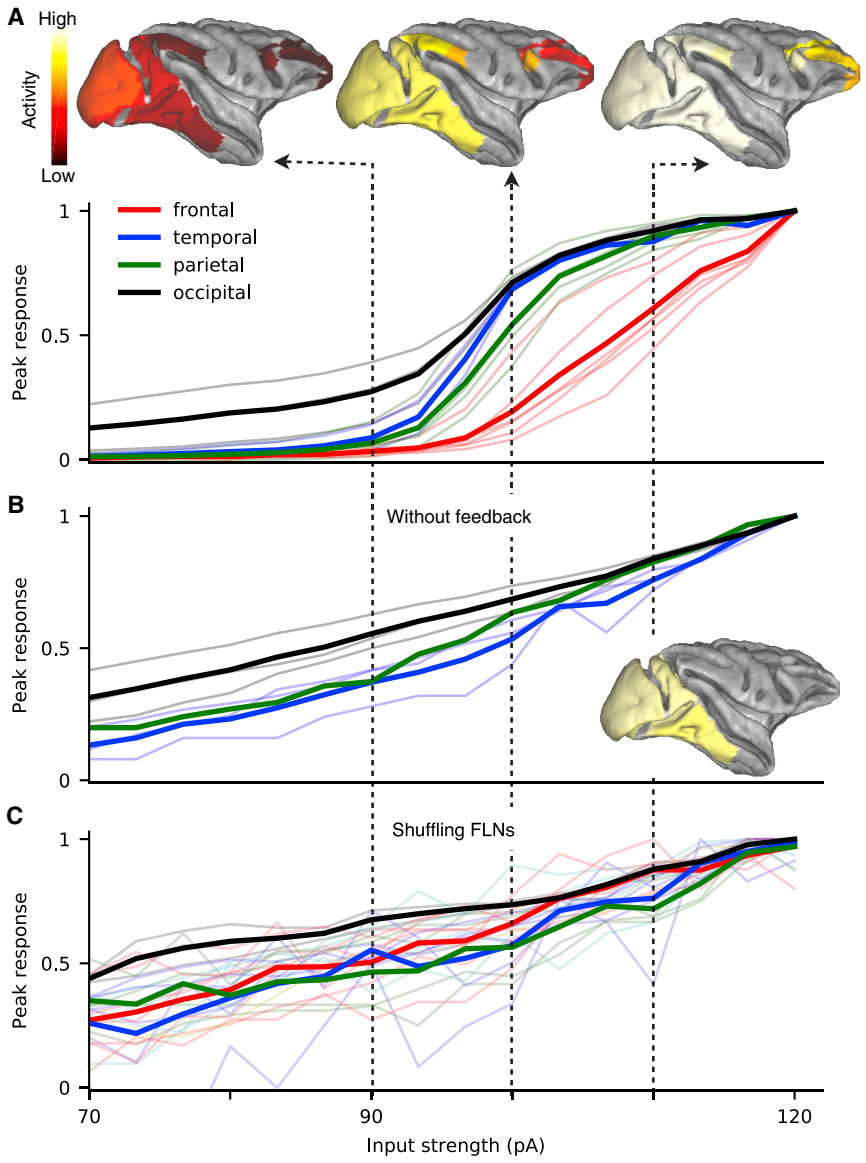

Figure 7. Threshold Crossing of Input Strength Is Required to Engage the Parietal and Frontal Lobes, in Support of the Global Ignition Model of Consciousness

(A) Normalized peak population activity (lower) in the active areas across cortical lobes, as a function the input current strength, using the parameters corresponding to Figure 5C. The light colored lines correspond to individual areas. The thick lines show the mean activity across all the active areas in each cortical lobe, revealing that activation of the temporal lobe (blue), parietal lobe (green), and frontal lobe (red) requires threshold crossing of input strength. Spatial activity (upper; indicated by color) across the macaque cortical surface, for the current strength of 90, 100, 110 pA respectively.

(B) Similar to (A) (left) with the deletion of all feedback projections in the model. Spatial activity pattern (inset) of the macaque cortical surface at input strength of 110 pA, to be compared with the intact model (top right in A).

(C) Similar to (A) with scrambling the long-range inter-areal connectivity while maintaining the network topology. In both (B) and (C) the input threshold crossing for activation of association areas is no longer present, indicating the importance of the inter-areal cortical connectomics of macaque monkey including feedback loops for this phenomenon. See also Figure S7.

incorporate heterogeneity across areas and mesoscopic inter-areal connectivity data (Markov et al., 2014a). The central idea of this paper is a generalized, balanced amplification mechanism in which strong long-range excitatory coupling produces a transient amplification of signals balanced by enhanced local inhibitory-to-excitatory strength to ensure network stability. Generalized, balanced amplification provides a solution to the tradeoff between the need for sufficiently strong excitation for reliable signal transmission and the risk that inter-areal recurrent excitation potentially destabilizes the entire system. We found that this mechanism improves signal propagation by up to 100-fold in large-scale network models of population firing rates with (Mejias et al., 2016) and without (Chaudhuri et al., 2015) a cortical laminar structure, as well as a network model of spiking neurons. Furthermore, inter-areal connection strengths along the ventral visual stream areas are stronger than the overall long-distance connection strengths; this underlies more effective signal propagation along the ventral pathway than the dorsal pathway—a prediction that can be tested experimentally. Finally, surprisingly, our model reproduces signature dynamics of "global ignition" associated with conscious report of a sensory stimulus.

Propagation of spiking activity in feedforward networks has been extensively examined theoretically, in terms of the propagation of both asynchronous (Shadlen and Newsome, 1998; van Rossum et al., 2002; Vogels and Abbott, 2009; Cortes and van Vreeswijk, 2015), and synchronous (Diesmann et al., 1999; Kumar et al., 2008) spiking activity. Synchronous propagation in a feedforward chain can be analyzed in detail in terms of temporal jitter (Marsálek et al., 1997), refractoriness (Kistler and Gerstner, 2002), and the distribution of synaptic strengths (Yazdanbakhsh et al., 2002). In addition, feedback from a higher area to the inhibitory interneurons in a lower area has been proposed to improve synchronous propagation in a multilayer network (Moldakarimov et al., 2015). From a more general point of view, Kumar and collaborators (Kumar et al., 2008) studied synchronous and asynchronous propagation in a feedforward network embedded in a recurrent network. Gating mechanisms to control propagation of asynchronous (Vogels and Abbott, 2009) and synchronous (Kremkow et al., 2010) activity have been proposed, based on the balance (Vogels and Abbott, 2009) and latency (Kremkow

et al., 2010) between excitation and inhibition and on other biophysical properties (Mejias and Longtin, 2014). The asynchronous and synchronous modes of propagation have been integrated in a recent study on feedforward propagation (Kumar et al., 2010), which argues that the two modes represent two extremes of a continuum parametrized by the model parameters. While previous theoretical works largely investigate either synchronous or asynchronous propagation, our propagation mechanism is effective for both the modes (Figures 5 and 6).

Another important insight from this work is to elucidate how signal propagation depends on the degree of non-normality of the underlying network connectivity. Signals in a dynamical system by a normal connection matrix will inevitably attenuate or grow exponentially, leading to instability. However, as pointed out in previous work (Murphy and Miller, 2009), connection matrices of biological neural circuits are always non-normal due to Dale's law. When connectivity is governed by strong recurrent excitation balanced by strong inhibition, biological circuits can transiently amplify incoming signals. This phenomenon, termed balanced amplification (Murphy and Miller, 2009), is proposed to ubiquitously contribute to neural dynamics across the brain (Murphy and Miller, 2009). For example, transient amplification of specific activity states based on a general form of this phenomenon has recently been employed to explain complex dynamics in the motor cortex (Hennequin et al., 2014). The corresponding network architecture uses control-theoretic optimization to balance strong excitatory recurrence with inhibition and is referred to as a stability-optimized circuit. The amplification enabled by this circuit facilitates generation of complex muscle movements. Balanced amplification has also recently been shown to improve memory replay through signal amplification in a hippocampus model (Chenkov et al., 2017). Our work extends this basic dynamical motif of balanced amplification characterized by an inhibition-stabilized network (Murphy and Miller, 2009) from a local circuit to a multi-regional large-scale system to explain reliable cortical signal transmission.

Future Extension of our Cortical Circuit Model

In this work, simulations have been limited to one sensory modality at a time. This can be easily extended by considering multiple input modalities simultaneously presented to the model in order to study how signals from different sensory modalities may interact and aid one another. Other mechanisms besides GBA that improve propagation should also be examined. For example, increasing the local recurrent inhibition and balancing it with local excitatory-to-inhibitory strength is expected to have a similar effect. This is because an increase in the excitatory-to-inhibitory connection strength would have a stronger effect along the hierarchy following from the scaling of excitatory projections. This would consequently scale the suppression of inhibition, resulting in a higher disinhibition along the hierarchy and, ultimately, to a relative increase of excitation in higher cortical areas.

While the anatomical projection strengths used in our model (Markov et al., 2014a) span five orders of magnitude (Figure S1), the role of weak connections in propagation remains unclear. For instance, in a phenomenological network model using the same dataset (Ercsey-Ravasz et al., 2013), the authors found that their measure of information transfer was robust to the removal of a significant fraction of the weakest connections and depended largely on the stronger connections. Similarly, we observe a minor change in the propagation ratio after removing the weakest connections in the rate model (Figure S3E), even as the network density is reduced to 50% of its original value. We also observe that the response latencies for synchronous propagation (Figure 6B) can be predicted by thresholding the anatomical connection weights to include only the strongest projections (Figure 6E). Weaker inter-areal projections can, however, be functionally important if they target a small neural population that has a strong impact in the local circuit. More data on cell-type-specific cortical connectivity is needed to shed light on this issue.

While our results provide reasonable response onset latencies in general, the signal propagation along the ventral pathway is faster than the dorsal pathway, which is inconsistent with previous experimental findings (Bullier, 2001; Bullier et al., 2004; Schmolesky et al., 1998; Schroeder et al., 1998). This discrepancy can be explained by several factors: (1) axons along the dorsal pathway have larger diameters, hence a higher action potential conduction velocity (Nowak and Bullier, 1997), a feature not included in our model. (2) Our model does not contain different types of cells within an area; however, magnocellular cells in lateral geniculate nucleus (LGN), which project mainly to the dorsal pathway, already have shorter latencies compared to parvocellular cells (Schmolesky et al., 1998; Bullier, 2001; Nowak and Bullier, 1997). (3) Our model does not account for the high degree of myelination resulting in faster conduction speeds in many areas of the dorsal stream (Bullier, 2001). (4) The structural connectivity data, on which our model is based (Markov et al., 2014a), does not contain areas MST and LIP along the dorsal pathway. Incorporating these areas could strengthen the dorsal pathway and reduce its response latency.

Our large-scale model, like most previous models that study signal propagation, does not perform complex information transformation on its inputs. It is important to study whether our proposed mechanism can help signal propagation in a large-scale system that performs complex computation as well. A good candidate framework for this would be artificial neural networks, which have proven useful to investigate neural mechanisms behind sensory processing and cognitive tasks (Mante et al., 2013; Song et al., 2016, 2017; Yamins et al., 2014). Deep feedforward neural networks suffer from a similar problem of signal propagation to the one described here, and a popular remedy for this problem is to normalize the neural activity at each network layer (Ioffe and Szegedy, 2015). However, the absolute strength of a signal is largely or entirely lost after normalization. Our proposed mechanism allows for transient propagation of input signals while maintaining its overall strength, therefore avoiding the need for normalization in these models. Existing artificial neural networks typically do not differentiate between excitatory and inhibitory neurons, so our mechanism does not directly apply to these models, but extensions have already been proposed (Song et al., 2016). We suggest that future large-scale, excitatory-inhibitory information-processing network models may benefit from our mechanism if they have homeostatic plasticity that balances strong long-range excitatory inputs with strong local inhibitory connections.

Role of Subcortical Structures

Although our model takes into account cortical connectivity data, which forms a large part of the input to cortical areas (Markov et al., 2011), the role of subcortical structures remains to be explored, and emerging subcortical connectivity data should be incorporated. For instance, cortico-thalamo-cortical interactions have been shown to drive robust activity in the higher-order somatosensory cortex (Theyel et al., 2010). Experimental studies have also shown that the pulvinar synchronizes activity (Saalmann et al., 2012) between interconnected cortical areas, indicating its role in regulating information transmission across the visual cortex. Recent work (Cortes and van Vreeswijk, 2015) studies asynchronous cortical transmission through two multilayered feedforward networks and explores the role of long-range inhibitory pulvinar connections in linking multiple cortical stages to boost propagation. While our large-scale models only employ anatomical connectivity from macaque cortex, novel thalamocortical connectivity data from the mouse (Oh et al., 2014) could be used to examine our mechanism in a large-scale corticothalamic model of the mouse brain.

Ignition Theory

Recent work suggests that the emergence of parieto-frontal activity could be viewed as a precursor to conscious perception (Dehaene et al., 2014). Such an event would follow from the input exceeding a threshold, leading to a reverberating neuronal assembly (Dehaene et al., 2006); note, however, that reverberation could be in the form of transients rather than a steady-state attractor (Baria et al., 2017). In this sense, subliminal, preconscious, and conscious processing would be associated with different levels of top-down attention and bottom-up stimulus strength (Dehaene et al., 2006). Our large-scale model, which is able to propagate signals efficiently (Figure 5), can be used as a framework for a computational examination of these phenomena. Subliminal processing, characterized by weak bottom-up activation insufficient to trigger large-scale reverberation (Dehaene et al., 2006), resembles the weak input case in Figure 7A. Preconscious processing is characterized by the absence of top-down attention (Dehaene et al., 2006). Correspondingly, we observe a disruption of global reverberation after removing top-down feedback projections (Figure 7B) that mostly originate from the strongly connected core of prefrontal and association areas (Markov et al., 2013). Besides their role in attention, feedback signals have been proposed to play a major role in discerning intrinsically ambiguous sensory input (Gilbert and Sigman, 2007; Scocchia et al., 2014). For example, top-down input is known to influence perception of visually ambiguous stimuli during binocular rivalry (Paffen and Alais, 2011; Wang et al., 2013a), hence reinforcing the idea suggested by our results that feedback input is crucial for conscious perception. Another important role of feedback has been proposed in the context of predictive coding, where top-down influences may be seen as "predictions" aimed to match the incoming bottom-up sensory stimuli (Rao and Ballard, 1999; Bastos et al., 2012). A mismatch in those top-down predictions would lead to a "prediction error" signal moving downstream of the hierarchy toward association areas. One could expect that, in the presence of top-down prediction signals, partial mismatches with the incoming sensory input could be fairly common, leading to relatively small prediction errors. Although our modeling framework does not have the level of structural detail needed to explain this complex matching process between different signal streams, it could in principle provide a mechanism for these small signals to efficiently propagate across the cortical hierarchy by means of the GBA principle. In some experiments on conscious processing (Dehaene et al., 2006; Del Cul et al., 2007; Dehaene and Changeux, 2011; Dehaene et al., 2014), the input strength can be adjusted such that the stimulus is detected stochastically 50% of the time. Further research is needed to test whether spiking stochasticity in our system could lead the model to such switching for a fixed input strength.

Although both conscious perception and attention involve an interplay between bottom-up and top-down signaling, attention and consciousness are dissociable (Li et al., 2002; Koch and Tsuchiya, 2007; Wyart and Tallon-Baudry, 2008). In the intact network model, strong input elicits activity across several prefrontal and parietal areas (Figure 7A). However, presently it is still a matter of debate whether conscious report necessarily requires the engagement of the prefrontal cortex (Tsuchiya et al., 2015; Boly et al., 2017). Regardless, a complete characterization of the observed neural activity related to conscious perception is beyond the scope of this work; future extensions of our network model would constitute a powerful tool to understand this phenomenon. In particular, NMDA receptors may be more prominent at synapses of top-down projections than those of bottom-up projections (Self et al., 2012). Implementing NMDA dynamics in the spiking network model could provide insight on the nature of the top-down amplification of posterior areas following prefrontal activation, as observed in conscious perception tasks (Dehaene et al., 2006). In addition, NMDA synapses could contribute to a stronger sigmoidal-shaped activation of frontal areas during the ignition phenomenon described in Figure 7A, given the important role of NMDA receptors in persistent activity in frontal areas (Wang et al., 2013b; Wang, 1999).

We are ushering in a new era of understanding large-scale brain systems beyond local circuits. Whereas brain connectomics is essential, structural connectivity is insufficient to predict dynamical behavior of recurrent neural circuits. Our work on reliable signal propagation offers another demonstration of this principle and represents an important step in our investigations of cognitive processes in a large-scale brain circuit.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- METHOD DETAILS
  - The rate model
  - Heterogeneity across areas
  - Local balanced amplification
  - Steady-state excitatory rate along stability boundary in 2 population case

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and can be found with this article online at https://doi.org/10.1016/j.neuron.2018.02.031.

ACKNOWLEDGMENTS

Funding was provided by The Swartz foundation, ONR Grant no. N00014-17-1-2041, Simons Collaborative Global Brain (SCGB) Program Grant 543057SPI to X.-J.W., STCSM grant 15JC1400104, and a Samuel J. and Joan B. Williamson Fellowship to G.R.Y. The authors thank Biyu He for her comments on the manuscript.

AUTHOR CONTRIBUTIONS

M.R.J., J.F.M., G.R.Y., and X.-J.W. designed the research, had regular discussions throughout the project, and wrote the manuscript. M.R.J. performed the computational research.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 21, 2017
Revised: December 27, 2017
Accepted: February 27, 2018
Published: March 22, 2018

REFERENCES

Barbas, H., and Rempel-Clower, N. (1997). Cortical structure predicts the pattern of corticocortical connections. Cereb. Cortex *7*, 635–646.

Baria, A.T., Maniscalco, B., and He, B.J. (2017). Initial-state-dependent, robust, transient neural dynamics encode conscious visual perception. PLoS Comput. Biol. *13*, e1005806.

Barone, P., Batardiere, A., Knoblauch, K., and Kennedy, H. (2000). Laminar distribution of neurons in extrastriate areas projecting to visual areas V1 and V4 correlates with the hierarchical rank and indicates the operation of a distance rule. J. Neurosci. *20*, 3263–3281.

Bastos, A.M., Usrey, W.M., Adams, R.A., Mangun, G.R., Fries, P., and Friston, K.J. (2012). Canonical microcircuits for predictive coding. Neuron *76*, 695–711.

Bastos, A.M., Vezoli, J., Bosman, C.A., Schoffelen, J.M., Oostenveld, R., Dowdall, J.R., De Weerd, P., Kennedy, H., and Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. Neuron *85*, 390–401.

Binzegger, T., Douglas, R.J., and Martin, K.A. (2009). Topology and dynamics of the canonical circuit of cat V1. Neural Netw. *22*, 1071–1078.

Boly, M., Massimini, M., Tsuchiya, N., Postle, B.R., Koch, C., and Tononi, G. (2017). Are the neural correlates of consciousness in the front or in the back of the cerebral cortex? Clinical and neuroimaging evidence. J. Neurosci. *37*, 9603–9613.

Brunel, N. (2000a). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. J. Comput. Neurosci. *8*, 183–208.

Brunel, N. (2000b). Persistent activity and the single-cell frequency-current curve in a cortical network model. Network *11*, 261–280.

Bullier, J. (2001). Integrated model of visual processing. Brain Res. Brain Res. Rev. *36*, 96–107.

Bullier, J., Deriche, R., Faugeras, O., Girard, P., Kornprobst, P., Papadopoulo, T., Viéville, T., Thorpe, S., Fize, D., and Guyonneau, R. (2004). RIVAGe Feedback during Visual Integration: towards a Generic Architecture. Research Report RR-5451. INRIA.

Chaudhuri, R., Knoblauch, K., Gariel, M.-A., Kennedy, H., and Wang, X.-J. (2015). A large-scale circuit mechanism for hierarchical dynamical processing in the primate cortex. Neuron *88*, 419–431.

Chenkov, N., Sprekeler, H., and Kempter, R. (2017). Memory replay in balanced recurrent networks. PLoS Comput. Biol. *13*, e1005359.

Cortes, N., and van Vreeswijk, C. (2015). Pulvinar thalamic nucleus allows for asynchronous spike propagation through the cortex. Front. Comput. Neurosci. *9*, 60.

D'Souza, R.D., Meier, A.M., Bista, P., Wang, Q., and Burkhalter, A. (2016). Recruitment of inhibition and excitation across mouse visual cortex depends on the hierarchy of interconnecting areas. eLife *5*, e19332.

Dehaene, S., and Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. Neuron *70*, 200–227.

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., and Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. Trends Cogn. Sci. *10*, 204–211.

Dehaene, S., Charles, L., King, J.-R., and Marti, S. (2014). Toward a computational theory of conscious processing. Curr. Opin. Neurobiol. *25*, 76–84.

Del Cul, A., Baillet, S., and Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. PLoS Biol. *5*, e260.

DeMarse, T.B., Pan, L., Alagapan, S., Brewer, G.J., and Wheeler, B.C. (2016). Feed-forward propagation of temporal and rate information between cortical populations during coherent activation in engineered in vitro networks. Front. Neural Circuits *10*, 32.

Diesmann, M., Gewaltig, M.O., and Aertsen, A. (1999). Stable propagation of synchronous spiking in cortical neural networks. Nature *402*, 529–533.

Elston, G.N. (2007). Specialization of the neocortical pyramidal cell during primate evolution. In Evolution of Nervous Systems, eds., J.H. Kass and T.M. Preuss, eds. (Elsevier), pp. 191–242.

Ercsey-Ravasz, M., Markov, N.T., Lamy, C., Van Essen, D.C., Knoblauch, K., Toroczkai, Z., and Kennedy, H. (2013). A predictive network model of cerebral cortical connectivity based on a distance rule. Neuron *80*, 184–197.

Felleman, D.J., and Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex *1*, 1–47.

Gilbert, C.D., and Sigman, M. (2007). Brain states: top-down influences in sensory processing. Neuron *54*, 677–696.

Golomb, D., Wang, X.J., and Rinzel, J. (1994). Synchronization properties of spindle oscillations in a thalamic reticular nucleus model. J. Neurophysiol. *72*, 1109–1126.

Goodman, D.F., and Brette, R. (2009). The brian simulator. Front. Neurosci. *3*, 192–197.

Hennequin, G., Vogels, T.P., and Gerstner, W. (2014). Optimal control of transient dynamics in balanced networks supports generation of complex movements. Neuron *82*, 1394–1406.

Henrici, P. (1962). Bounds for iterates, inverses, spectral variation and fields of values of non-normal matrices. Numer. Math. *4*, 24–40.

Histed, M.H., Pasupathy, A., and Miller, E.K. (2009). Learning substrates in the primate prefrontal cortex and striatum: sustained activity related to successful actions. Neuron *63*, 244–253.

Ioffe, S., and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. Proceedings of the 32nd International Conference on Machine Learning, PMLR *37*, 448–456.

King, J.R., Pescetelli, N., and Dehaene, S. (2016). Brain mechanisms underlying the brief maintenance of seen and unseen sensory information. Neuron *92*, 1122–1134.

Kistler, W.M., and Gerstner, W. (2002). Stable propagation of activity pulses in populations of spiking neurons. Neural Comput. *14*, 987–997.

Koch, C., and Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. Trends Cogn. Sci. *11*, 16–22.

Kopell, N., Ermentrout, G.B., Whittington, M.A., and Traub, R.D. (2000). Gamma rhythms and beta rhythms have different synchronization properties. Proc. Natl. Acad. Sci. USA *97*, 1867–1872.

Kremkow, J., Aertsen, A., and Kumar, A. (2010). Gating of signal propagation in spiking neural networks by balanced and correlated excitation and inhibition. J. Neurosci. *30*, 15760–15768.

Kumar, A., Rotter, S., and Aertsen, A. (2008). Conditions for propagating synchronous spiking and asynchronous firing rates in a cortical network model. J. Neurosci. *28*, 5268–5280.

Kumar, A., Rotter, S., and Aertsen, A. (2010). Spiking activity propagation in neuronal networks: reconciling different perspectives on neural coding. Nat. Rev. Neurosci. *11*, 615–627.

Li, F.F., VanRullen, R., Koch, C., and Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. Proc. Natl. Acad. Sci. USA *99*, 9596–9601.

Luczak, A., McNaughton, B.L., and Harris, K.D. (2015). Packet-based communication in the cortex. Nat. Rev. Neurosci. *16*, 745–755.

Mante, V., Sussillo, D., Shenoy, K.V., and Newsome, W.T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature *503*, 78–84.

Markov, N.T., Misery, P., Falchier, A., Lamy, C., Vezoli, J., Quilodran, R., Gariel, M.A., Giroud, P., Ercsey-Ravasz, M., Pilaz, L.J., et al. (2011). Weight consistency specifies regularities of macaque cortical networks. Cereb. Cortex *21*, 1254–1272.

Markov, N.T., Ercsey-Ravasz, M., Van Essen, D.C., Knoblauch, K., Toroczkai, Z., and Kennedy, H. (2013). Cortical high-density counterstream architectures. Science *342*, 1238406.

Markov, N.T., Ercsey-Ravasz, M.M., Ribeiro Gomes, A.R., Lamy, C., Magrou, L., Vezoli, J., Misery, P., Falchier, A., Quilodran, R., Gariel, M.A., et al. (2014a). A weighted and directed interareal connectivity matrix for macaque cerebral cortex. Cereb. Cortex *24*, 17–36.

Markov, N.T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., Lamy, C., Misery, P., Giroud, P., Ullman, S., et al. (2014b). Anatomy of hierarchy: feedforward and feedback pathways in macaque visual cortex. J. Comp. Neurol. *522*, 225–259.

Marsálek, P., Koch, C., and Maunsell, J. (1997). On the relationship between synaptic input and spike output jitter in individual neurons. Proc. Natl. Acad. Sci. USA *94*, 735–740.

Mejias, J.F., and Longtin, A. (2014). Differential effects of excitatory and inhibitory heterogeneity on the gain and asynchronous state of sparse cortical networks. Front. Comput. Neurosci. *8*, 107.

Mejias, J.F., Murray, J.D., Kennedy, H., and Wang, X.-J. (2016). Feedforward and feedback frequency-dependent interactions in a large-scale laminar network of the primate cortex. Sci. Adv. *2*, e1601335.

Moldakarimov, S., Bazhenov, M., and Sejnowski, T.J. (2015). Feedback stabilizes propagation of synchronous spiking in cortical neural networks. Proc. Natl. Acad. Sci. USA *112*, 2545–2550.

Murphy, B.K., and Miller, K.D. (2009). Balanced amplification: a new mechanism of selective amplification of neural activity patterns. Neuron *61*, 635–648.

Nowak, L.G., and Bullier, J. (1997). The timing of information transfer in the visual system. In Extrastriate Visual Cortex in Primates, eds., K.S. Rockland, J.H. Kaas, and A. Peters, eds. (Springer), pp. 205–241.

Oh, S.W., Harris, J.A., Ng, L., Winslow, B., Cain, N., Mihalas, S., Wang, Q., Lau, C., Kuan, L., Henry, A.M., et al. (2014). A mesoscale connectome of the mouse brain. Nature *508*, 207–214.

Ozeki, H., Finn, I.M., Schaffer, E.S., Miller, K.D., and Ferster, D. (2009). Inhibitory stabilization of the cortical network underlies visual surround suppression. Neuron *62*, 578–592.

Paffen, C.L.E., and Alais, D. (2011). Attentional modulation of binocular rivalry. Front. Hum. Neurosci. *5*, 105.

Perkel, D.H., and Bullock, T.H. (1968). Neural coding: A report based on an nrp work session. Neurosci. Res. Program Bull. *6*, 219–349.

Rao, R.P., and Ballard, D.H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat. Neurosci. *2*, 79–87.

Reyes, A.D. (2003). Synchrony-dependent propagation of firing rate in iteratively constructed networks in vitro. Nat. Neurosci. *6*, 593–599.

Romo, R., Brody, C.D., Hernández, A., and Lemus, L. (1999). Neuronal correlates of parametric working memory in the prefrontal cortex. Nature *399*, 470–473.

Saalmann, Y.B., Pinsk, M.A., Wang, L., Li, X., and Kastner, S. (2012). The pulvinar regulates information transmission between cortical areas based on attention demands. Science *337*, 753–756.

Schmolesky, M.T., Wang, Y., Hanes, D.P., Thompson, K.G., Leutgeb, S., Schall, J.D., and Leventhal, A.G. (1998). Signal timing across the macaque visual system. J. Neurophysiol. *79*, 3272–3278.

Schroeder, C.E., Mehta, A.D., and Givre, S.J. (1998). A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. Cereb. Cortex *8*, 575–592.

Scocchia, L., Valsecchi, M., and Triesch, J. (2014). Top-down influences on ambiguous perception: the role of stable and transient states of the observer. Front. Hum. Neurosci. *8*, 979.

Self, M.W., Kooijmans, R.N., Supèr, H., Lamme, V.A., and Roelfsema, P.R. (2012). Different glutamate receptors convey feedforward and recurrent processing in macaque V1. Proc. Natl. Acad. Sci. USA *109*, 11031–11036.

Shadlen, M.N., and Newsome, W.T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. J. Neurosci. *18*, 3870–3896.

Song, S., Sjöström, P.J., Reigl, M., Nelson, S., and Chklovskii, D.B. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. PLoS Biol. *3*, e68.

Song, H.F., Yang, G.R., and Wang, X.-J. (2016). Training excitatory-inhibitory recurrent neural networks for cognitive tasks: A simple and flexible framework. PLoS Comput. Biol. *12*, e1004792.

Song, H.F., Yang, G.R., and Wang, X.-J. (2017). Reward-based training of recurrent neural networks for cognitive and value-based tasks. eLife *6*, e21492.

Swadlow, H.A. (1990). Efferent neurons and suspected interneurons in S-1 forelimb representation of the awake rabbit: receptive fields and axonal properties. J. Neurophysiol. *63*, 1477–1498.

Theyel, B.B., Llano, D.A., and Sherman, S.M. (2010). The corticothalamocortical circuit drives higher-order cortex in the mouse. Nat. Neurosci. *13*, 84–88.

Tsuchiya, N., Wilke, M., Frässle, S., and Lamme, V.A.F. (2015). No-Report Paradigms: Extracting the True Neural Correlates of Consciousness. Trends Cogn. Sci. *19*, 757–770.

Vaadia, E., Haalman, I., Abeles, M., Bergman, H., Prut, Y., Slovin, H., and Aertsen, A. (1995). Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. Nature *373*, 515–518.

Van Essen, D.C., Drury, H.A., Dickson, J., Harwell, J., Hanlon, D., and Anderson, C.H. (2001). An integrated software suite for surface-based analyses of cerebral cortex. J. Am. Med. Inform. Assoc. *8*, 443–459.

van Rossum, M.C., Turrigiano, G.G., and Nelson, S.B. (2002). Fast propagation of firing rates through layered networks of noisy neurons. J. Neurosci. *22*, 1956–1966.

Vogels, T.P., and Abbott, L.F. (2009). Gating multiple signals through detailed balance of excitation and inhibition in spiking networks. Nat. Neurosci. *12*, 483–491.

Wang, X.-J. (1999). Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory. J. Neurosci. *19*, 9587–9603.

Wang, X.-J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. Physiol. Rev. *90*, 1195–1268.

Wang, X.-J., and Kennedy, H. (2016). Brain structure and dynamics across scales: in search of rules. Curr. Opin. Neurobiol. *37*, 92–98.

Wang, M., Arteaga, D., and He, B.J. (2013a). Brain mechanisms for simple perception and bistable perception. Proc. Natl. Acad. Sci. USA *110*, E3350–E3359.

Wang, M., Yang, Y., Wang, C.J., Gamo, N.J., Jin, L.E., Mazer, J.A., Morrison, J.H., Wang, X.-J., and Arnsten, A.F. (2013b). NMDA receptors subserve persistent neuronal firing during working memory in dorsolateral prefrontal cortex. Neuron *77*, 736–749.

Wyart, V., and Tallon-Baudry, C. (2008). Neural dissociation between visual awareness and spatial attention. J. Neurosci. *28*, 2667–2679.

Yamins, D.L.K., Hong, H., Cadieu, C.F., Solomon, E.A., Seibert, D., and DiCarlo, J.J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. Proc. Natl. Acad. Sci. USA *111*, 8619–8624.

Yazdanbakhsh, A., Babadi, B., Rouhani, S., Arabzadeh, E., and Abbassian, A. (2002). New attractor states for synchronous activity in synfire chains with excitatory and inhibitory coupling. Biol. Cybern. *86*, 367–378.

Zingg, B., Hintiryan, H., Gou, L., Song, M.Y., Bay, M., Bienkowski, M.S., Foster, N.N., Yamashita, S., Bowman, I., Toga, A.W., and Dong, H.W. (2014). Neural networks of the mouse neocortex. Cell *156*, 1096–1111.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
Brian2	Goodman and Brette (2009). The Brian simulator. Front Neurosci https://doi.org/10.3389/neuro.01.026.2009	https://brian2.readthedocs.io/en/stable/
Anatomical Connectivity data	Markov et al. (2014a) Cereb. Cortex	http://core-nets.org/

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Xiao-Jing Wang (xjwang@nyu.edu).

METHOD DETAILS

The rate model

The model examined in Figures 2, 3, based on a recently published large-scale model of the macaque cortex (Chaudhuri et al., 2015), is described below. The dynamics of each area is described as a threshold linear recurrent network, with interacting excitatory and inhibitory populations, as follows

$$\tau_E \frac{d\nu_E}{dt} = -\nu_E + \beta_E [I_E]_+$$

$$\tau_I \frac{d\nu_I}{dt} = -\nu_I + \beta_I [I_I]_+$$

where $[I_E]_+ = \max(I_E, 0)$. ν_E, ν_I denote the firing rates of the excitatory and inhibitory populations respectively, and τ_E, τ_I are the corresponding intrinsic time constants. β_E, β_I are the slopes of the f-I curves. The local-microcircuit is qualitatively the same across areas, with quantitative inter-areal differences as stated below.

Heterogeneity across areas

The laminar pattern of inter-areal projections is used to assign a hierarchical position to each area (Barone et al., 2000; Markov et al., 2014b; Chaudhuri et al., 2015). This is based on the notion that feedforward projections tend to originate from the superficial cortical layer, and feedback projections from the deep layer (Barbas and Rempel-Clower, 1997). Thus the hierarchical distance between a source and target area is computed based on the fraction of projections originating in the superficial layer of the source area (Chaudhuri et al., 2015). An area's hierarchical position is found to be strongly correlated with the number of basal-dendritic spines of layer 3 pyramidal neurons in that area (Chaudhuri et al., 2015; Elston, 2007). The pyramidal-cell spine count increases with the hierarchical position of the cortical area by a factor of 6∼7 (Elston, 2007; Chaudhuri et al., 2015), and it is used as a proxy for the total excitatory drive. Thereby, heterogeneity across areas is introduced in the form of a gradient of excitatory connection strengths along the hierarchy (Chaudhuri et al., 2015).

Inter-areal projection strength is based on a recently published anatomical connectivity dataset from the macaque cortex (Markov et al., 2014a). The data measures the number of neurons labeled by a retrograde tracer injected in 29 widely distributed cortical areas. To control for the injection size, the neuron counts are normalized by the net number of neurons labeled by the injection, resulting in an FLN (Fraction of Labeled Neurons) across two areas. Thus, given areas *i*,*j*, the FLN_{ij} is the number of neurons projecting from area *j* to area *i* weighted by the net number of neurons projecting to area *i* from all the areas. The net incoming current is given by

$$I_E^i = (1 + \eta h_i)\left(w_{EE}\nu_E^i + I_{lr,E}^i\right) - w_{EI}\nu_I^i + I_{ext,E}^i$$

$$I_I^i = (1 + \eta h_i)\left(w_{IE}\nu_E^i + I_{lr,I}^i\right) - w_{II}\nu_I^i + I_{ext,I}^i$$

where I_E^i, I_I^i denote the input currents to the excitatory and inhibitory populations respectively for area *i* and w_{ij} denotes the local-circuit connection strength from the population type *j* to the population type *i*. $I_{lr,E}^i, I_{lr,I}^i$ correspond to the long-range input currents,

assumed to be purely excitatory. $I^i_{ext,E}, I^i_{ext,I}$ correspond to the external inputs. The hierarchical position h_i is normalized to lie between 0 and 1. η scales the excitatory connection strengths based on the hierarchical position of the area. We set η to 0.68 (Chaudhuri et al., 2015). The background firing rate, is set to an excitatory rate of 10 Hz, and an inhibitory rate of 35 Hz (Chaudhuri et al., 2015). The background rate is subtracted when monitoring the response activity in Figures 2 and 3 and Figure S3. The parameter values are set as $\tau_E = 20, \tau_I = 10$ (in ms), $\beta_E = 0.066, \beta_I = 0.351$ (Chaudhuri et al., 2015; Binzegger et al., 2009). The long-range input currents are given by,

$$I^i_{lr,E} = \mu_{EE} \sum_{j=1}^{N} FLN_{ij} \nu^j_E$$

$$I^i_{lr,I} = \mu_{IE} \sum_{j=1}^{N} FLN_{ij} \nu^j_E.$$

Thus, the inter-areal connectivity depends on the corresponding FLNs and is scaled by the global scaling parameters μ_{EE} and μ_{IE} corresponding to long-range E to E and long-range E to I coupling respectively. The connectivity parameters are set as $\mu_{IE} = 25.3, w_{EE} = 24.3, w_{IE} = 12.2, w_{II} = 12.5$ (in pA/Hz) (Chaudhuri et al., 2015). The connectivity parameters corresponding to the local I to E connection strength and the global excitatory coupling are (in pA/Hz) $w_{EI} = 19.7, \mu_{EE} = 33.7$ for weak GBA and $w_{EI} = 25.2, \mu_{EE} = 51.5$ for strong GBA (Figure 3).

Local balanced amplification

For Figure 1B, recurrent excitation and lateral inhibition values are (6, 6.7) and (4.45, 4.7) for strong LBA and weak LBA respectively. The excitatory to inhibitory connection strength, and the inhibitory recurrent strength values are fixed at 4.29, 4.71 respectively (Figures 1B and 1C).

Steady-state excitatory rate along stability boundary in 2 population case

Here we consider the dynamics of a 2 population nonlinear rate model wherein the local circuit lies in an inhibition stabilized regime, that is, it has strong recurrent excitation balanced by strong lateral inhibition. We show that for this circuit, increasing the recurrent excitation and balancing the system with stronger inhibition results in a higher steady state excitatory firing rate on moving along the stability boundary in parameter space. This provides an intuitive understanding of the higher transient amplification following from stronger local balanced amplification as seen in Figure 1B.

Consider a local circuit lying in an inhibition-stabilized regime ($w_{EE}, w_{EI} > 1$, where w_{ij} is the connection strength from population type j to population type i) as in Figure 1A, with excitatory and inhibitory firing rates denoted by E, I respectively. Say the excitatory population is acted upon by a delta-pulse input, to set the initial firing rate to some $E_0 > 0$ while $I_0 = 0$. We henceforth refer to this initial condition as $\{E_0, I_0\}$.

The dynamics is described by,

$$\begin{aligned} \tau \frac{dE}{dt} &= -E + [w_{EE}E - w_{EI}I]_+ \\ \tau \frac{dI}{dt} &= -I + [w_{IE}E - w_{II}I]_+ \end{aligned} \tag{1}$$

where τ denotes the intrinsic time constant and $[x]_+ = \max(x, 0)$ denotes rectification.

For fixed w_{IE}, w_{II}, the system stability depends on the local-circuit recurrent excitation w_{EE} and the feedback-inhibition w_{EI}. For our analysis, we assume that w_{IE}, w_{II} are fixed. We define the $w_{EE} - w_{EI}$ stability boundary as the set of points in the $w_{EE} - w_{EI}$ parameter space for which given the initial condition $\{E_0, I_0\}$, the system evolves to a non-zero steady state. We want to show that,

Given w_{IE}, w_{II}, the steady state E value corresponding to the $w_{EE} - w_{EI}$ stability boundary (Figure 1C) increases on moving along the direction of increasing w_{EE}, w_{EI}.

Let us compute the solutions for the steady state of the system. Denote steady state excitatory and inhibitory values as E_s, I_s respectively. Then besides the trivial steady state, we show that the only other solution is when both $(w_{IE}E - w_{II}I) > 0$ and $(w_{EE}E - w_{EI}I) > 0$. Say this were not true. Then consider cases

$$\tau \frac{dE}{dt} = -E, \ \tau \frac{dI}{dt} = -I + [w_{IE}E - w_{II}I]_+$$

$$\tau \frac{dE}{dt} = -E + [w_{EE}E - w_{EI}I]_+, \ \tau \frac{dI}{dt} = -I$$

But both the above cases would correspond to $\{E_s, I_s\} = \{0, 0\}$. Thus, besides the trivial solution, any other steady state would have $E_s > 0, I_s > 0$.

Let us solve for the non-zero steady state. Then $\{E_s, I_s\}$ lies in an open set for which $w_{IE}E - w_{II}I > 0$ and $w_{EE}E - w_{EI}I > 0$. Now

$$\tau\frac{dE}{dt} = -E + [w_{EE}E - w_{EI}I] = 0,\ \tau\frac{dI}{dt} = -I + [w_{IE}E - w_{II}I] = 0$$

gives $I_s = E_s(w_{IE}/1 + w_{II}) = E_s(w_{EE} - 1/w_{EI})$. Say $c = (1 + w_{II})/w_{IE} = w_{EI}/(w_{EE} - 1)$. Then c is a constant since w_{IE}, w_{II} are fixed. Let,

$$A = (1/\tau)\begin{bmatrix} w_{EE} - 1 & -w_{EI} \\ w_{IE} & -1 - w_{II} \end{bmatrix}$$

We evaluate the behavior around this steady state by finding the eigenvalues of matrix A. The eigenvalue equation is,

$$\lambda^2\tau^2 + \lambda\tau(2 - w_{EE} + w_{II}) + w_{EI}w_{IE} - (w_{EE} - 1)(1 + w_{II}) = 0$$

Since $c = (1 + w_{II})/w_{IE} = w_{EI}/(w_{EE} - 1)$, the equation reduces to $\lambda^2\tau^2 + \lambda\tau(2 - w_{EE} + w_{II}) = 0$ and thus, $\lambda_1 = 0$, $\lambda_2 = -(2 - w_{EE} + w_{II})/\tau$.

Thus this steady state corresponds to real eigenvalues, one of them being 0, and cannot have a pair of complex conjugate eigenvalues. The eigenvectors corresponding to λ_1, λ_2 will be

$$v_1 = \begin{bmatrix} c \\ 1 \end{bmatrix}, v_2 = \begin{bmatrix} \frac{w_{EI}}{1 + w_{II}} \\ 1 \end{bmatrix}.$$

We can simply set the second entry of each eigenvector to 1 because w_{IE} and w_{II} are fixed. Thus in some open set around the steady state, the dynamics are given by,

$$\begin{bmatrix} E(t) \\ I(t) \end{bmatrix} = a_1 e^{\lambda_1 t} v_1 + a_2 e^{\lambda_2 t} v_2 = a_1 \begin{bmatrix} c \\ 1 \end{bmatrix} + a_2 e^{\lambda_2 t} \begin{bmatrix} \frac{w_{EI}}{1 + w_{II}} \\ 1 \end{bmatrix}$$

where a_1, a_2 depend on the initial condition of the system and $\lambda_2 = -(2 - w_{EE} + w_{II})/\tau$.

We now show that when the system has a non-zero steady state, the dynamics of the system for initial condition $\{E_0, I_0\}$ evolves according to the equation above. Thus, the system trajectory does not encounter "rectification" since $w_{IE}E - w_{II}I > 0, w_{EE}E - w_{EI}I > 0$ for $t > 0$. Note that $w_{IE}E_0 - w_{II}I_0 > 0, w_{EE}E_0 - w_{EI}I_0 > 0$. Thus, the initial dynamics evolves according to

$$\begin{bmatrix} E(t) \\ I(t) \end{bmatrix} = a_1 \begin{bmatrix} c \\ 1 \end{bmatrix} + a_2 e^{\lambda_2 t} \begin{bmatrix} \frac{w_{EI}}{1 + w_{II}} \\ 1 \end{bmatrix} = \frac{E_0}{c - \frac{w_{EI}}{1 + w_{II}}} \left(\begin{bmatrix} c \\ 1 \end{bmatrix} - e^{\lambda_2 t} \begin{bmatrix} \frac{w_{EI}}{1 + w_{II}} \\ 1 \end{bmatrix} \right).$$

where $c = (1 + w_{II})/w_{IE} = w_{EI}/(w_{EE}\text{-}1)$ and $\lambda_2 = -(2 - w_{EE} + w_{II})/\tau$. Thus,

$$\begin{bmatrix} E(t) \\ I(t) \end{bmatrix} = \frac{-E_0 w_{IE}}{\lambda_2 \tau} \left(\begin{bmatrix} c \\ 1 \end{bmatrix} - e^{\lambda_2 t} \begin{bmatrix} \frac{w_{EI}}{1 + w_{II}} \\ 1 \end{bmatrix} \right)$$

Thus,

$$\begin{aligned} E(t) &= \frac{-E_0 w_{IE}}{\lambda_2 \tau} \left(c - e^{\lambda_2 t} \frac{w_{EI}}{1 + w_{II}} \right) = \frac{-E_0}{\lambda_2 \tau} \left((1 + w_{II}) - e^{\lambda_2 t}(w_{EE} - 1) \right) \\ I(t) &= \frac{-E_0 w_{IE}}{\lambda_2 \tau} \left(1 - e^{\lambda_2 t} \right) \end{aligned} \tag{2}$$

To show that rectification is not encountered, we need to show that for $t > 0$, $w_{IE}E - w_{II}I > 0$ and $w_{EE}E - w_{EI}I > 0$, that is, $w_{IE}E > w_{II}I$ and $w_{EE}E > w_{EI}I$. We show the case $w_{IE}E > w_{II}I$, first for $\lambda_2 < 0$, then for $\lambda_2 > 0$. Thereafter we show the case $w_{EE}E > w_{EI}I$, first for $\lambda_2 < 0$, then for $\lambda_2 > 0$. Finally we show cases $w_{IE}E > w_{II}I$ and $w_{EE}E > w_{EI}I$ for $\lambda_2 = 0$.

Case $\boldsymbol{w_{IE}E > w_{II}I}$ when $\lambda_2 < 0$. From Equation (2), it is enough to show,

$$w_{IE}\left((1 + w_{II}) - e^{\lambda_2 t}(w_{EE} - 1) \right) > w_{II}w_{IE}\left(1 - e^{\lambda_2 t} \right)$$

$$\Leftrightarrow \left((1 + w_{II}) - e^{\lambda_2 t}(w_{EE} - 1) \right) > w_{II}\left(1 - e^{\lambda_2 t} \right)$$

$$\Leftrightarrow 1 - e^{\lambda_2 t}(w_{EE} - 1) > -w_{II}e^{\lambda_2 t}$$

But $\lambda_2 = -(2 - w_{EE} + w_{II})/\tau < 0$, so $(1 + w_{II}) > (w_{EE} - 1)$, so $-e^{\lambda_2 t}(w_{EE} - 1) > -e^{\lambda_2 t}(1 + w_{II})$, so

$$1 - e^{\lambda_2 t}(w_{EE} - 1) > 1 - e^{\lambda_2 t}(1 + w_{II}) = \left(1 - e^{\lambda_2 t}\right) - w_{II}e^{\lambda_2 t} > -w_{II}e^{\lambda_2 t} \text{ for } t > 0 \text{ as required.}$$

Case $\boldsymbol{w_{IE}E > w_{II}I}$ when $\lambda_2 > 0$. From Equation (2), it is enough to show,

$$w_{IE}\left((1 + w_{II}) - e^{\lambda_2 t}(w_{EE} - 1)\right) < w_{II}w_{IE}\left(1 - e^{\lambda_2 t}\right)$$

$$\Leftrightarrow \left((1 + w_{II}) - e^{\lambda_2 t}(w_{EE} - 1)\right) < w_{II}\left(1 - e^{\lambda_2 t}\right)$$

$$\Leftrightarrow 1 - e^{\lambda_2 t}(w_{EE} - 1) < -w_{II}e^{\lambda_2 t}$$

But $\lambda_2 = -(2 - w_{EE} + w_{II})/\tau > 0$, so $(1 + w_{II}) < (w_{EE} - 1)$, so $-e^{\lambda_2 t}(w_{EE} - 1) < -e^{\lambda_2 t}(1 + w_{II})$, so

$$1 - e^{\lambda_2 t}(w_{EE} - 1) < 1 - e^{\lambda_2 t}(1 + w_{II}) = \left(1 - e^{\lambda_2 t}\right) - w_{II}e^{\lambda_2 t} < -w_{II}e^{\lambda_2 t} \text{ for } t > 0 \text{ as required.}$$

Case $\boldsymbol{w_{EE}E > w_{EI}I}$ when $\lambda_2 < 0$. From Eqn. (2), it is enough to show,

$$w_{EE}\left((1 + w_{II}) - e^{\lambda_2 t}(w_{EE} - 1)\right) > w_{IE}w_{EI}\left(1 - e^{\lambda_2 t}\right)$$

$$\Leftrightarrow w_{EE}\left((1 + w_{II}) - e^{\lambda_2 t}(w_{EE} - 1)\right) > (1 + w_{II})(w_{EE} - 1)\left(1 - e^{\lambda_2 t}\right) \text{ since } c = (1 + w_{II})/w_{IE} = w_{EI}/(w_{EE} - 1)$$

$$\Leftrightarrow -w_{EE}e^{\lambda_2 t}(w_{EE} - 1) > -(1 + w_{II}) - e^{\lambda_2 t}(1 + w_{II})(w_{EE} - 1)$$

$$\Leftrightarrow (1 + w_{II}) > e^{\lambda_2 t}(w_{EE}(w_{EE} - 1) - (1 + w_{II})(w_{EE} - 1))$$

$$\Leftrightarrow (1 + w_{II}) > e^{\lambda_2 t}(w_{EE} - 1)(w_{EE} - 1 - w_{II})$$

But $\lambda_2 = -(2 - w_{EE} + w_{II})/\tau < 0$, so $e^{\lambda_2 t} < 1$ for $t > 0, w_{EE} - 1 < 1 + w_{II}, w_{EE} - 1 - w_{II} < 1$. Thus,

$$e^{\lambda_2 t}(w_{EE} - 1)(w_{EE} - 1 - w_{II}) < (1 + w_{II}) \text{ for } t > 0 \text{ as required.}$$

Case $\boldsymbol{w_{EE}E > w_{EI}I}$ when $\lambda_2 > 0$. From Eqn. (2), it is enough to show,

$$w_{EE}\left((1 + w_{II}) - e^{\lambda_2 t}(w_{EE} - 1)\right) < w_{IE}w_{EI}\left(1 - e^{\lambda_2 t}\right)$$

$$\Leftrightarrow w_{EE}\left((1 + w_{II}) - e^{\lambda_2 t}(w_{EE} - 1)\right) < (1 + w_{II})(w_{EE} - 1)\left(1 - e^{\lambda_2 t}\right) \text{ since } c = (1 + w_{II})/w_{IE} = w_{EI}/(w_{EE} - 1)$$

$$\Leftrightarrow -w_{EE}e^{\lambda_2 t}(w_{EE} - 1) < -(1 + w_{II}) - e^{\lambda_2 t}(1 + w_{II})(w_{EE} - 1)$$

$$\Leftrightarrow (1 + w_{II}) < e^{\lambda_2 t}(w_{EE}(w_{EE} - 1) - (1 + w_{II})(w_{EE} - 1))$$

$$\Leftrightarrow (1 + w_{II}) < e^{\lambda_2 t}(w_{EE} - 1)(w_{EE} - 1 - w_{II})$$

But $\lambda_2 = -(2 - w_{EE} + w_{II})/\tau > 0$, so $e^{\lambda_2 t} > 1$ for $t > 0, w_{EE} - 1 > 1 + w_{II}, w_{EE} - 1 - w_{II} > 1$. Thus,

$$e^{\lambda_2 t}(w_{EE} - 1)(w_{EE} - 1 - w_{II}) > (1 + w_{II}) \text{ for } t > 0 \text{ as required.}$$

Now we consider the cases when $\lambda_2 = 0$. Then in some open set around the steady state $E_s > 0, I_s > 0$, the dynamics corresponding to Equation (1) are given by

$$\begin{bmatrix} E(t) \\ I(t) \end{bmatrix} = a_1 \begin{bmatrix} c \\ 1 \end{bmatrix} + a_2 \left(t \begin{bmatrix} c \\ 1 \end{bmatrix} + \begin{bmatrix} \rho \\ 1 \end{bmatrix} \right)$$

where a_1, a_2 depend on the initial state of the system and ρ is such that $A\begin{bmatrix}\rho\\1\end{bmatrix} = \begin{bmatrix}c\\1\end{bmatrix}$. This gives $\rho = (\tau/w_{IE}) + c$.

Initial condition $\{E_0, I_0\}$ gives $a_1 = -E_0 w_{IE}/\tau, a_2 = -a_1$ so that

$$\begin{bmatrix}E(t)\\I(t)\end{bmatrix} = \frac{E_0 w_{IE}}{\tau}\left(\begin{bmatrix}c\\1\end{bmatrix}(t-1) + \begin{bmatrix}\rho\\1\end{bmatrix}\right) \text{ where } \rho = \frac{\tau}{w_{IE}} + c \tag{3}$$

$$\text{so that } E(t) = \frac{E_0 w_{IE}}{\tau}\left(ct + \frac{\tau}{w_{IE}}\right), I(t) = \frac{E_0 w_{IE} t}{\tau}.$$

Case $\boldsymbol{w_{IE}E > w_{II}I}$ when $\lambda_2 = \mathbf{0}$. From Equation (3), it is enough to show

$$w_{IE}\left(ct + \frac{\tau}{w_{IE}}\right) > w_{II}t \Leftrightarrow (w_{IE}c - w_{II})t + \tau > 0.$$

But $c = (1 + w_{II})/w_{IE}$, so $(w_{IE}c - w_{II}) = (1 + w_{II} - w_{II}) = 1$. Thus for $t > 0$, $w_{IE}E > w_{II}I$.

Case $\boldsymbol{w_{EE}E > w_{EI}I}$ when $\lambda_2 = 0$. From Equation (3), it is enough to show

$$w_{EE}\left(ct + \frac{\tau}{w_{IE}}\right) > w_{EI}t \Leftrightarrow (w_{EE}c - w_{EI})t + \frac{\tau w_{EE}}{w_{IE}} > 0$$

But $c = w_{EI}/(w_{EE} - 1)$, so $(w_{EE}c - w_{EI}) = cw_{EE} - c(w_{EE} - 1) = c > 0$. Therefore, for $t > 0$, $w_{EE}E > w_{EI}I$.

Thus we have shown that $\lambda_1 = 0$ for the system to have a non-zero steady state, and in this case, for any real λ_2, a trajectory starting from $\{E_0, I_0\}$ does not encounter the "rectification" since for $t > 0$, $w_{IE}E - w_{II}I > 0$ and $w_{EE}E - w_{EI}I > 0$. Since we want the trajectory to evolve to the steady state instead of diverging, we now only consider the case $\lambda_1 = 0, \lambda_2 = -(2 - w_{EE} + w_{II})/\tau < 0$. Note that any point on the stability boundary would correspond to $\lambda_1 = 0, \lambda_2 < 0$.

Let P be a point on the stability boundary corresponding to parameters w_{EE}^P, w_{EI}^P. Then at P, from Equation (2), the trajectory starting from $\{E_0, I_0\}$ converges to $\{E_s, I_s\} = \{E_0(1 + w_{II})/(2 - w_{EE}^P + w_{II}), E_0 w_{IE}/(2 - w_{EE}^P + w_{II})\}$. Note that $\lambda_2 = -(2 - w_{EE} + w_{II})/\tau < 0$, so $(2 - w_{EE} + w_{II}) > 0$ along the stability boundary. Let E_P denote the steady state excitatory response corresponding to P, given initial condition $\{E_0, I_0\}$. Thus $E_P = E_0(1 + w_{II})/(2 - w_{EE}^P + w_{II})$.

Say Q is another point on the stability boundary with parameters w_{EE}^Q, w_{EI}^Q such that $w_{EE}^Q > w_{EE}^P$. Note that $c = (1 + w_{II})/w_{IE} = w_{EI}/(w_{EE} - 1)$ is constant along the stability boundary and $w_{EE} > 1, w_{EI} > 1$ since the system is in the inhibition-stabilized regime. Thus, $w_{EE}^Q > w_{EE}^P$ implies $w_{EI}^Q > w_{EI}^P$. Let E_Q denote the steady state excitatory response at Q. Then $E_Q = E_0(1 + w_{II})/(2 - w_{EE}^Q + w_{II}) > E_P$ since $w_{EE}^Q > w_{EE}^P$.

Thus, given w_{IE}, w_{II}, moving along the stability boundary in the direction of increasing recurrent excitation w_{EE} and lateral inhibition w_{EI} (Figure 1C) shows a progressive increase in the steady state excitatory firing rate. This can be used to intuitively understand the higher transient amplification prior to decay (Figure 1B) achievable with stronger recurrent excitation balanced by stronger lateral inhibition.

The laminar model

The model examined in Figure 4 is based on a recently published large-scale rate model of the macaque cortex (Mejias et al., 2016), which incorporates a cortical laminar structure. The intra-laminar cortical circuit in each area consists of a recurrently connected excitatory and inhibitory population, with dynamics described by the following Wilson-Cowan equations,

$$\tau_E \frac{dr_E}{dt} = -r_E + \phi\left(I_E^{net} + I_E^{ext}\right) + \sqrt{\tau_E}\xi_E(t)$$

$$\tau_I \frac{dr_I}{dt} = -r_I + \phi\left(I_I^{net} + I_I^{ext}\right) + \sqrt{\tau_I}\xi_I(t)$$

where $r_{E,I}$ denote the dimensionless mean firing rates of the excitatory and inhibitory populations respectively, $\tau_{E,I}$ denote the corresponding time constants, $\xi_{E,I}$ denote Gaussian white noise terms with strengths $\sigma_{E,I}$, and $\phi(x) = x/(1 - \exp(-x))$ is the transduction function. The network input, denoted by $I_{E,I}^{net}$, is the input arriving to the E, I populations respectively from the other populations in the network, and includes the inputs from the same layer, a different layer, and from different areas. The external input, denoted by $I_{E,I}^{ext}$, is the input arriving from external sources such as sensory stimuli, thalamic input and other cortical areas not explicitly included in the model. The network input taking into account only local contributions, that is, on assuming an isolated intra-laminar population, is given by,

$$I_E^{net} = J_{EE} r_E + J_{EI} r_I$$

$$I_I^{net} = J_{IE} r_E + J_{II} r_I$$

where $J_{\alpha\beta}$ denotes the mean synaptic strength from population β to population α. The parameter values for the circuit in the superficial layer are $\tau_E = 6$ ms, $\tau_I = 15$ ms, $J_{EE} = 1.5, J_{IE} = 3.5, J_{EI} = -3.25, J_{II} = -2.5$ and $\sigma_{E,I} = 0.3$. The parameters for the circuit in the deep layer are the same except for $\tau_E = 30$ ms, $\tau_I = 75$ ms, and $\sigma_{E,I} = 0.45$.

The inter-laminar interactions assume only the strongest connections between the superficial and deep layer, that is, the excitatory projections from the pyramidal neurons of the superficial layer to the pyramidal neurons of the deep layer, and those from the pyramidal neurons of the deep layer to the interneurons of the superficial layer. The long-range connections are assumed to be excitatory. The inter-areal interactions assume that feedforward projections originate from the superficial layer and target the excitatory population in the superficial layer across areas. Feedback projections, are assumed to originate from the deep layer and target both the E and I populations in the superficial and deep layer across areas. The feedforward connections from the excitatory neurons of the superficial layer to the inhibitory ones (Figure 4A) were added to test our mechanism and are not present in the original model (Mejias et al., 2016). These connections are assumed to have 20% of the strength of the feedforward connections targeting the excitatory populations across areas. For consistency with the other models examined in the present work, we simulate the large-scale laminar model considering 29 areas as opposed to the 30 areas (Mejias et al., 2016) (i.e., we remove area LIP).

The spiking network model

We build a spiking network model examined in Figures 5, 6, 7. Simulations are performed using a network of leaky integrate-and-fire neurons, with the local-circuit and long-range connectivity structure similar to the rate model. Each of the 29 areas consists of 2000 neurons, with 1600 excitatory and 400 inhibitory neurons. Connection density, both intra and inter-areal is set at 10%. The membrane constant values are $\tau_E = 20$ ms for excitatory, and $\tau_I = 10$ ms for inhibitory neurons. The resting membrane potential V_r, reset potential V_{reset}, and threshold potential V_t are given by $V_r = -70$ mV, $V_{reset} = -60$ mV and $V_t = -50$ mV respectively, and the absolute refractory period $\tau_{ref} = 2$ ms. Background currents are injected to yield firing rates in the 0.75-1.5 Hz range for the excitatory, and 5-6 Hz for the inhibitory population in the absence of input. We introduce distance-dependent inter-areal synaptic delays by assuming a conduction velocity of 3.5 m/sec (Swadlow, 1990; Mejias et al., 2016) and using a distance matrix based on experimentally measured wiring distances across areas (Markov et al., 2014a). Inter-areal delays are assumed to have a Gaussian distribution with mean based on the inter-areal wiring distance and variance given by 10% of the mean. Intra-areal delays are set to 2 ms. For neuron i, the depolarization voltage at the soma follows

$$\tau_i \dot{V_i}(t) = -(V_i(t) - V_r) + RI_{syn,i}(t) + RI_{ext,i}(t)$$

where $I_{syn,i}(t)$ is the post-synaptic current and $I_{ext,i}(t)$ is the external input. The post-synaptic current corresponds to a summation of spike-contributions of spikes arriving at different synapses at different time intervals, where the spikes are modeled as delta functions (Brunel, 2000a). Thus,

$$RI_{syn,i}(t) = \tau_i \sum_j J_{ij} \sum_k \delta\left(t - t_j^k - D\right)$$

where τ_i is the membrane constant, D is the transmission delay, t_j^k is the emission time of the k^{th} spike at the j^{th} synapse, and J_{ij} is the synaptic strength of the j^{th} synapse to neuron i. We choose $R = 50M\Omega$; synaptic strengths for the local and global coupling parameters are chosen in the range 0.01 - 1 mV (Song et al., 2005; Brunel, 2000b). Simulations are performed using the Python library Brian2 (Goodman and Brette, 2009), using a time-step of 0.1ms. Population firing rate in each case is calculated using sliding time window with a bin size of 10 ms and the sliding window width of 1 ms. For the spiking model simulations, the parameter η, governing the gradient of excitatory strengths, is set to 4.

For the asynchronous regime (Figure 5), the synaptic strengths are set to (in mV) $w_{EE}, w_{IE}, w_{II}, \mu_{IE} = 0.01, 0.075, 0.075, 0.19/4$. The global excitatory coupling and local I to E strength are set to (in mV) $\mu_{EE}, w_{EI} = 0.0375, 0.0375$ for weak GBA, and (in mV) $\mu_{EE}, w_{EI} = 0.05, 0.05$ for strong GBA (Figure 5). The input current duration is set to 150 ms, based on recent work suggesting that activity packets of 50-200 ms duration serve as basic building blocks of global cortical communication (Luczak et al., 2015). The input current strength is set to 300 pA in Figure 5B and 126 pA in Figure 5C, resulting in a peak firing rate of 82-87 Hz in V1. For weak and strong GBA in Figure S4, connectivity parameters were used as in Figures 5B and 5C. A pulse input of 150 ms duration is injected to the excitatory population of area 2. The input current strength is set to 138 pA for weak GBA and 140 pA for strong GBA, each resulting in a peak firing rate of ~40-41 Hz in area 2.

For the synchronous regime in the spiking model, we set the synaptic strengths $w_{EE}, w_{IE}, w_{II}, \mu_{IE}$ to be a multiple of the values used in the asynchronous case, the values are (in mV) $w_{EE}, w_{IE}, w_{II}, \mu_{IE} = 0.04, 0.3, 0.3, 0.19$. The global excitatory coupling and local I to E strength are set to (in mV) $\mu_{EE}, w_{EI} = 0.16, 0.56$ for weak GBA, and $\mu_{EE}, w_{EI} = 0.25, 0.98$ for strong GBA (Figures 6A and 6B). The input current duration is set to 8 ms, and the input current strength is set to 200 pA, resulting in a peak firing rate of ~24 Hz in V1. For

Figure 5D, we map the attenuation in Figure 5E logarithmically to a heatmap. For better visualization, we threshold the areas showing strong propagation and zoom in on the other areas, and plot the values using Caret (Van Essen et al., 2001). For Figure 6C, we map the attenuation in Figure 6D logarithmically to a heatmap and plot the values using Caret (Van Essen et al., 2001).

QUANTIFICATION AND STATISTICAL ANALYSIS

Non-normality measure

Since inhibition-stabilized networks showing balanced amplification are known to be characterized by strong non-normality of the underlying connectivity matrix (Murphy and Miller, 2009), we examine the non-normality of the large-scale connectivity matrix on increasing GBA (Figure S2B) corresponding to the parameters used in Figure 3F. Non-normality can be examined through Schur decomposition (Murphy and Miller, 2009; Henrici, 1962), which can be used to express the effective connectivity across network basis patterns through self-connections and feedforward connections. For a given matrix A, the Schur decomposition of A can be expressed as, $A = UTU^*$ where U is a unitary matrix, and T is an upper triangular matrix such that $T = \Sigma + R$, where Σ is a diagonal matrix and R is strictly upper triangular. The departure of A from normality can be approximated (Henrici, 1962) as

$$\mathrm{dep}_F(A) = \sqrt{\|A\|_F^2 - \|\Sigma\|_F^2} = \left(\sum_{j=1}^{N} \sigma_j^2 - \sum_{j=1}^{N} |\lambda_j|^2\right)^{1/2}$$

where σ_j, λ_j denotes the singular value and the eigenvalue of A respectively, for a given j and $\| \|_F$ denotes the Frobenius norm. For normal matrices, $\sigma_j = |\lambda_j|$ for each j, thus $\mathrm{dep}_F(A) = 0$. An improvement in propagation is accompanied by an increase in the non-normality measure of the underlying connectivity matrix governing the large-scale dynamics (Figure S2B).

Incorporating structural asymmetries

Recent work (D'Souza et al., 2016) observes asymmetries in the recruitment of inhibition and excitation by feedforward and feedback connections, noting a gradual scaling down of the inhibition/excitation ratio from the most feedforward to the most feedback pathway. Figure S3F examines the robustness of our mechanism on incorporating these structural asymmetries in our model. To do this, we first compute how feedforward or feedback a long-range connection is. Say, there is a long-range connection from area X to Y. Say, $\gamma_{YX} = \mathrm{hierarchy(receiving\ area\ Y)} - \mathrm{hierarchy(projecting\ area\ X)}$ is the feedforward degree of the X-to-Y connection. Then the long-range excitatory to excitatory connection strength from area X to Y is kept the same, while the long-range excitatory to inhibitory connection strength is scaled by $(1 + \mu \max(\gamma_{YX}, 0))$ where μ is a scaling parameter governing the strength of long-range excitatory to inhibitory connections based on hierarchy. Thus, if the hierarchical position of the receiving area is lower than that of the projecting area, then the connection strength is left unchanged. On varying μ and computing the propagation ratio, namely the ratio of the peak response in area 24c, which is highest in the hierarchy, to that in V1, we observe that the mechanism consistently improves propagation.

Predicted onset times

The response onset time for a given area in Figure 6B is defined as the time at which the response activity starts building up in that area prior to the peaking of activity. To compute the predicted onset times (Figure 6E), the FLNs are thresholded to 0.02, that is, connections with FLN values below 0.02 are removed, following which the predicted onset time is computed assuming that the signal follows the shortest possible path to reach a given area. For a positive integer k, a k-step path, given by $\{A = A_0, A_1, A_2, \ldots, A_k = B\}$ is said to exist between areas A and B if there exist anatomical connections from area A_m to area A_{m+1} from $m = 0, 1, \ldots, k-1$, that is, if $FLN_{A_{m+1},A_m} > 0$. We use Dijkstra's algorithm to compute the shortest path from V1 to a given area based on the inter-areal wiring distances (Markov et al., 2014a); the predicted onset time is computed assuming a constant conduction velocity of 3.5 m/s (Swadlow, 1990; Mejias et al., 2016). Let $S_{A,V1}$ be the length of the shortest path from $V1$ to a given area A, divided by the conduction velocity. We assume that the signal undergoes local-circuit processing at each step while traveling from V1 to area A, and assume a processing time of 1 ms in each area. (This is also assumed in Figure S6.) Assuming that the shortest path from V1 to area A is a k-step path, the predicted onset time (in ms) for the signal to reach area A from V1 is given by $S_{A,V1} + k$. While computing the predicted onset times, we eliminate those areas from our computations wherein the signal does not elicit response activity. In particular, as the background firing rate for the excitatory population lies in the range 0.75-1.5 Hz, we eliminate those areas with a peak firing rate <1.5 Hz. The only such area we find is F2 (in the premotor region). For F2, the predicted onset time for visualization purposes (Figure 6E) is set as the mean of the predicted onset times corresponding to its neighboring areas in the hierarchy.

From the asynchronous to synchronous regime

The degree of population synchrony χ, in the network is measured based on the variance of the average population voltage in comparison to the variance of the individual neuron voltage. We compute χ, which takes a value between 0 and 1, defined (Golomb et al., 1994) as

$$\chi = \sigma_V \Big/ \sqrt{\sum_{i=1}^{N} \sigma_{V_i}^2 \Big/ N}.$$

χ is computed (Figure S5) on moving from the parameter set in Figure 5B to the parameter set in Figure 6A. The parameter set comprises of the inter-areal and intra-areal connectivity parameters, the current duration and input current strength. The current duration corresponding to the parameter set in Figure 6A is set to 100 ms to characterize the fluctuations in the average population voltage. For each parameter set, the simulation is run using 5 random seeds, to get the mean value and standard deviation of χ.

Activity in cortical lobes on varying input

Active areas in Figure 7A are defined as those for which the peak firing rate in Figure 5C is at least 5% that of the corresponding peak rate in V1. The population activity of these areas is monitored for the input current strength varying from 70 to 120 pA. At each current value, we subtract the background firing rate for each of the active areas. To compare activation across areas, the normalized activity for each active area is computed. The normalized peak firing rate for a given current value is computed by dividing the peak firing rate at that value by the maximum peak firing rate over all the current values, so that the normalized rate lies between 0 and 1 as the current strength is varied. Active areas in Figure 7A are occipital areas V1, V2, V4, parietal areas DP, 7A, 7 m and 7B, temporal areas MT, TEO and TEpd and prefrontal areas 8 m, 8l, 46d, 10, 9/46d and 8B. In the absence of feedback in Figure 7B, normalized activity is examined only in those areas deemed ''active,'' for which the peak firing rate at current strength 120 pA is at least 1% that of the peak firing rate in V1. For Figure 7C, on scrambling the anatomical connectivity, the areas which show a strongly non-monotonic change in normalized activity in response to an increase in current strength, are regarded as not receiving the input signal, and are not plotted. For each input current value in Figure 7, the simulation is run using 5 random seeds; the curves plotted for each area correspond to the mean from the 5 simulations. For Figures 7A (upper) and B (inset), we map the difference between the normalized activity of an area and 1, for the given current value, to a heatmap and plot using Caret (Van Essen et al., 2001).

Figure S7 quantifies global ignition by computing the number of areas for which the activity exceeds 15% of peak activity, as a function of input current strength. Areas in the occipital and temporal lobes are activated for smaller values of input current strength. A sudden jump is observed beyond 95 pA, resulting from parietal lobe activation in area 7 (7A, 7 m, 7B), and frontal lobe activation in the frontal eye fields (8l, 8 m), the dorsolateral prefrontal cortex (area 46d, 9/46d) and the frontopolar cortex (area 10).