

# Chapter 5

## Synaptic Excitatory-Inhibitory Balance Underlying Efficient Neural Coding



Shanglin Zhou and Yuguo Yu

### 5.1 Introduction

Neural information coding is one of the central topics in neuroscience. The brain utilizes some features of action potential sequences (spike trains) to encode sensory and cognitive information. The algorithm operating within those features is called the neural code. Half a century ago, Perkel and Bullock (1968) noted that a potential neural code must serve at least four functions: stimulus representation, interpretation, transformation and transmission. Stimulus representation indicates that the neural activity should be altered by the stimulus properties needed to be coded, and therefore, the neural code can represent this stimulus (Perkel and Bullock 1968; Kumar et al. 2010). Due to its basic role, neural representation has been extensively studied using experimental and theoretical approaches. Barlow in 1961 proposed a theoretical framework which hypothesized that the action potentials in the sensory neurons formed a **neural code** for efficiently representing sensory information. By efficient Barlow meant that the code minimized the number of neurons and spikes needed to represent an input signal. This is the origin of sparse coding or efficient neural coding (Barlow 1961). Barlow's model treats the sensory pathway as a communication channel where neuronal spiking is an efficient code for representing sensory signals. The spiking code aims to maximize available channel capacity by minimizing the redundancy between representational units (Simoncelli and Olshausen 2001). In addition, one of the major components of a typical neural code also include reliable information transmission. The brain is highly modular, and a successful neural code should be able to be transmitted (propagated) from one module to another with high fidelity (Perkel and Bullock 1968; Kumar et al. 2010).

---

S. Zhou · Y. Yu (✉)

State Key Laboratory of Medical Neurobiology, School of Life Science and Human Phenome Institute, Institutes of Brain Science, Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, China  
e-mail: [yuyuguo@fudan.edu.cn](mailto:yuyuguo@fudan.edu.cn)

The transmission property of neural coding has also drawn significant attention recently (Diesmann et al. 1999; Kistler and Gerstner 2002; van Rossum et al. 2002; Litvak et al. 2003; Vogels and Abbott 2005; Kumar et al. 2008, 2010).

The balance of excitatory and inhibitory synaptic membrane currents (E/I balance) received by a neuron underlying its spontaneous firing and/or responding to sensory inputs has been widely observed (van Vreeswijk and Sompolinsky 1996; Brunel 2000; Shu et al. 2003; Wehr and Zador 2003; Froemke et al. 2007; Murphy and Miller 2009). Here, E/I balance generally refers to excitatory-inhibitory balance in terms of either overall global balance or temporal balance on a fine time scale. Global E/I balance refers to the bulk measurement of relative contributions of excitatory and inhibitory synaptic currents received by a specific neuron. It is called global E/I balance if across a range of spatio-temporal conditions of interest, the ratio between the synaptic excitation and inhibition is kept approximately constant. In some situations, even the measurement of firing rates of excitatory and inhibitory neurons or excitatory and inhibition synaptic conductances received by a neuron can represent E/I balance for individual neurons within the cortical network circuit. Temporal balance indicates that the relative magnitudes of excitatory and inhibitory synaptic currents are matched in a point-to-point manner on a fast time scale. Global E/I balance is often used to examine pathological or dysfunctional brain states, whereas temporal E/I balance can be used to examine the effect of synaptic correlation on spiking timing to sensory input and stimulus feature selectivity. Both global and temporal E/I balances enable cortical operation in a precise manner to represent sensory inputs. The disruption of the cortical E/I balance has been demonstrated to cause cognitive dysfunction, such as schizophrenia (Yizhar et al. 2011; Murray et al. 2014). Because the E/I balance may be the key structure underlying the neural code and cognition, multiple questions arise: (1) How is the E/I balance achieved? (2) Why does the neural system choose such a scenario to function? (3) How does the E/I balance evolve during neural plasticity and coding? Specifically, how does the E/I balance influence information representation and propagation across regions?

Recently, more and more studies are conducted to answer these questions. Here, we briefly summarize the evidence for the existence of an E/I balance in the cortex and the mechanisms by which the E/I balance is achieved. We then review the experimental and computational development on the impact of the E/I balance on neural coding, especially the processes of stimulus representation and information propagation.

## 5.2 E/I Balance Is Ubiquitous in Cortical Circuits

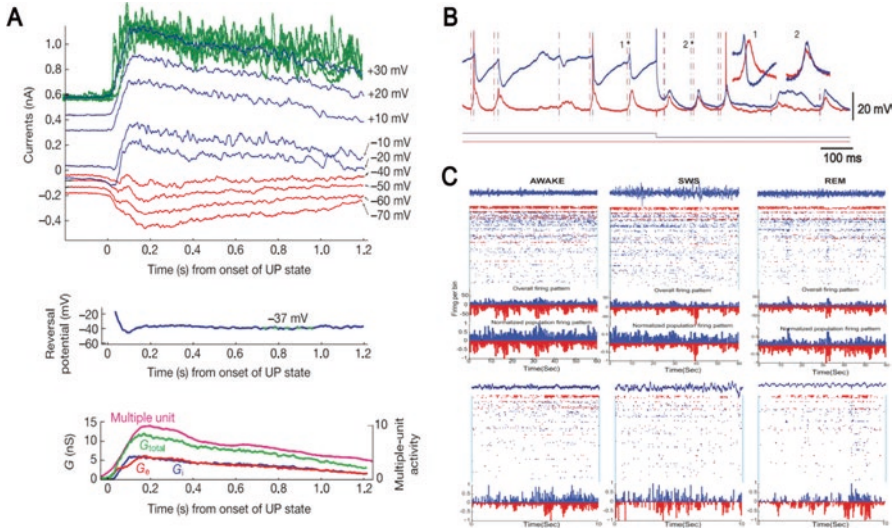
Over the last decades, E/I balance has been found to exist in many situations including ongoing spontaneous activity, sensory-evoked activity and storage of memories. Synaptic plasticity at both excitatory and inhibitory synapses is suggested to play a central role in balancing the excitatory and inhibitory inputs to a targeted cell during the training or learning process (Vogels et al. 2011; Yu et al. 2014). The level of the

developed balance depends on the time scale of the correlation between the excitatory and inhibitory inputs to the cell, ranging from a global balance, either without a correlation or with a correlation at a slower time scale, to a fine-scale balance for strong correlations with a fast time scale.

Global balance is quantified by using global measures of excitatory and inhibitory synaptic currents, including measuring spontaneous or ongoing excitatory and inhibitory postsynaptic currents (mEPSC and mIPSC) and the field potential, which is considered a rough signature of the relative timing and magnitude of excitation and inhibition. In fact, it is difficult to simultaneously measure excitatory and inhibitory current inputs on the same neuron. However, researchers can overcome this by measuring the excitatory and inhibitory currents separately at different holding potentials and then calculating the average conductance of both currents ( $g_E$  and  $g_I$ ) (Shu et al. 2003; Haider et al. 2006; Monier et al. 2008). By using this method, Shu et al. (2003) found that the received synaptic conductance values of  $g_E$  and  $g_I$  were always balanced with a certain ratio during the up state generated by recurrent connection patterns in the *in vitro* brain slice (Fig. 5.1a). Other experimental results also support the idea that the ratio of  $g_E$  and  $g_I$  of a given neuron remains constant across different conditions and in many systems (Wehr and Zador 2003; Haider et al. 2006; Xue et al. 2014). Additionally, many studies have demonstrated that the E/I balance still exists even when the system is driven by external inputs (Anderson et al. 2000; Martinez et al. 2002; Tan et al. 2004, 2011; Wilent and Contreras 2005; Cardin et al. 2007; Wu et al. 2008; Tan and Wehr 2009; Runyan et al. 2010; Liu et al. 2011). In fact, some studies have shown that the tuning curves of the excitatory and inhibitory conductance are similar to one another (Anderson et al. 2000; Wehr and Zador 2003; Cardin et al. 2007; Runyan et al. 2010).

To understand the E/I balance on the fine time scale (the fine-scale balance), researchers have tried to simultaneously record the time series of both the excitatory and inhibitory currents and then obtain the correlation between them. Since adjacent neurons in the cortex generally receive strongly correlated synaptic inputs, researchers can record both excitatory and inhibitory currents separately and simultaneously, each in a single neuron in a pair of neighboring cells, and the correlation between the excitatory and inhibitory currents onto a single cell can be inferred from the correlation between the time series from the two cells (Okun and Lampl 2008). Based on this method, researchers have found that the excitatory and inhibitory inputs from ongoing spontaneous activity or sensory-evoked activity are strongly correlated with one another, with inhibitory currents tracking excitatory currents closely with a few milliseconds of a delay (Fig. 5.1b) (Okun and Lampl 2008). More evidence has also shown that a fine-scale E/I balance exists during oscillations in the gamma and beta frequencies (Atallah and Scanziani 2009; Poo and Isaacson 2009).

Recently, in an interesting study using *in vivo* recordings with dense multielectrodes in the neocortex of higher level mammals (including human and primate), Dehghani et al. (2016) found that excitatory and inhibitory ensembles are well-balanced and co-fluctuate instantaneously in all states of the wake-sleep cycle (wake, slow-wave sleep and rapid-eye movement sleep) at different temporal scales



**Fig. 5.1** Experimental evidence of the E/I balance. (a) Average currents during the up state in recordings clamped at different membrane potentials from *in vitro* brain slices (top, red and blue curves showing the average currents, the green curves showing the raw traces at +30 mV), the reversal potential of the average synaptic currents (middle), and additional conductances during the up state (bottom). Adapted from Shu et al. (2003). (b) Simultaneous *in vivo* recordings from two cortical cells. One cell (red) was continuously recorded in a hyperpolarized mode, and the other cell (blue) was switched between depolarized and hyperpolarized modes (current depicted below the traces). Dashed lines mark the onset of synaptic events. Insets show examples of two events (marked by asterisks). Adapted from Okun and Lampl (2008). (c) Recordings in humans during awake (left), slow-wave sleep (SWS) (middle) and rapid-eye movement (REM) (right) states. Top row shows 60-s windows; bottom row shows a 10-s window of the same state. Putative inhibitory neurons (FS cells) are shown in red. Putative excitatory neurons (RS) are depicted in blue. At the top of each panel, a sample LFP trace (in blue) accompanies the spiking activity. Histograms show the overall activity of the RS (blue) and FS (red) cells. Adapted from Dehghani et al. (2016)

(Fig. 5.1c). Beyond the temporal view of E/I balance, the spatial properties of E/I balance are also important in information processing. For instance, researchers observed that local E/I imbalance coexisting with overall balance facilitates neural network creating novel features selectivity (Wu et al., 2008). However, we will mainly focus on temporal E/I balance in the following discussion.

### 5.3 Mechanisms to Achieve E/I Balance

To achieve global E/I balance in a dynamic neural network, several theoretical studies have shown that a neural network needs to be equipped with the following properties: (1) neurons in the network must be connected sparsely (the number of

connections per neuron should be much smaller than the total number of neurons in the network) and randomly; (2) the strength of the inhibitory connections must be higher than the excitatory connections (van Vreeswijk and Sompolinsky 1996, 1998; Brunel 2000). Under such conditions, the average of the excitatory and inhibitory synaptic currents could be well-balanced, and the network dynamics could be stable. In such a balanced network, the membrane potentials and spike trains of the individual neurons may be highly uncorrelated (van Vreeswijk and Sompolinsky 1996, 1998; Brunel 2000). Beyond the two aspects mentioned above, synaptic plasticity may also play a vital role in the formation of the E/I balance (Froemke 2015).

In contrast to requirements for global balance of a network, Renart et al. (2010) proposed a neural network with random and dense connections (with the number of connections per neuron comparable to the total number of neurons) to achieve a more fine-scale E/I balance by setting the synaptic conductances and connection structure in the optimal range. In such a network, the excitatory and inhibitory currents received by each neuron are strongly correlated on a fast time scale. If excitatory and inhibitory currents cancel each other, then the net input current will be highly random, resulting in highly variable neural responses. Boerlin et al. (2013) have demonstrated that both variable neural responses and balanced excitation/inhibition are necessary consequences of neural networks that represent information efficiently in their spike trains. However, the Boerlin model assumes instantaneous synapses (transmission without delays) and only achieves balance because of this assumption. Further work allowed to relax this assumption and introduce realistic synapses (Koren and Deneve 2017). With realistic synapses it is however required that parameters that weight the cost on spiking are fine-tuned. Those parameters can be interpreted in biological terms as determining the excitability of the network.

Furthermore, a theoretical study proved that a network with synaptic plasticity of inhibitory synapses could evolve into a fine-scale E/I-balanced state with sparse connections (Vogels et al. 2011), and a later experimental study demonstrated the existence of this form of synaptic plasticity (D'amour and Froemke 2015). Further information about inhibitory synaptic plasticity could be found in other works (Vogels et al., 2011).

Beyond the theoretical work, experimental studies have provided some additional insights into the development of E/I balance (Froemke 2015). For example, Liu (2004) found that the ratio of the number of the excitatory and inhibitory synapses on the dendrites of cultured hippocampal neurons remained constant along different developmental stages, which suggests that the E/I balance may be related to an anatomical basis. In addition, sensory experiences at different developmental stages may play important roles in shaping the final E/I balance level (Froemke 2015).

## 5.4 E/I Balance and Information Representation

One of the fundamental functions of the neural systems is to represent the sensory information and make use of it for guiding action, a process termed neural coding. Representing neural signals is the process of interpreting prominent features of external sensory inputs with individual or population neuronal activity. Experimental and theoretical studies have demonstrated that action potential generation is an energy-expensive process (Attwell and Laughlin 2001; Alle et al. 2009; Yu et al. 2012). Therefore, efficient coding can be defined as coding with as few neurons and action potentials as possible, while not losing fidelity in representation of certain stimulus features (equivalent to sparse coding as defined by Barlow, 1961) or minimized coding error during stimulus representation (Deneve and Machens, 2016). Although there is no strict theoretical proof, the minimal coding error and information maximization or redundancy reduction are related in some aspects. Intuitively, coding error reduction means a decrease in noise information, which would increase the mutual information between the neural response and the input signal, thus increasing the information coding efficiency. Because the E/I balance is ubiquitous in neural systems, there must be some strategic benefits of the E/I balance for efficient representation. Here, we summarize the evidence for this as follows.

### 5.4.1 Irregular Spike Trains and Global E/I Balance

A typical well-known property of the firing pattern of an individual neuron recorded *in vivo* is its irregularity or stochasticity, which is similar to Poisson-like time sequences. Revealing how individual neurons establish such irregular firing patterns is important for understanding the network states with spontaneous firing and how such states could be used to represent stimulus inputs. In fact, it has been widely shown that irregular firing patterns could be achieved by a neuron with balanced excitatory and inhibitory synaptic inputs on multiple time scales (Shadlen and Newsome 1994, 1998; van Vreeswijk and Sompolinsky 1996; Amit and Brunel 1997; Brunel 2000). There is an intuitive explanation to why such a globally balanced network would lead to the irregular firing of a single neuron. Imagine that there is a neural network where each neuron is bombarded with noisy, Poisson-distributed synaptic inputs from both excitatory and inhibitory sources. When the excitatory input values exceed the inhibitory inputs, then the net mean positive input would depolarize the neuron to fire quasi-regularly. However, if the excitatory inputs and inhibitory inputs cancel each other out in a slow time scale without correlation on a fast time scale, then the membrane potential of each neuron would randomly cross the threshold dependent on the fast noise, resulting in a spiking pattern with a high level of irregularity (Denève and Machens 2016).

Although such a network architecture would capture the irregularity of the neural firing pattern, the network behavior would become very sensitive to even a small

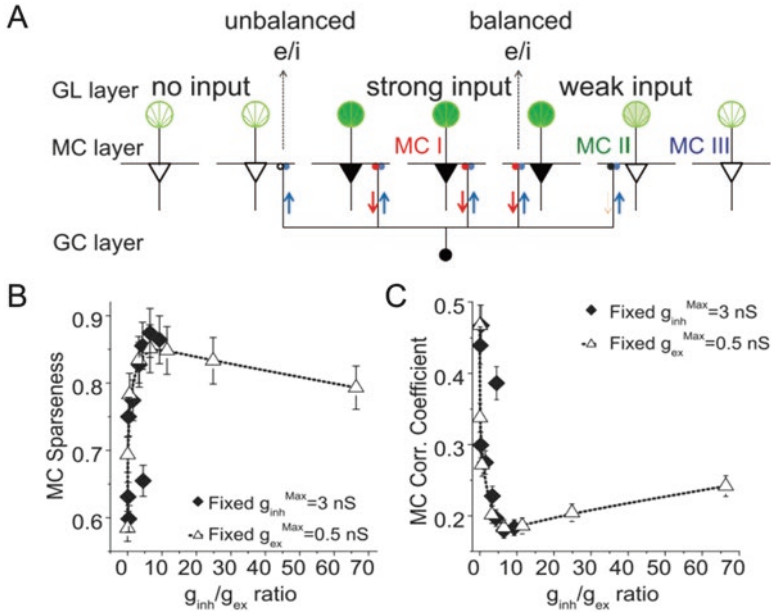
perturbation due to its chaotic dynamics (Shadlen and Newsome 1998; Brunel 2000). This hypothesis suggests a low reliability of the neural network in response to sensory input, which makes such an E/I-balanced network represent stimulus features in a poor fidelity.

### 5.4.2 Sparse Coding, E/I Balance and Energy Efficiency

Sparse coding implies that only a small fraction of cells in the network displays a transient response to an input signal (Vinje and Gallant 2000) and such an energy-efficient paradigm greatly extends the coding capacity of a large family of sensory inputs (Olshausen and Field 1996; Dhawale et al. 2010; Wolfe et al. 2010; Koutrakov and Rinberg 2011). Yu et al. (2013, 2014) implemented a large-scale olfactory bulb model with mitral cell and granule cell connected by dendro-dendritic synapses with regular LTP/LTD synaptic plasticity, and they found that balanced excitation/inhibition in strongly activated mitral cells leads to a sparse representation of odorant inputs (Fig. 5.2a). They further found that such a network with synaptic plasticity could always evolve into a sparsely oscillatory state to represent the input signal efficiently. During the evolving process, global synaptic excitation and inhibition gradually reach an optimal balance with which the network produces firing patterns with the highest level of sparseness (Fig. 5.2b) (Yu et al. 2014). The optimal level of synaptic excitation and inhibition could produce the highest level of sparseness and decorrelation in the network response and reduce energy cost (Nawroth et al. 2007).

Interestingly, the formation of response sparseness in such an olfactory bulb network does not depend on a specific type of synaptic plasticity, meaning either Hebbian or non-Hebbian rules can both develop the network dynamics into sparseness during the training process (Migliore et al. 2010; Yu et al. 2013, 2014). In a recent work by Vogels et al. (2011), the sparsely connected network endowed with plasticity of inhibitory synapses could evolve to a sparse response to natural stimuli. In addition, this type of network can accommodate synaptic memories with activity similar to the background activity; and same activity can be reactivated by external stimuli (Vogels et al. 2011).

Note that the above approaches assumed sparse network connection, i.e., neurons receive few connections  $K$  compared to the size of the network  $N$ , so that  $K \ll N$ . Such a sparse connectivity usually leads to uncorrelated excitation and inhibition, resulting in random fluctuations as input to neurons within network. Indeed, achieving efficient representation of input signal does not have to set the network to be sparsely connected. Recent theoretical works investigated network with dense connections. In such a scenario excitation and inhibition received in a neuron are strongly correlated while neuronal output spike trains are highly uncorrelated. Such a balanced network could represent information efficiently in their spikes (Boerlin et al. 2013). To ensure the network performance, two assumptions are required: (1) information on dynamic variables can be read out linearly from spike trains, and (2) neurons fire a spike only if it improves the representation of



**Fig. 5.2** The correlation of firing sparseness and mitral cell spiking in a large-scale olfactory bulb model. **(a)** Schematic representation of balanced and unbalanced excitation and inhibition in the MC-GC circuit. Three activated middle MCs (solid black triangles) receive strong input from glomeruli (solid deep green color); through back-propagation of APs in their lateral dendrites, they distribute the excitation (red) through reciprocal synapses, activating lateral inhibition in the surrounding MCs through the reciprocal inhibitory synapses. This mode of excitation and inhibition is balanced, and these MCs are called MC type I. The activated GCs (small blue spheres) deliver lateral inhibition to other surrounding MCs with weak or no excitatory inputs, making their reciprocal synapses unbalanced. These MCs are called MC type II. MCs that do not receive lateral inhibition are MC type III. **(b)** The MC network sparseness level as a function of reciprocal inhibitory weight to excitation weight ratio  $g_{inh}/g_{ex}$  for the cases of different  $g_{inh}^{Max}$  with a fixed  $g_{ex}^{Max}=0.5$  nS and different  $g_{ex}^{Max}$  with a fixed  $g_{inh}^{Max}=0.3$  nS. **(c)** Same in **(b)** but shows the mitral cell spiking correlation. Adapted from Yu et al. (2014)

dynamical variables. When a network satisfies these conditions, it evolves naturally to achieve the objective representation of a time-varying input with a minimum number of spikes with maximal efficiency (Boerlin et al. 2013; Denève and Machens 2016). In the following work, they further revealed that the maximally efficient network is right at the transition from synchronous to asynchronous network states (Koren and Deneve 2017). Moreover, there is a tight relationship between coding efficiency and energy efficiency, it was observed that there is a “sweet spot” where maximal coding efficiency coincides with rather low number of spikes (Boerlin et al. 2013; Koren and Deneve 2017).

Using single-compartment computational models with stochastic voltage-gated ion channels, Sengupta et al. (2013) calculated information content under either E/I-balanced or unbalanced conditions. They found that balanced synaptic currents



evoke fewer spikes per second than the unbalanced conditions but with more information content in a single spike (bits/spike) in the balanced conditions. The total informative rate is similar in the two conditions (Sengupta et al. 2013). These results strongly support the hypothesis that E/I balance can promote both coding efficiency and energy efficiency.

Indeed, maximizing the ratio of the coding capacity to energy cost has been suggested to be one of the key principles chosen by the nervous system to evolve under selective pressure, and the metabolic energy efficiency demands of the nervous system could be sufficiently large to influence the design, function and evolution of the brain (Niven and Laughlin 2008). A recent theoretical work revealed a general rule for population coding in which the neuronal number should be sufficiently large to ensure reliable information transmission that is robust to the noisy environment but small enough to minimize energy cost (Yu et al. 2016). Experiments in cortex cultures, anesthetized rats, and awake monkeys, as well as computer models, have shown that balanced excitation/inhibition (E/I) could lead to a critical dynamic of avalanches in the cortical neural network (Shew et al. 2009; Poil et al. 2012; Yang et al. 2012). The number of metastable states (Haldeman and Beggs 2005) and the dynamic range to the input stimuli (Shew et al. 2009), as well as the information capacity and transmission (Beggs 2008; Shew et al. 2011) of the cortical neural networks, could be maximized at the critical point. The developed E/I balance within brain circuits during rest, learning and memory states may be beneficial for the brain to maintain an optimal state based on the theory of criticality. A large E/I ratio leads to a super-critical state whereby the neurons are highly activated and spikes among neurons are highly correlated. However, a small E/I ratio leads to a sub-critical state whereby the overall neural activity level drops and the spikes among neurons are random and not correlated (Yang et al. 2012). For information processing, highly correlated spikes reduce entropy in the former case, and in the latter case, the reduced correlation increases entropy, but this increase is counteracted by the concurrent drop in total information, resulting in maximal information transmission at a moderate E/I ratio (Shew et al. 2011). However, energy expenditure increases monotonically as the E/I ratio increases due to the increasing overall neural activity level. Therefore, a relatively large information transmission while relatively low energy cost, is expected to be maximized around an optimal E/I ratio (Poo and Isaacson 2009; Yu et al. 2014; Denève and Machens 2016).

### 5.4.3 *Decorrelation and E/I Balance*

The correlations among the spiking trains of all individual neurons in a network in response to sensory input can either help or harm the information transfer (Averbeck et al. 2006). More specifically, if positive signal correlations (i.e., neurons with similar selectivities of stimulus features) are linked to positive noise correlation, this would harm the information transfer. On the contrary, neurons have opposite stimulus selectivities, positive noise correlation help the information transfer. In many

cases, correlations will not influence the information transfer (see for example, a theoretical study by Moreno-bote et al. 2014). To overcome the spiking correlation problem induced by correlated presynaptic input, Renart et al. (2010) built a densely connected neural network with excitatory and inhibitory currents canceling each other on a fast time scale (fine-scale balance). By using such mechanism, they showed that, theoretically, a fine-scale balanced network could generate an asynchronous state of population activity with a low mean spiking correlation despite correlated inputs (Renart et al. 2010). In the same study described above, Yu et al. (2014) found the E/I balance-induced sparse representation of odorant inputs was accompanied by a decorrelated state of mitral cell firing patterns, and the maximal decorrelation value existed at the optimal level for synaptic excitation and inhibition for the sparseness (Fig. 5.2c). In another interesting experimental work, researchers manipulated the excitation/inhibition ratio (E/I ratio) to obtain an optimal E/I ratio that maximized the information capacity by trading off between a lower correlation state (induced by low E/I ratio) and moderate activity (induced by a relatively high E/I ratio) (Shew et al. 2011).

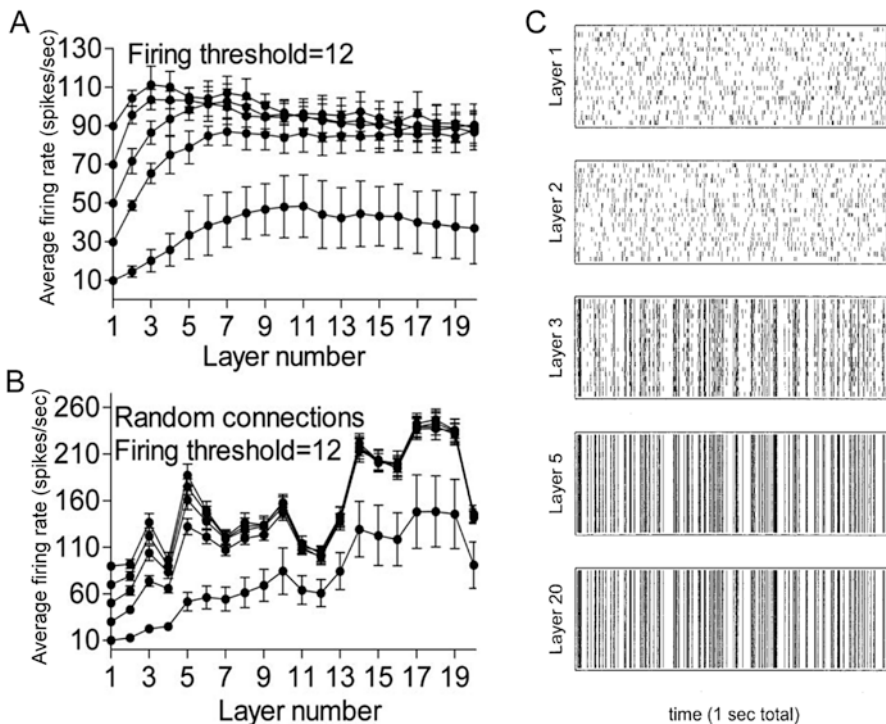
## 5.5 E/I Balance and Information Propagation

Because the brain is highly modular, and spiking activity may carry a lot of neural information, it is important that the spiking activity can be transmitted from one module to another with high fidelity. Indeed, Perkel and Bullock (1968) noted that one of the major components of a typical neural code should be the inclusion of reliable information transmission or information propagation. The identification of the conditions under which spiking activity can propagate with high fidelity has attracted the attention of many theoretical researchers in the recent decade (Diesmann et al. 1999; Kistler and Gerstner 2002; van Rossum et al. 2002; Litvak et al. 2003; Kumar et al. 2008, 2010). Researchers usually address the propagation topic using a model of a cascade of neural assemblies in which a single neuron can participate at multiple levels (termed a feedforward network). To construct a more biologically oriented neural network, theoretical works tend to embed a feedforward sub-network into a larger recurrent neural network. However, neurons in the feedforward sub-network receive stronger correlated excitation than the rest of the recurrent network. This may destabilize the activity of the recurrent network. To solve this defect, Aviel et al. (2003) added inhibitory neurons into the subset network to balance the extra excitation.

Researchers have identified two modes of spiking activity propagation: the asynchronous mode (rate code, with information about the stimulus is carried and propagated in the firing rate of a neuron or the average of population activity) (van Rossum et al. 2002; Litvak et al. 2003; Vogels and Abbott 2009) and synchronous mode (temporal code, with information about the stimulus is carried and propagated by the precise timing of action potentials) (Aertsen et al. 1996; Diesmann et al. 1999; Gewaltig et al. 2001; Litvak et al. 2003; Kumar et al. 2008). For a network with a

feedforward configuration, the firing rates at each layer could be stabilized to a constant level after an initial increase (Fig. 5.3a) (Litvak et al. 2003). The network inhibition precisely balanced with excitation plays a key role in modulating the mean firing rate level. Small deviations from the precise balance would result in a large fluctuation in the firing rate at each layer (Fig. 5.3b) (Litvak et al. 2003). Litvak et al. (2003) showed that the population synchrony could be formed after a few layers and then propagate stably through many layers in such a feedforward network with the excitation firing rate balanced with the inhibitory firing rate (Fig. 5.3c).

Beyond the fidelity of information propagation, the regulation of the spiking activity is also important for neural coding. A given module of the neural system has the potential to respond to several different signal pathways. To accomplish a single task, some mechanisms must exist to selectively block or boost some signal pathways. Recently, Vogels and Abbott (2009) showed that a detailed balance of excitation and inhibition in the target feedforward network group could be a potential



**Fig. 5.3** Propagation of firing rate in a multilayer feedforward network. (a) Average firing rate of different layers in the precisely balanced network. (b) Average firing rate of different layers in the feedforward network with small deviations from the precise balance. (c) Raster plot showing the firing pattern of excitatory neurons in different layers in a feedforward network with balanced firing rates between excitation and inhibition. Adapted from Litvak et al. (2003)

gating mechanism; there information transmission can be gated ‘on’ by adjusting the excitatory and inhibitory gains to upset this detailed balance.

## 5.6 Conclusion

Stimulus representation and information propagation are two basic functions of neural coding. E/I balance, which is acknowledged as a fundamental paradigm for many brain functions, has been demonstrated to play a fundamental role in shaping the neural coding process. On one hand, the E/I balance can significantly increase the coding efficiency and energy efficiency to extend the coding capacity by promoting a sparse representation and signal decorrelation. Intuitively, an E/I ratio that is too high leads to excitatory dominance, resulting in high correlation (low level of coding efficiency) and activity (high level of energy consumed); however, an E/I ratio that is too low leads to suppressed activity with low information content. Therefore, the tradeoff between these two aspects requires the balance of the excitatory and inhibitory currents. On the other hand, based on recent theoretical studies, the E/I balance also plays a vital role in determining the fidelity of spiking activity propagation and gating of the multiple signal pathways. More experimental investigations are expected to test theoretical hypotheses and predictions in the near future. As discussed above, the implementations and functions of the two types of E/I balance—global balance and fine-scale balance, are different from one another. More studies are also needed to demonstrate the exact differences between the effects of these two types of balance on neural coding.

Here, we mainly discussed the roles of E/I balance in neural coding, while many additional studies have focused on the role of E/I balance in other brain functions, e.g., the formation of memories (Vogels et al. 2011; Lim and Goldman 2013) and the information storage process. More studies are expected to clarify the effects of E/I balance in memory formation and examine how the information storage process benefits from the E/I balance.

**Acknowledgements** YY thanks for the support from the National Natural Science Foundation of China (81761128011, 31571070), Shanghai Science and Technology Committee support (16410722600), the program for the Professor of Special Appointment (Eastern Scholar SHH1140004) at Shanghai Institutions of Higher Learning, and Omics-based precision medicine of epilepsy entrusted by the Key Research Project of the Ministry of Science and Technology of China (Grant No. 2016YFC0904400) for their support.

**Author Contributions** Yuguo Yu and Shanglin Zhou designed research; Yuguo Yu and Shanglin Zhou performed research; Shanglin Zhou and Yuguo Yu wrote the paper. All authors reviewed the manuscript.

**Conflict of Interest** The authors declare no competing financial interests.

## References

- Aertsen A, Diesmann M, Gewaltig MO. Propagation of synchronous spiking activity in feedforward neural networks. *J Physiol Paris*. 1996;90(3–4):243–7.
- Alle H, Roth A, Geiger JRP. Energy-efficient action potentials in hippocampal mossy fibers. *Science*. 2009;325:1405–8. Available at: <http://www.sciencemag.org/content/325/5946/1405>
- Amit DJ, Brunel N. A model of spontaneous activity and local delay activity during delay periods in the cerebral cortex. *Cereb Cortex*. 1997;7:237–52. Available at: <http://cercor.oxfordjournals.org/content/7/3/237?related-urls=yes&legid=cercor;7/3/237>
- Anderson JS, Carandini M, Ferster D. Orientation tuning of input conductance, excitation, and inhibition in cat primary visual cortex. *J Neurophysiol*. 2000;84:909–26.
- Atallah BV, Scanziani M. Instantaneous modulation of gamma oscillation frequency by balancing excitation with inhibition. *Neuron*. 2009;62:566–77. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0896627309003511>
- Attwell D, Laughlin SB. An energy budget for signaling in the grey matter of the brain. *J Cereb Blood Flow Metab*. 2001;21:1133–45. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11598490>
- Averbeck BB, Latham PE, Pouget A. Neural correlations, population coding and computation. *Nat Rev Neurosci*. 2006;7(5):358–66.
- Aviel Y, Mehring C, Abeles M, Horn D. On embedding synfire chains in a balanced network. *Neural Comput*. 2003;15:1321–40. Available at: <http://www.mitpressjournals.org/doi/abs/10.1162/089976603321780290>
- Barlow HBH. Possible principles underlying the transformation of sensory messages. In: Rosenblith W, editor. *Sensory communication*. Cambridge, MA: MIT Press; 1961. p. 217–34.
- Beggs JM. The criticality hypothesis: how local cortical networks might optimize information processing. *Philos Trans R Soc A Math Phys Eng Sci*. 2008;366:329–43. Available at: <http://rsta.royalsocietypublishing.org/content/366/1864/329.abstract>
- Boerlin M, Machens CK, Denève S. Predictive coding of dynamical variables in balanced spiking networks. *PLoS Comput Biol*. 2013;9:e1003258.
- Brunel N. Dynamics of sparsely connected networks of excitatory and inhibitory neurons. *J Comput Neurosci*. 2000;8:183–208.
- Cardin JA, Palmer LA, Contreras D. Stimulus feature selectivity in excitatory and inhibitory neurons in primary visual cortex. *J Neurosci*. 2007;27:10333–44.
- D’Amour JA, Froemke RC. Inhibitory and excitatory spike-timing-dependent plasticity in the auditory cortex. *Neuron*. 2015;86:514–28.
- Dehghani N, Peyrache A, Telenczuk B, Le Van Quyen M, Halgren E, Cash SS, Hatsopoulos NG, Destexhe A. Dynamic balance of excitation and inhibition in human and monkey neocortex. *Sci Rep*. 2016;6:1–12. Available at: <http://arxiv.org/abs/1410.2610>
- Denève S, Machens CK. Efficient codes and balanced networks. *Nat Neurosci*. 2016;19:375–82. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/26906504>
- Dhawale AK, Hagiwara A, Bhalla US, Murthy VN, Albeanu DF. Non-redundant odor coding by sister mitral cells revealed by light addressable glomeruli in the mouse. *Nat Neurosci*. 2010;13:1404–12. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3208311&tool=pmcentrez&rendertype=abstract>
- Diesmann M, Gewaltig MO, Aertsen A. Stable propagation of synchronous spiking in cortical neural networks. *Nature*. 1999;402:529–33.
- Froemke RC. Plasticity of cortical excitatory-inhibitory balance. *Annu Rev Neurosci*. 2015;38:195–219.
- Froemke RC, Merzenich MM, Schreiner CE. A synaptic memory trace for cortical receptive field plasticity. *Nature*. 2007;450:425–9. <https://doi.org/10.1038/nature06289>.
- Gewaltig MO, Diesmann M, Aertsen A. Propagation of cortical synfire activity: survival probability in single trials and stability in the mean. *Neural Netw*. 2001;14:657–73.

- Haider B, Duque A, Hasenstaub AR, McCormick DA. Behavioral/systems/cognitive neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition. *J Neurosci*. 2006;26:4535–45.
- Haldeman C, Beggs JM. Critical branching captures activity in living neural networks and maximizes the number of metastable states. *Phys Rev Lett*. 2005;94(5):058101.
- Kistler WM, Gerstner W. Stable propagation of activity pulses in populations of spiking neurons. *Neural Comput*. 2002;14:987–97. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11972904>
- Koren V, Deneve S. Computational account of spontaneous activity as a signature of predictive coding. *PLoS Comput Biol*. 2017;13:e1005355.
- Koulakov AA, Rinberg D. Sparse incomplete representations: a potential role of olfactory granule cells. *Neuron*. 2011;72:124–36.
- Kumar A, Rotter S, Aertsen A. Conditions for propagating synchronous spiking and asynchronous firing rates in a cortical network model. *J Neurosci*. 2008;28:5268–80. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18480283>
- Kumar A, Rotter S, Aertsen A. Spiking activity propagation in neuronal networks: reconciling different perspectives on neural coding. *Nat Rev Neurosci*. 2010;11:615–27. Available at: <http://www.nature.com/doi/10.1038/nrn2886>
- Lim S, Goldman MS. Balanced cortical microcircuitry for maintaining information in working memory. *Nat Neurosci*. 2013;16:1306–14. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23955560>
- Litvak V, Sompolinsky H, Segev I, Abeles M. On the transmission of rate code in long feedforward networks with excitatory-inhibitory balance. *J Neurosci*. 2003;23:3006–15.
- Liu BH, Li YT, Ma WP, Pan CJ, Zhang LI, Tao HW. Broad inhibition sharpens orientation selectivity by expanding input dynamic range in mouse simple cells. *Neuron*. 2011;71:542–54.
- Liu G. Local structural balance and functional interaction of excitatory and inhibitory synapses in hippocampal dendrites. *Nat Neurosci*. 2004;7:373–9.
- Martinez LM, Alonso J-M, Reid RC, Hirsch JA. Laminar processing of stimulus orientation in cat visual cortex. *J Physiol*. 2002;540:321–33.
- Migliore M, Hines ML, Mctavish TS, Shepherd GM. Functional roles of distributed synaptic clusters in the mitral-granule cell network of the olfactory bulb. *Front Integr Neurosci*. 2010;4:122.
- Monier C, Fournier J, Frégnac Y. In vitro and in vivo measures of evoked excitatory and inhibitory conductance dynamics in sensory cortices. *J Neurosci Methods*. 2008;169:323–65.
- Moreno-Bote R, Beck J, Kanitscheider I, Pitkow X, Latham P, Pouget A. Information-limiting correlations. *Nat Neurosci*. 2014;17:1410–7. <https://doi.org/10.1038/nn.3807>
- Murphy BK, Miller KD. Balanced amplification: a new mechanism of selective amplification of neural activity patterns. *Neuron*. 2009;61:635–48. <https://doi.org/10.1016/j.neuron.2009.02.005>
- Murray JD, Anticevic A, Gancsos M, Ichinose M, Corlett PR, Krystal JH, Wang XJ. Linking microcircuit dysfunction to cognitive impairment: effects of disinhibition associated with schizophrenia in a cortical working memory model. *Cereb Cortex*. 2014;24:859–72.
- Nawroth JC, Greer CA, Chen WR, Laughlin SB, Shepherd GM. An energy budget for the olfactory glomerulus. *J Neurosci*. 2007;27:9790–800. Available at: <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.1415-07.2007>
- Niven JE, Laughlin SB. Energy limitation as a selective pressure on the evolution of sensory systems. *J Exp Biol*. 2008;211:1792–804. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18490395>
- Okun M, Lampl I. Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nat Neurosci*. 2008;11:535–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18376400>
- Olshausen BA, Field DJ. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*. 1996;381:607–9. <https://doi.org/10.1038/381607a0>
- Perkel DH, Bullock TH. Neural coding: a report based on an NRP work session. *Neurosci Res Progr Bull*. 1968;6:219–349.

- Poil S-S, Hardstone R, Mansvelder HD, Linkenkaer-Hansen K. Critical-state dynamics of avalanches and oscillations jointly emerge from balanced excitation/inhibition in neuronal networks. *J Neurosci*. 2012;32:9817–23. Available at: <http://www.jneurosci.org/content/32/29/9817.short>
- Poo C, Isaacson JS. Odor representations in olfactory cortex: “sparse” coding, global inhibition, and oscillations. *Neuron*. 2009;62:850–61. <https://doi.org/10.1016/j.neuron.2009.05.022>.
- Renart A, de la Rocha J, Bartho P, Hollender L, Parga N, Reyes A, Harris KD. The asynchronous state in cortical circuits. *Science*. 2010;327:587–90.
- Runyan CA, Schummers J, Van Wart A, Kuhlman SJ, Wilson NR, Huang ZJ, Sur M. Response features of parvalbumin-expressing interneurons suggest precise roles for subtypes of inhibition in visual cortex. *Neuron*. 2010;67:847–57.
- Sengupta B, Laughlin SB, Niven JE. Balanced excitatory and inhibitory synaptic currents promote efficient coding and metabolic efficiency. *PLoS Comput Biol*. 2013;9(10):e1003263.
- Shadlen MN, Newsome WT. Noise, neural codes and cortical organization. *Curr Opin Neurobiol*. 1994;4:569–79.
- Shadlen MN, Newsome WT. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J Neurosci*. 1998;18:3870–96. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9570816>
- Shew WL, Yang H, Petermann T, Roy R, Plenz D. Neuronal avalanches imply maximum dynamic range in cortical networks at criticality. *J Neurosci*. 2009;29:15595–600.
- Shew WL, Yang H, Yu S, Roy R, Plenz D. Information capacity and transmission are maximized in balanced cortical networks with neuronal avalanches. *J Neurosci*. 2011;31:55–63. Available at: [https://www.researchgate.net/publication/49731639\\_Information\\_Capacity\\_and\\_Transmission\\_Are\\_Maximized\\_in\\_Balanced\\_Cortical\\_Networks\\_with\\_Neuronal\\_Avalanches](https://www.researchgate.net/publication/49731639_Information_Capacity_and_Transmission_Are_Maximized_in_Balanced_Cortical_Networks_with_Neuronal_Avalanches)
- Shu Y, Hasenstaub A, McCormick DA. Turning on and off recurrent balanced cortical activity. *Nature*. 2003;423:288–93.
- Simoncelli EP, Olshausen BA. Natural image statistics and neural representation. *Annu Rev Neurosci*. 2001;2001:1193–216.
- Tan AY, Zhang LI, Merzenich MM, Schreiner CE. Tone-evoked excitatory and inhibitory synaptic conductances of primary auditory cortex neurons. *J Neurophysiol*. 2004;92:630–43. Available at: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=14999047](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=14999047)
- Tan AYY, Wehr M. Balanced tone-evoked synaptic excitation and inhibition in mouse auditory cortex. *Neuroscience*. 2009;163:1302–15.
- Tan AYY, Brown BD, Scholl B, Mohanty D, Priebe NJ. Orientation selectivity of synaptic input to neurons in mouse and cat primary visual cortex. *J Neurosci*. 2011;31:12339–50.
- van Rossum MCW, Turrigiano GG, Nelson SB. Fast propagation of firing rates through layered networks of noisy neurons. *J Neurosci*. 2002;22:1956–66.
- van Vreeswijk C, Sompolinsky H. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*. 1996;274:1724–6. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8939866>
- van Vreeswijk C, Sompolinsky H. Chaotic balanced state in a model of cortical circuits. *Neural Comput*. 1998;10:1321–71. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9698348>
- Vinje WE, Gallant JL. Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*. 2000;287:1273–6.
- Vogels TP, Abbott LF. Signal propagation and logic gating in networks of integrate-and-fire neurons. *J Neurosci*. 2005;25:10786–95. Available at: <http://www.jneurosci.org/content/25/46/10786>
- Vogels TP, Abbott LF. Gating multiple signals through detailed balance of excitation and inhibition in spiking networks. *Nat Neurosci*. 2009;12:483–91. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19305402>
- Vogels TP, Sprekeler H, Zenke F, Clopath C, Gerstner W. Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. *Science*. 2011;334:1569–73. Available at: <http://science.sciencemag.org/content/334/6062/1569.abstract>

- Wehr MS, Zador AM. Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature*. 2003;426:442–6. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/14647382>
- Wilent WB, Contreras D. Dynamics of excitation and inhibition underlying stimulus selectivity in rat somatosensory cortex. *Nat Neurosci*. 2005;8:1364–70.
- Wolfe J, Houweling AR, Brecht M. Sparse and powerful cortical spikes. *Curr Opin Neurobiol*. 2010;20:306–12. <https://doi.org/10.1016/j.conb.2010.03.006>.
- Wu GK, Arbuckle R, Liu BH, Tao HW, Zhang LI. Lateral sharpening of cortical frequency tuning by approximately balanced inhibition. *Neuron*. 2008;58:132–43.
- Xue M, Atallah BV, Scanziani M. Equalizing excitation-inhibition ratios across visual cortical neurons. *Nature*. 2014;511:596–600. <https://doi.org/10.1038/nature13321>.
- Yang H, Shew WL, Roy R, Plenz D. Maximal variability of phase synchrony in cortical networks with neuronal avalanches. *J Neurosci*. 2012;32:1061–72. Available at: <http://www.jneurosci.org/content/32/3/1061.abstract>
- Yizhar O, Fenno LE, Prigge M, Schneider F, Davidson TJ, O’Shea DJ, Sohal VS, Goshen I, Finkelstein J, Paz JT, Stehfest K, Fudim R, Ramakrishnan C, Huguenard JR, Hegemann P, Deisseroth K. Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature*. 2011;477:171–8. <https://doi.org/10.1038/nature10360>.
- Yu L, Zhang C, Liu L, Yu Y. Energy-efficient population coding constrains network size of a neuronal array system. *Sci Rep*. 2016;6:19369. Available at: <http://www.nature.com/srep/2016/160119/srep19369/full/srep19369.html>
- Yu Y, Hill AP, McCormick DA. Warm body temperature facilitates energy efficient cortical action potentials. *PLoS Comput Biol*. 2012;8:e1002456.
- Yu Y, Mctavish TS, Hines ML, Shepherd GM, Valenti C, Migliore M. Sparse distributed representation of odors in a large-scale olfactory bulb circuit. *PLoS Comput Biol*. 2013;9:1–20.
- Yu Y, Migliore M, Hines ML, Shepherd GM. Sparse coding and lateral inhibition arising from balanced and unbalanced dendrodendritic excitation and inhibition. *J Neurosci*. 2014;34:13701–13. Available at: <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.1834-14.2014>