Augmenting Flexibility: Mutual Inhibition Between Inhibitory Neurons Expands Functional Diversity

Belle Liu[1, 2, *], Alexander James White[1, *, ✉], and Chung-Chuan Lo[1, 3, ✉]

[1] *Institute of Systems Neuroscience, National Tsing Hua University, Hsinchu, Taiwan*

[2] *Department of Physics, National Tsing Hua University, Hsinchu, Taiwan*

[3] *Brain Research Center, National Tsing Hua University, Hsinchu, Taiwan*

Abstract

In the face of a complex environment, it is necessary for neural circuits to develop a "Swiss army knife" toolkit, in which neural networks flexibly switch between different functionalities to meet the demands placed on them. Here, we develop a theoretical framework for how recurrent circuits give rise to this flexibility. Moreover, mutual inhibition, an abundant yet less studied structure, is critical in expanding the functionality of the network, far beyond what feedback inhibition alone can accomplish. Using dynamical systems theory, we show that mutual inhibition doubles the number of cusp bifurcations in small neural circuits. As a concrete example, we build a class of functional motifs we call Coupled Recurrent Inhibitory and Recurrent Excitatory Loops (CRIRELs). These CRIRELs have the advantage of being multi-functional, performing a plethora of functions, including decisions, memory, toggle, etc. Finally, we demonstrate how this trend holds for larger networks.

Introduction

Organisms often exhibit a vast range of behaviors, and these behaviors require the neural network to perform different tasks. In light of this, an efficient design is for neural networks to be capable of flexibly switching between various functions upon demand [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11], instead of having each task be performed by a different network [1, 4, 12, 13, 14]. This flexibility has indeed been observed experimentally [3, 4, 15, 16, 17, 18, 19, 20], but how this flexibility came to be remains a topic of investigation.

* Contributed equally to this work.

✉ Send Correspondence to Chung-Chuan Lo: cclo@mx.nthu.edu, or Alexander White: ajw131@case.edu

It has been hypothesized that a key component of this flexibility is the recurrent nature of neural connections [12, 13, 21]. Quite often, the focus of recurrent structures is on mutual excitation [6, 22, 23, 24, 25, 26] or feedback inhibition [6, 23, 27, 28, 29, 30, 31], and the role of mutual inhibition has received less attention (the term "recurrent" has been used interchangeably across literature – see Fig. 1a for how it is defined in this study). However, there are several pieces of evidence suggesting that mutual inhibition plays a more crucial role than previously thought [20, 32, 33, 34].

First of all, it has been widely reported that inhibitory neurons increase the variability of a network [3, 20, 27, 35, 36, 37, 38], and supporting evidence from connectome studies revealed that mutual inhibition is abundant and formed locally [34, 35, 39], possibly performing local computations [27]. Secondly, evidence shows that there are functional differences between feedback and mutual inhibition. For example, feedback inhibition is composed of an interconnected pair of excitatory and inhibitory neurons, which balances the network [30, 31, 40], allowing it to approximate arbitrary functions [21] and are useful in gain control. On the other hand, many studies have examined inhibition in a more general way. Recurrent inhibition (either feedback or mutual inhibition) increases the number of basins of attraction [6], allowing the network to perform functions such as winner-take-all decision [6, 41, 42, 43], bistable perception [32, 33], oscillations [44, 13, 12], associative memory [45, 20] and grid formation [46]. Interestingly, we show a wide variety of these computations can be performed by introducing mutual inhibition into the network. This, along with supporting connectome studies, all point to the idea that mutual inhibition is crucial in expanding a network's functional repertoire.

At last, we note that a network's functional capacity is not determined solely by the number of synapses [12, 40, 35, 47]. In particular, while each type of synapse increases the amount of potential computations a network can have, the network is not necessarily flexible if it cannot seamlessly transition between different functional modes by modulating the parameters. Whether the network is capable of easy transitions is highly dependent on the underlying bifurcation, which is the phenomenon of switching between behavioral characteristics through the change of model parameters [12, 48, 49, 50]. Specifically, a previous theoretical study states that networks near complicated bifurcations are computationally useful [12], because only networks near a bifurcation have non-trivial dynamics [12, 21, 50]. In short, one can control the specific operations a neuron can perform by controlling which bifurcations a circuit is near.

Given the lack of attention on mutual inhibition, we set out to systematically model its effect on a network's functionality and dynamics. We hypothesize that mutual inhibition can increase the flexibility of a network by increasing the number of bifurcations the network is near. In particular, it can subsume the functions originally present in the network, and introduce new functions on top of that. To further illustrate this argument, we

investigated the dynamics of a family of recurrent circuits we dubbed CRIREL (Coupled Recurrent Inhibition and Recurrent Excitation Loop). This type of structure is used across literature regarding a myriad of phenomena, such as decision making and deviance detection [51, 52].

To showcase the effect of different recurrent structures on a circuit's dynamics, we first conducted a series of statistical analyses. Based on these statistical results, we then delve into a detailed dynamical analysis of the CRIREL circuit. Our analysis shows how, by including mutual inhibition, even a small 4-neuron circuit is capable of a wide range of computations. Finally, we extend our study to a large random network and show that mutual inhibition is critical for increasing the entropy of its working memory state, which implies that mutual inhibition makes the dynamical landscape of a network more diverse.

Results

Effects of recurrent structures on complexity of circuit dynamics

We compared several circuit architectures to investigate how different recurrent structures contribute to a circuit's dynamics (Fig. 1a,b). The first metric we employed to characterize the diversity of circuit behavior was the number of equilibrium points. If the circuit contains multiple equilibrium points, it is more likely to lead to complex activity patterns in response to different inputs, an indication of the circuit's computational ability.

We begin with the CRIREL circuit, which contains all three recurrent structures, then remove these connections one at a time. Motif A contains all the recurrent structures contained in the other motifs (Fig. 1a). Motif B does not contain mutual inhibition, and motif C does not have mutual excitation. In motifs D and E, the feedback inhibition loop is broken – in motif D, feedback inhibition is removed, and in motif E, feedback excitation is removed. Motif F is a pure feedforward structure that was chosen as arbitrarily, and motif G is a feedforward structure that forms a single loop.

For each of the motifs, we performed a parameter sweep, where the synaptic weights of the circuit and bias currents were varied. We ran a total of 18900 trials. The protocol is designed so that it pushes the system in different directions in the phase space, ensuring that as much as possible is explored, which allows the algorithm to detect different equilibrium points. The firing rate sequence of the trial is then put though the Elbow method to judge the number of clusters that are formed, which correlates to the amount of stable equilibrium points in the system.

The results show that motif A has the largest parameter regime in which the system not monostable, i.e. the number of equilibrium points exceed 1 (Fig. 1c; monostable in blue). Furthermore, motifs that have mutual

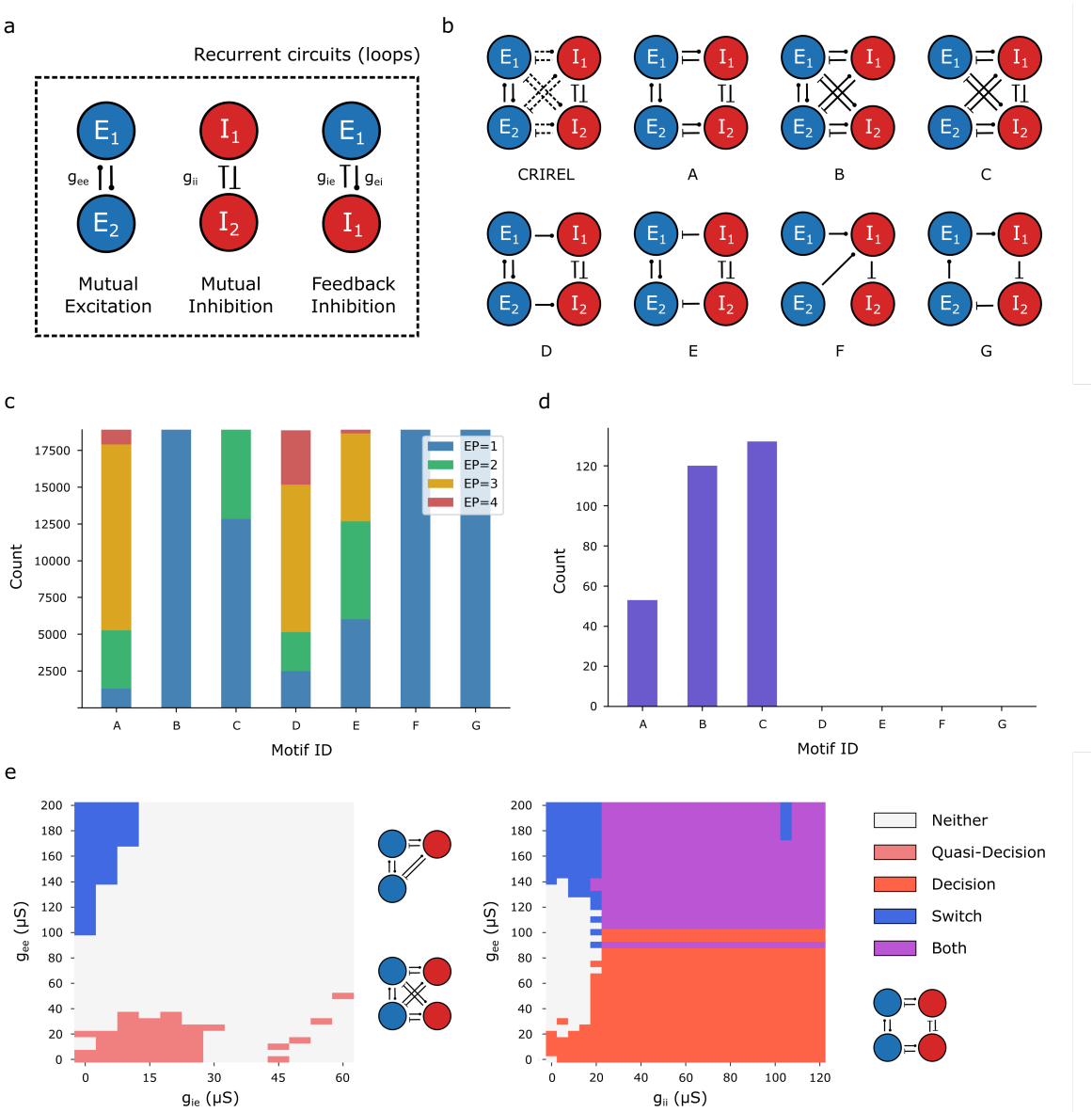

Figure 1: Motifs and statistical analysis of dynamical diversity. (a) Definition of the terms "recurrent" circuits, "mutual" excitation/inhibition and "feedback" inhibition used in this paper. Synaptic weights used in part of the results are assigned based on its pre- and post-synaptic neurons, hence there are four types of weights: excitatory-to-excitatory (g_{ee}), excitatory-to-inhibitory (g_{ei}), inhibitory-to-inhibitory (g_{ii}) and inhibitory-to-excitatory (g_{ie}). (b) The motif labeled "CRIREL" is the definition of CRIREL circuits. The rest of the motifs are indexed by alphabets in no particular order. Blue represents excitation neurons, and red inhibition. (c) Parameter sweep for equilibrium points in different motifs across some parameter space. In this graph, 1 equilibrium point is in blue (EP=1), 2 is green (EP=2), 3 is yellow (EP=3) and 4 is green (EP=4). (d) Parameter sweep for CPGs in different motifs across some parameter space. (e) Parameter sweep for motifs A and B. For motif B (left), the relevant parameters concerning quasi-decision making and switch are g_{ee} and g_{ie}. For motif A (right), the relevant parameters are g_{ee} and g_{ii}. Blue represents regions where the circuit can perform switching, coral-red represents quasi-decision making, red represents decision making, and white represents no functions present. The purple region is the regime in which the system can perform both functions simultaneously.

inhibition (motifs A, C, D, E) have parameter sets with equilibrium point counts larger than two, while those with mutual excitation but not mutual inhibition has at most counts of two, and feedforward motifs have no

bistability at all. In contrast, the presence or absence of feedback inhibition does not make such a large difference. For example, motifs A, D and E all have high numbers of equilibrium points, but D and E do not contain feedback inhibition while A does.

While the first metric shows the number of equilibrium points in the network, it does not reveal whether other stable structures are possible within the network, such as oscillation given constant input. Therefore, the second metric we employed tested whether central pattern generators (CPG) are present in the network. We count the number of distinct inter-spike intervals (ISI) within the inhibitory neurons, and if the number of ISIs is greater than 1, that implies that there is a CPG (see methods for more details). This analysis shows that CPGs are only found in A, B and C – the ones that contain feedback inhibition (Fig. 1d).

In summary, we illustrated the functional differences between mutual and feedback inhibition – mutual inhibition increases the number of equilibrium points in the network, while feedback inhibition allows the network to oscillate when given constant input.

Functional differences between feedback and mutual inhibition

The statistical results indicate that mutual inhibition increases the dynamical states, but does not explain the mechanism behind it. Hence here we consider an explicit example that compares feedback inhibition with mutual inhibition, and compare motif A with motif B. Specifically, we compared the two motifs' ability to perform different functions within some relevant parameter regime.

Motif B contains mutual excitation, which allows it to perform a switch-like function, where some brief positive current pulse can transfer the system to the ON state, and a negative pulse can turn it off. Motif B also has feedback inhibition, which skews the basin of attraction one way, therefore making one of the excitatory neuron fire faster than the other, performing a quasi-decision like function. However, these two functions cannot coexist within the circuit (Fig. 1e, left panel). If we add recurrent inhibition to the circuit, however, this problem can be resolved. Consider motif A, which has mutual excitation, feedback inhibition, as well as mutual inhibition. This gives a large parameter regime in which motif A contains both functions simultaneously (Fig. 1e, right panel). Here, we stress that even though motif A has more synapses, it is motif B (with mutual inhibition) that is more functional. Therefore, presence of mutual inhibition can add more functions to the underlying neural network.

Dynamics of CRIREL circuits

All of the statistical analysis above point to mutual inhibition being crucial in expanding the functional repertoire of a network. However, to have a better grasp on the functions CRIREL circuits actually bring about, we must look into its dynamics. To this end, we reduced the 4-dimensional CRIREL circuit into a 2-dimensional reduced model, with each of the recurrent circuits compressed into 1 dimension (see Methods and Supplementary Note). Conceptually, it can be understood as preserving the dimensions that are important to the dynamics of interest by considering the properties of the two mutual connections. Because each LIF neuron as a class 1 excitable neuron [48], we can approximate each neuron as a Wilson-Cowen [12, 53] type firing rate model, with each neuron's firing rate as e_1, e_2, i_1 and i_2. Since mutual excitation tends to synchronize, the dimension of $e_1 - e_2$ is less important than $e_1 + e_2$ if the coupling is much weaker than the mutual connection strength, where e_i represents the firing rate of excitatory neuron i, which is denoted as E_i. On the other hand, mutual inhibition desynchronizes, therefore $i_1 + i_2$ is less important than $i_1 - i_2$ when the coupling is weak. Similarly, i_j represents the firing rate of inhibitory neuron j, which is denoted as I_j. Hence, by retaining the dimensions of $e_1 + e_2$ and $i_1 - i_2$ only, we can gain a geometrical understanding of how the system operates, which in turn sheds light on the potential computations recurrent circuits can perform (Fig. 2a).

The resulting reduced model (Fig. 2b) can be visualized using the coordinates $\mathbf{E} \approx e_1 + e_2$ and $\mathbf{I} \approx i_1 - i_2$. The blue and red line represents the nullclines of $\mathbf{E}$ and $\mathbf{I}$ respectively. Note that both nullclines are of a cubic form, which is the normal form of a cusp bifurcation, thereby proving that mutual connections do indeed add cusp bifurcations to the circuit. Since there are two cubic functions, we know that this system has a double-cusp bifurcation. There are four free parameters in this reduced model, which comes from complicated nonlinear mappings of the original neuronal and bias current parameters. For physical intuition, a_1 and b_1 are nonlinear mappings of structural parameters of the neurons, such as capacitance or time constant or synaptic weights. a_2 and b_2 are mappings from the bias currents into the circuit. These parameters allow the cubic functions to change shape as well as translation, which gives rise to many different phase diagrams (Fig. 2c). These phase diagrams correspond to different "modes" of the circuit, in which it performs different functions. We will go over all of the functions we discovered below, but note that this is by no means a complete list of what the circuit is capable of doing.

The functions of CRIREL circuits include rudimentary ones that rely more heavily on mutual excitation, such as switch, toggle and synchronized CPG; it also includes ones in which computations are mainly performed in the mutual inhibition loop, such as decision-making and anti-synchronized CPG. Building from these concepts, we can also have computationally interesting functions, such as working memory, threshold-based filtering and

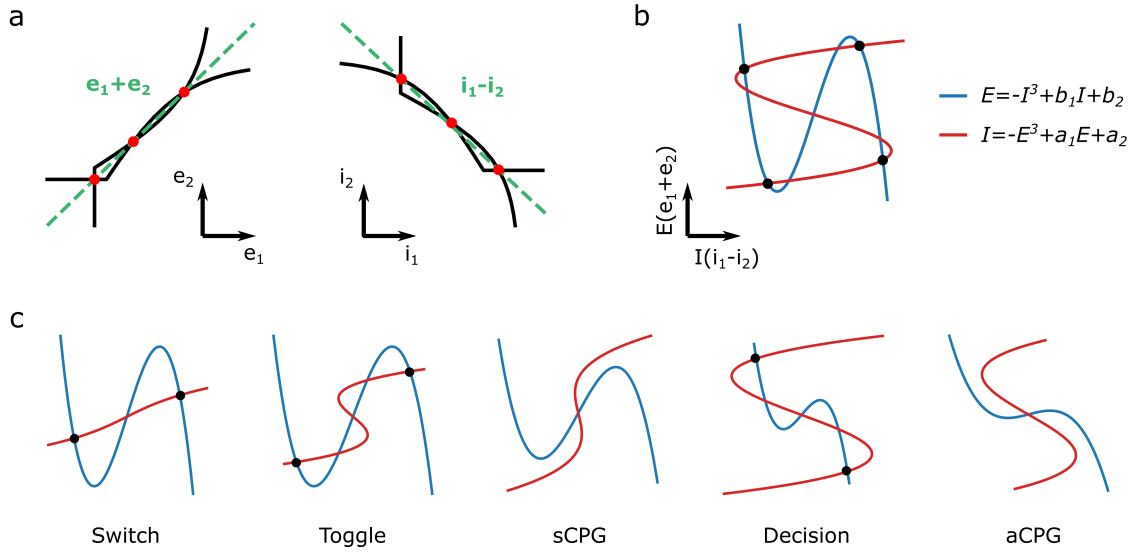

Figure 2: The reduced model of the CRIREL circuit. (a) The phase diagrams of the mutual excitation (left) and inhibition loop (right) when decoupled from one another. The black curves are nullclines and the spots represent equilibrium points. The green dashed line shows the direction in which in interesting dynamics lie on (it is only approximate for the inhibitory case). (b) When weakly coupled together, the CRIREL circuit can be reduced to two dimensions. The blue curve is the nullcline of the mutual excitation loop, red is the nullcline of mutual inhibition, and black dots are stable equilibria. (c) By adjusting the four parameters a_1, a_2, b_1 and b_2, the reduced model will change into different configurations, each of them corresponding to different functions.

timing-based decision-making.

Computations in the excitatory subsystem

Switch. We first demonstrate one of the simplest computations, which is a switch-like function (Fig. 3a). An intuitive understanding of switching and its behavior is already described in previous sections. Here we demonstrate how to comprehend this function geometrically using the reduced model. The reduced model for switch has only two stable equilibrium points separated by a saddle equilibrium (Fig. 3b). The simulation begins at the equilibrium point where all the neurons are silent, i.e. the OFF state. When a brief positive current pulse is given to one or both of the excitatory neurons, the **E**-nullcline shifts upwards, annihilating the OFF state. Therefore the system transits to the remaining equilibrium point, i.e. the ON state, and remains there even as the input is removed and the **E**-nullcline returns to its original position. This means that the circuit will continue firing even in the absence of input. When a negative current pulse is applied, the opposite happens, and the system returns to the OFF state. This shows that the circuit is bistable. We see that the reduced model is consistent with the intuitive explanation given above (Fig. 3b). Moreover, we tested that this switch into the upstate was present for a broad range of potential input stimuli (Fig. 3c).

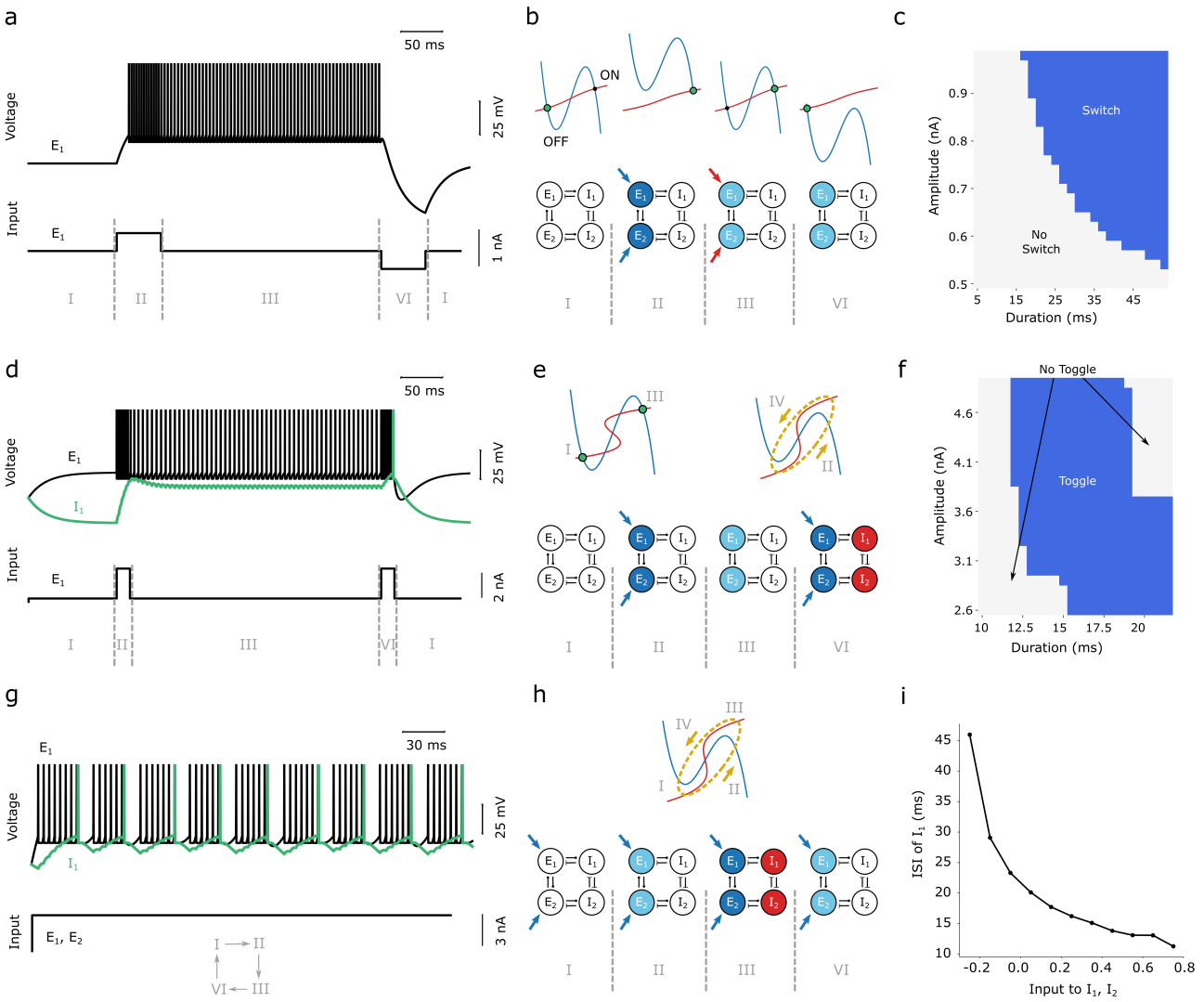

Figure 3: Switch, toggle and synchronous CPG. (a) Switch voltage traces. Top row: the voltage trace of neuron E_1. Bottom row: the current input into E_1. The roman numerals represent different phases of the input, and corresponds to the roman numerals shown in panel (b). (b) Switch phase diagrams and schematics of neural activity. Top row: switch function phase diagram. The green dot represents the state of the system. Bottom row: Schematic diagram of the relative activity of each neuron during different time periods. Blue represents excitation, and red inhibition. The darker the color, the higher the activity. (c) Inputs that induce a switch. A parameter sweep over duration and amplitude of the input pulse that induces a switch into the upstate. Blue represent stimuli that induce a switch, while gray represents those that do not. (d) Toggle voltage traces. Top row: the voltage traces of E_1 (in black) and I_1 (in green). Bottom row: the current given to E_1. The roman numerals correspond to the roman numerals shown in panel (e). (e) Toggle phase diagrams and schematics of neural activity. Top row: toggle function phase diagram. Bottom row: Schematic diagram of the firing rates of each neuron during different time periods. (f) Inputs that induce a toggle. A parameter sweep over duration and amplitude of the input pulse that toggles between the upstate and downstate. Blue represent stimuli that induce a switch, while gray represents those that do not. (g) synchronous CPG voltage traces. The rows are identical to those of (d). Here, however, the current into E1 is changed from pulses into a constant input. The roman numerals arranged in a loop means that the four phases are occurring repeatedly. (h) sCPG phase diagrams and schematics of neural activity. (i) ISI of inhibitory spikes. Parameter sweep over the inhibitory bias current, and its effect on the period of the CPG.

Computations in the excitatory subsystem with feedback inhibition

Toggle. A function that is similar to switch is toggle, except that it can be turned off by a positive current pulse instead of a negative pulse. We can visualize this by plotting the voltage traces of one of the excitatory neurons (which are synchronized) and one of the inhibitory neurons (Fig. 3d). The first pulse causes the system to go to the ON state, and the firing rate of the inhibitory neurons are insufficient to inhibit the excitatory loop. However, when a second identical pulse enters the system, the firing rate of the excitatory loop increases, and concurrently the inhibitory loop. This extra push is enough for the inhibitory neurons to turn off the excitatory neurons, thus returning the whole system to inactivity.

Again, we can gain insight by examining the phase portrait of the system. The simple shifting mechanism of the switch-system no longer applies here. As the impulse current excites the excitatory loop and consequently the inhibition loop, the inhibitory response deepens, causing the inhibitory nullcline to change its shape (Fig. 3e). This in turns causes both the ON and OFF state to annihilate, and the system goes from 5 equilibrium points to one unstable spiral node. This means that for a brief amount of time, there exists a limit cycle and the system's state rotates. When the impulse input ends, the inhibitory system relaxes and the 5 equilibrium points return, but by that time the system is already trapped in the ON state. The second pulse repeats the process and rotates the system yet again, allowing it to return to the OFF state. We also examine the required amplitude and duration of stimuli required to toggle the system (Fig. 3f). Interestingly, for very long duration the system may rotate a full rotation, rather than a half rotation. This will cause the system not to toggle for long duration stimuli.

Synchronized Central Pattern Generator. The presence of toggle implies that the circuit is also capable of sustaining periodic oscillations, that is, capable of becoming CPGs. Note that if the impulse current is changed into a constant bias current, the system will repeat the process of toggling over and over again, thus undergoing a synchronized bursting behavior (Fig. 3g). The phase diagram of synchronous CPG is the limit-cycle mode (Fig. 3h). Moreover, we showed that inhibitory bias current can tune the network over a wide operating frequency (Fig. 3i).

Computations in the inhibitory subsystem

Decision making. Switch, toggle and synchronous CPG rely heavily on the dynamics of the mutual excitation loop. Decision making, on the other hand, is mainly due to mutual inhibition. For a decision network, when two inputs are injected into the neurons E_1 and E_2, the larger of the two inputs will "choose" their corresponding inhibitory neuron, thus preventing the other inhibitory neuron from firing (Fig. 4a). Logically, this is the exact

counterpart of the switch system, the difference being that the nonlinearity now lies in the **I**-nullcline instead of the **E**-nullcline. While inputs shift the **E**-nullcline either up or down in the switch or toggle phase diagram, it shifts the inhibitory neuron either left or right in this case, hence forcing the system to decide between the two equilibrium points. Geometrically, this corresponds to "rotating" the nullclines of the synchronized CPG by 90 degrees (Fig. 4b). To quantify the performance of decision making, we examined the accuracy and reaction time of the circuit (Fig. 4c). The results are consistent with decision making networks containing two competing populations of excitatory neurons – as the input becomes more coherent, accuracy increases and reaction time decreases.

Anti-toggle (anti-synchronized toggle). Just as decision making is the counterpart of the switch function, anti-toggle is the counterpart of toggle. By taking advantage of the reduced model, we can rotate the nullclines by 90 degrees to generate an anti-toggle from toggle. This is achieved by swapping the parameters of the excitatory and inhibitory systems. As with the case of toggling, in this mode the excitatory neurons of the excitatory sub-circuit can be turned on and off by positive input pulses. The difference is that now when one inhibitory neuron is turned off, the other one automatically turns on (Fig. 4d, e). This function is extremely robust, as it works in a wide range of parameters (Fig. 4f).

Asynchronized Central Pattern Generator. Similar to how toggle gives rise to synchronized CPGs, we can do the same for anti-toggle to an anti-synchronized CPG. by changing the current pulses into a constant bias. This change will not directly yield anti-synchronized CPGs, since a constant bias would be reinforcing the system to stay within the same equilibrium point. However, if we introduce time-scale separation, e.g. mechanisms such as slow GABA synapses, calcium dynamics, and second-messenger systems, the circuit can oscillate. We can visualize this by taking advantage of a limit cycle in the reduced system (Fig. 4g, h). Again, we examined how the inhibitory bias current controls the operating frequency of the CPG (Fig. 4i), and showed that it operates over a wide rage of frequencies.

Computations in the CRIREL system

Working memory. The situation described above complicates when the bistability of the excitatory sub-system is combined with decision making of the inhibitory sub-system. Since the ON state exists, this means that the decision can be remembered even after the input is withdrawn because of the excitatory sub-system. When receiving strong input, the excitation sub-system enters the ON state asymmetrically. One of the excitatory neurons will have higher firing rates, thus allowing its neighboring inhibitory neuron to fire and inhibit the other one. Furthermore, the feedback inhibition strength is not strong enough to turn the excitatory sub-system off,

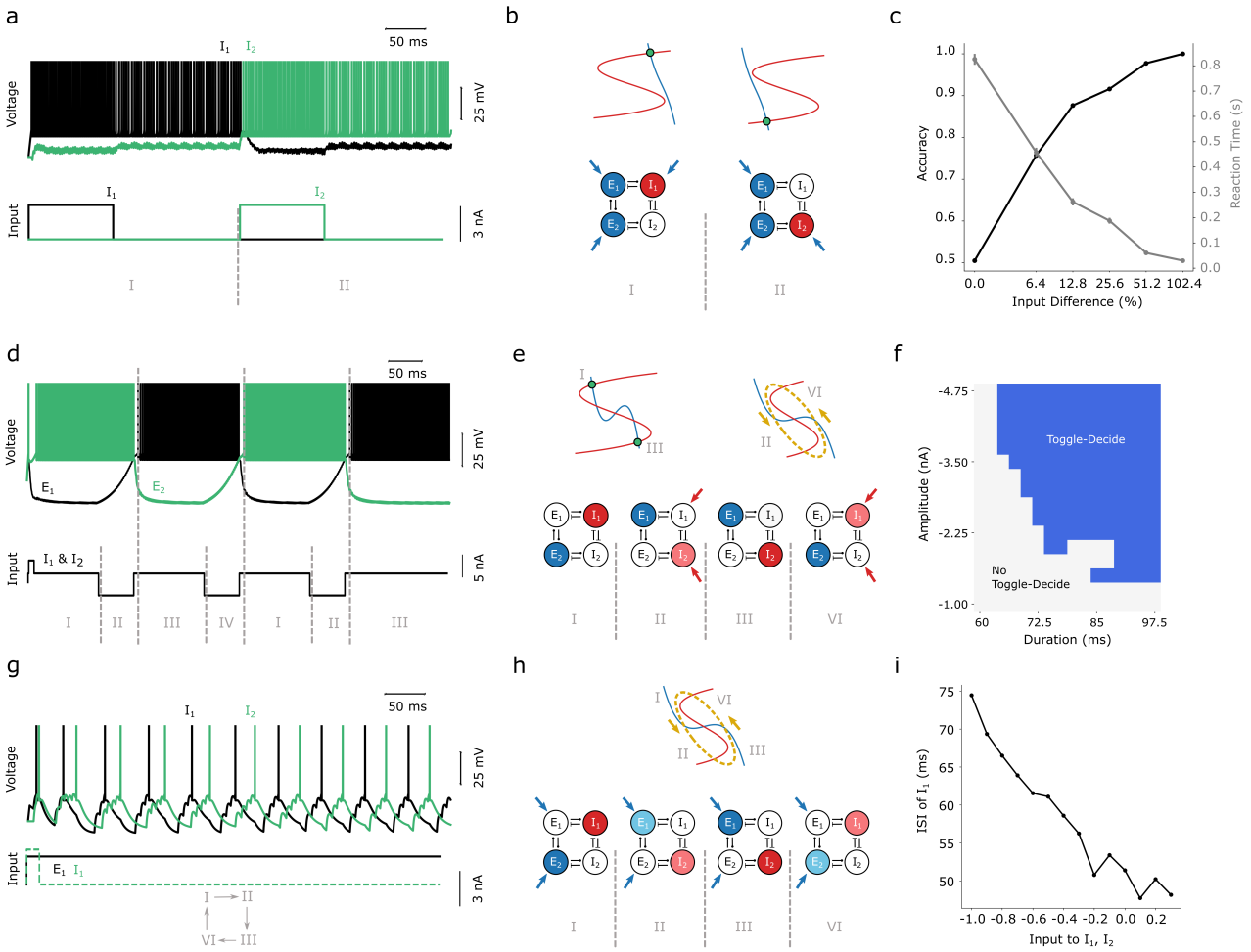

Figure 4: Decision making and asynchronous CPG. (a) Decision making voltage traces. Top row: the voltage traces of I_1 (in black) and I_2 (in green). Bottom row: the current given to I_1 (in black) and I_2 (in green). Here, E_1 and E_2 are given constant bias currents to maintain the decision after the input is removed. (b) Decision making phase diagrams and schematics of neural activity. (c) Psychometric function of decision making. As the input difference of the two inputs increases, the accuracy (black) increases and the reaction time (grey) decreases. (d) Anti-toggle voltage traces. (e) Anti-toggle phase diagrams and schematics of neural activity. (f) A parameter sweep over duration and amplitude of the input pulse that anti-toggle. Blue represent stimuli that induce a anti-toggle, while gray represents those that do not. (g) Asynchronous CPG voltage traces. (h) Asynchronous CPG phase diagrams and schematics of neural activity. (i) ISI of inhibitory spikes. Parameter sweep over the inhibitory bias current, and its effect on the period of the CPG.

thus the system "remembers" the input (Fig. 5).

There are two protocols to change the memory currently being maintained in the system. One can clear the system with a reset signal and the system will return to rest (Fig. 5a). The memory can also be overwritten directly (Fig. 5b-d).

The analysis of this system is identical to the decision-making system, except that there could be 4 equilibrium points (Fig. 5e). It can indeed be confirmed that this nullcline configuration is present in the full CRIREL system, where the inhibitory neurons can either be both on (i.e. (1, 1)), one on and one off (i.e. (1, 0) or (0, 1))

or both off (i.e. (0, 0)). Thus this system is capable of forming a working memory system with 2-bit memory. This is a perfect example of how simple functions such as decision making and switch can combine to yield a more interesting and useful function.

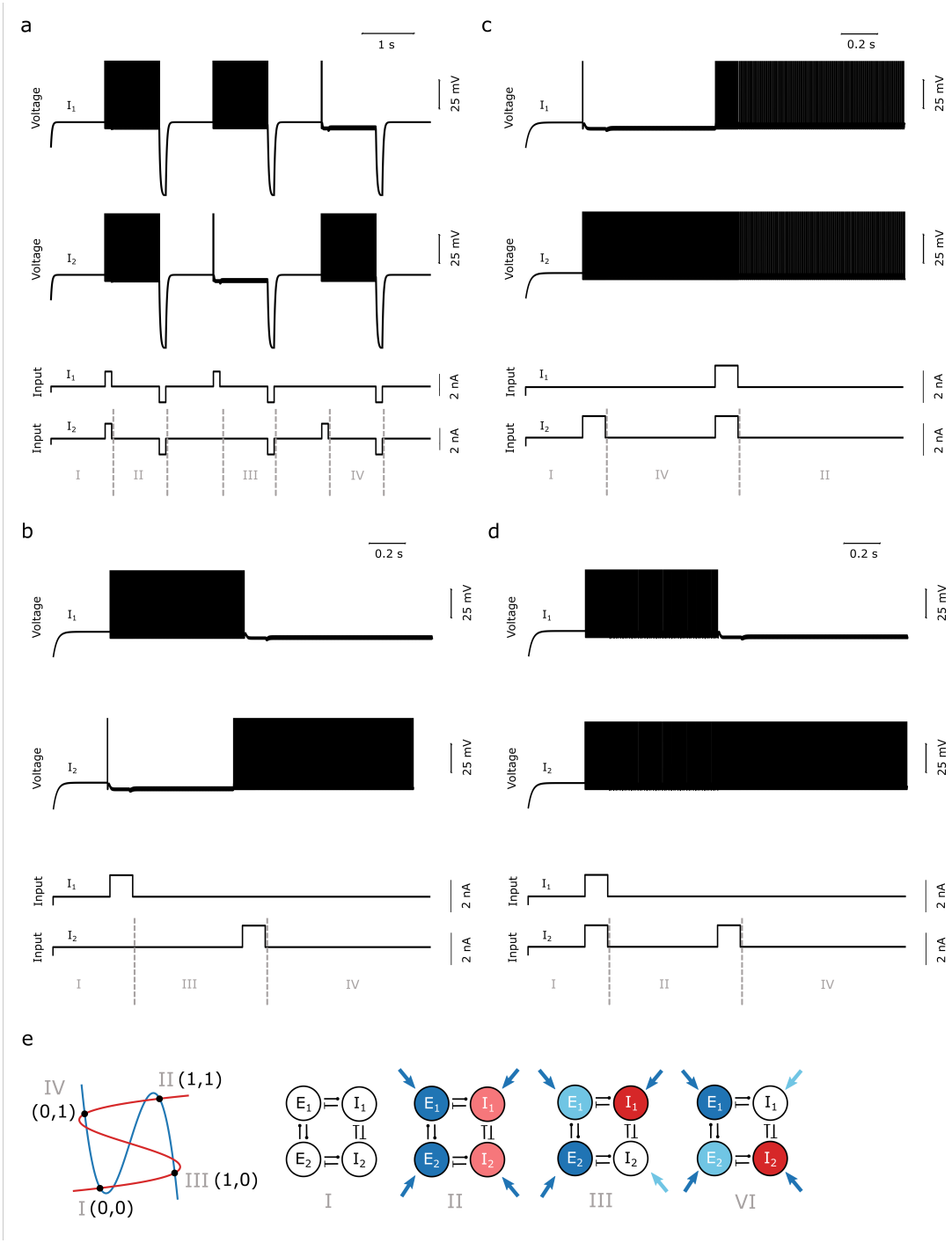

Figure 5: Working memory. (a) Voltage trace. This circuits stores a "two bit" signal in working memory, i.e. (1,1), (1,0), (0,1) and (0,0), corresponding to (I_1 on, I_2 on), (I_1 on, I_2 off), (I_1 off, I_2 on), (I_1 off, I_2 off). This can be seen in the top two rows of the graph, showing the voltage traces of I_1 and I_2. The bottom two rows are the inputs to the four neurons. (b-d) Switching memory without reset. (b) Switching from (1,0) to (0,1). (c) Switching from (0,1) to (1,1). (d) Switching from (1,1) to (0,1). (e) Working memory phase diagrams and schematics of neural activity. Top row: phase diagram. Bottom row: Schematic diagram of the neural activity.

Threshold-based filtering. The capability of making decisions allows CRIRELs to be utilized in many different ways, including useful functions such as creating a safety net for hyperactivity. If we allow the CRIREL circuit to take on asymmetrical parameters (Fig. 6a, left) of membrane and synapse, the circuit can become a low-pass filter in terms of amplitudes, where small inputs are remembered by the circuit (working memory like), and larger ones terminate the bistability (toggle like) (Fig. 6b). This computation is possible precisely due to the presence of asymmetries in the network . When one of the excitatory neurons receive a small input, since I_1 has stronger synaptic weights, I_1 will be chosen and thus inhibit I_2. However, when the input is sufficiently large, I_2 has smaller capacitance and time constants than I_1, and therefore will be excited faster than I_1, thus winning the decision. The feedback inhibition from I_1 is weak, therefore when I_1 is chosen, it does not affect the network's bistability; however, the feedback inhibition from I_2 is strong, therefore whenever I_2 is chosen, the network shuts down. Hence the circuit serves as a gateway that filters out inputs that are too large, which could be potentially useful for preventing epilepsy. The threshold of the circuit can be adjusted by manipulating the bias current of I_1 (Fig. 6c).

Timing-based detection. Aside from making decisions based on the magnitude of the input, what is perhaps even more intriguing is detecting the relative timing of inputs (Fig. 6 d,e). Even a sub millisecond difference between the inputs is sufficient for the network to perceive. Timing-based detections can be accomplished if we allow sufficiently strong feedback inhibition, as we did in toggling. Under this condition, two modes are allowed: the coincidence detection mode (Fig. 6d) and the anti-coincidence detection mode (Fig. 6e). In the anti-coincidence mode, the network will chose whichever input came first, but will turn off completely if inputs arrive at the same time. In the coincidence mode, the network will only fire if the inputs arrive simultaneously (Fig. 6e). Note that for these two modes, an additional neuron is added to the CRIREL circuit that combines the output of E_1 and E_2 (Fig. 6a, right). However, this extra neuron is only added for the sake of having a more clear-cut output. The computation for these two modes are still done within the CRIREL circuit itself.

Here, coincidence is controlled primarily by mutual excitation. The synaptic input is sufficient to turn the system on when both inputs arrive at the same time, but if there is a delay in the system then, the system will not fire. Conversely, the anti-coincidence system takes advantage of the toggle like dynamics, where it turns the system off if and only if both of the inputs arrive simultaneously. Otherwise, the inhibitory neurons do not fire with sufficient strength to turn the system off. Coincidence has a narrow range of operation (Fig. 6f) because it relies on a more delicate balance between the inhibitory and excitatory subsystems. One the other hand, anti-coincidence operates over a much broader range of parameters (Fig 6g).

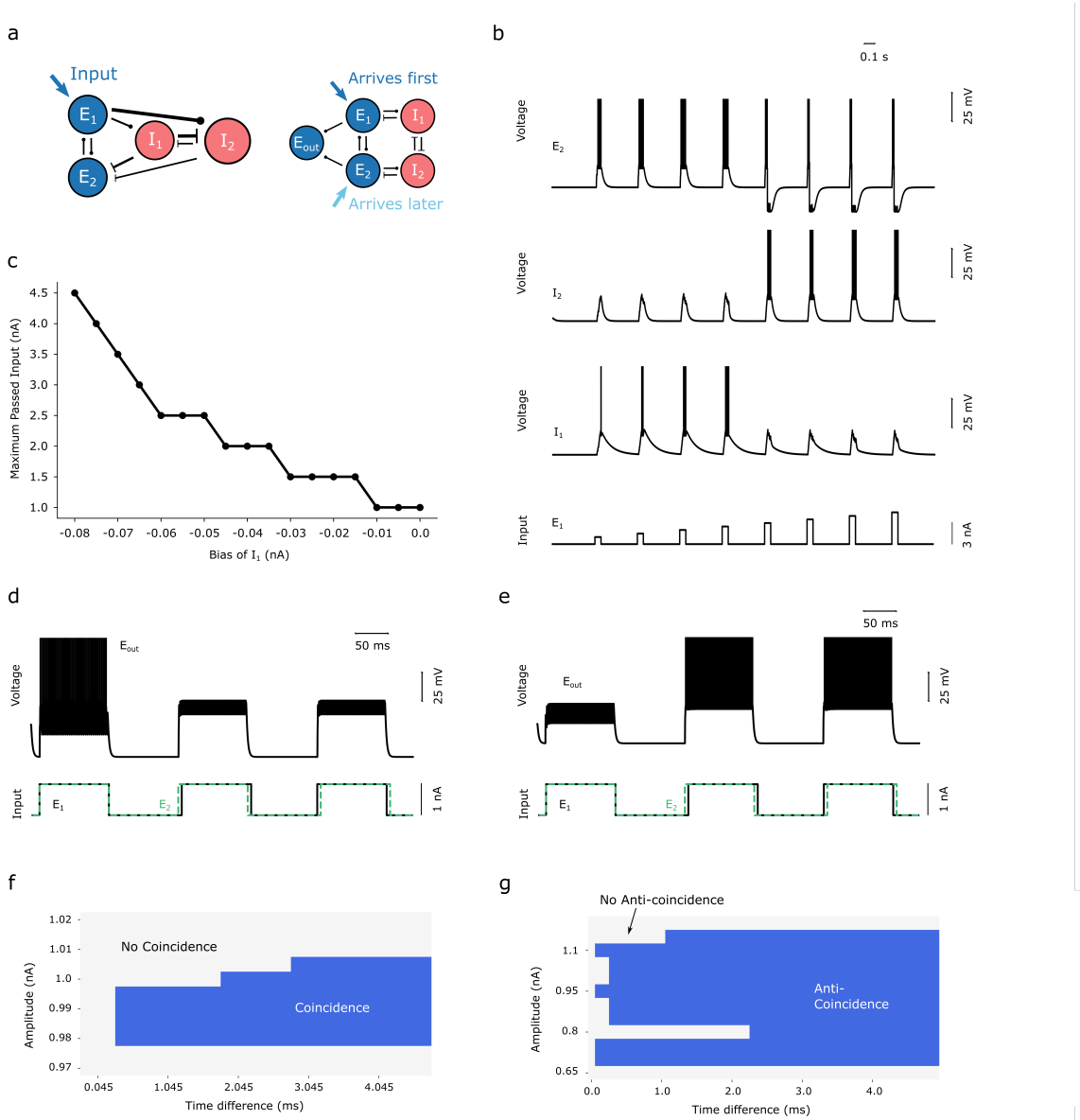

Figure 6: (a) Circuit motifs for threshold-based filtering (left) and timing based detection (right). The asymmetries of the parameters are shown schematically in the diagram: thicker lines represent larger weights, and vice versa. (b) Voltage traces of threshold-based filtering. The top three row shows the voltage traces of E_2, I_2 and I_1, and the bottom row shows the injected current into E_1. Here we inject neuron E_1 with different pulses of increasing amplitude. (c) Maximum amplitude of input that passes the filter as a function of bias into I_1. The steps are discrete because the input we used are discrete. (d) Timing based detection, coincidence-mode. The top row is the voltage trace for the output neuron. The bottom row is the input for E_2 (in black) and E_1 (in green). Since the timing difference between the input is small (5 ms), for visual demonstration the difference is exaggerated. (e) Timing based detection, anti-coincidence-mode. The top row is the voltage trace for the output neuron. The bottom row is the input for E_2 (in black) and E_1 (in green). (f-g) A parameter sweep for both coincidence (f) and anti-coincidence (g) over both amplitude of signal and timing difference. Blue represents the existence of the function, while light gray represents its absence.

Memory in large network

So far, we have explicitly explained how mutual inhibition adds a cusp bifurcation to the circuit, which allows the circuit to perform various functions. Here, we extend this argument to large random networks. We focused on

the function of working memory in particular, as working memory requires use of both on-off switches (excitatory cusps) as well as decision making (inhibitory cusps). We showed that an increase in mutual inhibition strength allows the network to have a wider variety of memory states. However, other functions that are present in the CRIREL circuit can also be reproduced in the large network as well (see Fig. S3).

The random network consists of 100 neurons, 75 excitatory and 25 inhibitory, where each connection has a 50 percent probability of being present (Fig. 7a). g_{ie} and g_{ii} are varied from 0 to approximately 30 (μS). For each parameter set, the excitatory neurons of the network are given random current pulses, and after the stimulation stops the network relaxes into a steady state. Since the stimulation is random, each time the steady state might be identical or different, and this is repeated for 100 times per parameter set. Each steady state is recorded as a "word", where different words represent different steady states (Fig. 7b; see Methods). The entropy for the distribution of the words across 100 trials, H_{word}, is calculated for the various parameter sets, which reflects how diverse the steady states are (Fig. 7c).

For the same g_{ie} value, H_{word} increases as g_{ii} increases (Fig. 7c,d). To understand the reason behind this, we turn to two other entropy measures, H_{symbol} and H_{binary} (Fig. 7b). Simply put, H_{symbol} measures how diverse the firing rate across population is, for a single trial. If the activities of all neurons are high (Fig. 7d left) then there are not a lot of symbols available to form a word, hence H_{symbol} would be low and consequentially H_{word} as well. Moreover, It can be shown that there is a regime of suitable g_{ii} to g_{ie} ratio where H_{symbol} is maximized (Fig. 7c). This trend is quite different from H_{word}, where it increases almost monotonically with g_{ii}. Therefore, an increase in the number of symbols is not the main mechanism behind what we observed in H_{word}.

Therefore, we used yet another method for calculating entropy, H_{binary}. Here, if a neuron fires, it is labeled as "1", while a silent neuron is labeled "0" (Fig. 7b). Each time the network relaxes into a steady state, it yields a distribution of 1s and 0s across the neuron population, and the entropy for that distribution is calculated. Essentially, if the neurons' firing rates are more "all or nothing" (or "winner takes all") – so that the ratio between the count of 1s and 0s are more equal – then H_{binary} will be higher. Moreover, H_{binary} is augmented as g_{ii} increases, which follows a similar trend to H_{word} (Fig. 7c), implying that the increase in memory capacity as shown by H_{word} may be caused by the neurons' firing becoming more "all or nothing", which is in turns due to an increase in g_{ii}.

Furthermore, we note that for large g_{ii} values, its entropy value is near maximal. That is, if all 100 trials result in unique states, then $H_{word} = log_2 \frac{1}{100} \approx 6.64$, and we see that the entropy values of large g_{ii} are near that value.

To conclude, increasing g_{ii} increases the number of memory states the network contains. Consistent with

how mutual inhibition introduces new bifurcations to the CRIREL circuit, mutual inhibition in a large network also complicates its dynamics and yields a more computationally useful product. The large network is also capable of performing other functions found in the CRIREL circuit (Fig. S3).

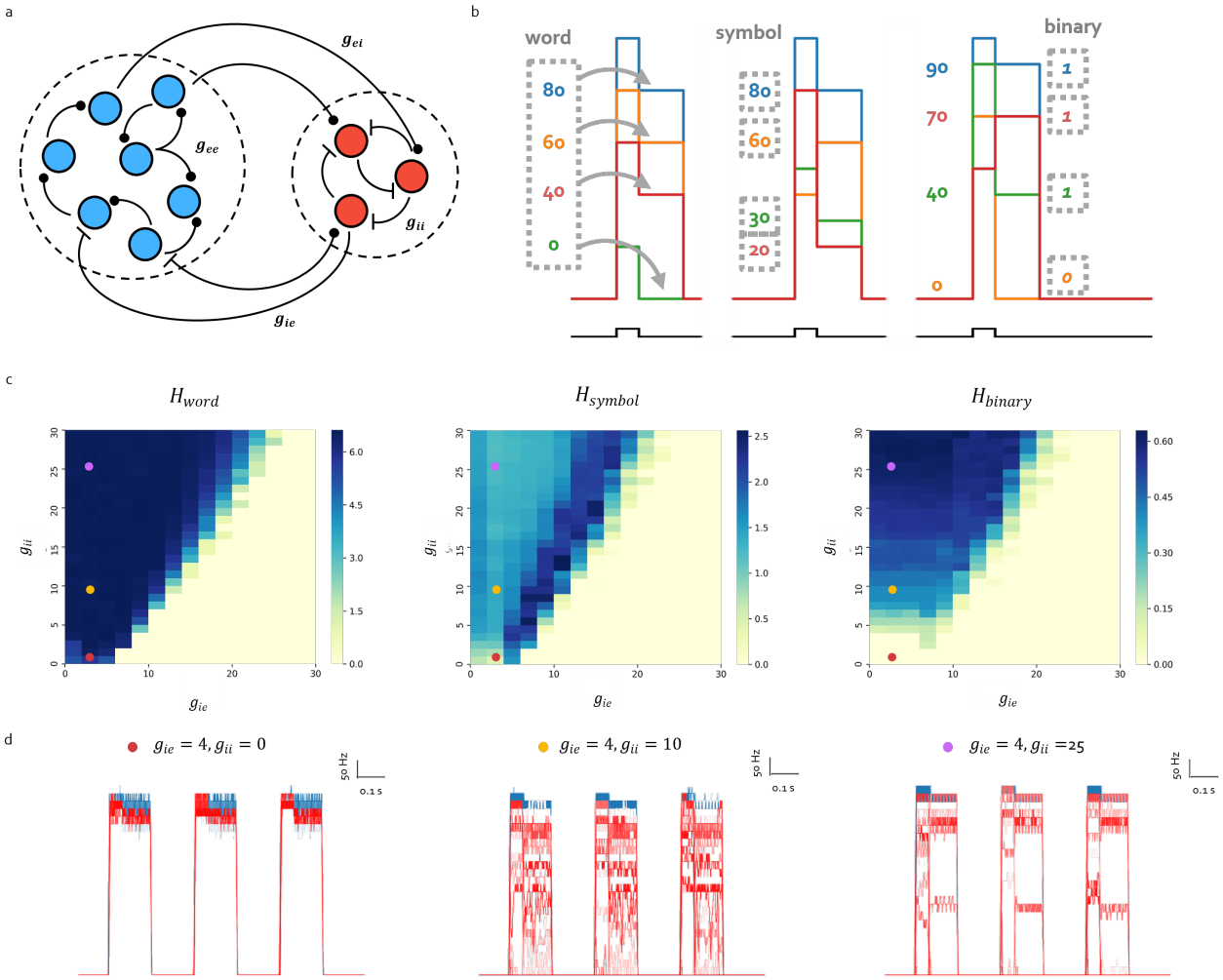

Figure 7: Memory in large network. (a) Schematic diagram of large random network. (b) Schematic diagram of how entropy is calculated. Each line represents the firing rate of a neuron, and the firing rate of the neuron population is defined as a "word". The words across 100 trials forms a probability mass function, which is used to calculate H_{word}. The word itself is also a probability mass function for different firing rates, which also has an associated entropy. The trial-average of this entropy is called H_{symbol}. Alternatively, each neuron can be labeled as on (1) or off (0), and the entropy associated with this is called H_{binary}. (c) From left to right: H_{word}, H_{symbol}, H_{binary} for different g_{ii} and g_{ie}. (d) Firing rate traces for neurons across 3 trials, with different synaptic weights. Each trial showcases how the steady state activity differs between having different g_{ii} values. Blue traces represent excitatory neurons, and red traces inhibitory neurons. From left to right: increasing g_{ii}.

Discussion

In the present study, we show that mutual inhibition plays a key role in increasing the computation complexity and functional repertoire. This is evident from: (1) motifs that contain mutual inhibition exhibit more functions than those without, (2) the ability of a specific class of the motif, called CRIREL, to flexibly perform multiple biologically relevant functions for the same synaptic weights, and (3) higher working memory capacity in large neural networks with mutual inhibition than those without. We also demonstrate that the key mechanism that underlying the functionality of mutual inhibition lies in its ability to increase the basins of attraction from an increase in the underlying cusp bifurcations.

This is inline with past research, as connections between inhibitory neurons has been shown in many different contexts [27, 28, 29, 34, 35, 54]. It is interesting to note that in a recent study [39] that reconstructed single-cell level brain-wide connectome of fruit flies[55, 56, 57], putative inhibitory neurons receive more inhibitory than the excitatory inputs. Moreover, studies in the hippocampus of rats have found a plethora of mutually inhibitory parvalbumin interneurons [34]. These findings imply the importance of mutual inhibition in the brain network. We have shown computationally that these types of connections are useful and perform a unique role in computational processing by expanding the number of computational states a network has. We also hypothesis the existence of a function called anti-toggle, that is capable of seamlessly switching between different basins of attraction, which is potentially useful in tracking changes of state, working memory based decisions, or even counting. Furthermore, the existence of decisions and anti-toggle in the inhibitory sub-network could potentially be one mechanism behind simultaneous suppression and excitation of inhibitory interneurons observed in various studies of cortical neurons [27, 35, 51].

Moreover, our results are in line with studies that have reported an increase in variability [35, 51] and working memory maintenance [20] when the connection strength of inhibition is increased. In our study, we showed that increases in g_{ii} is more responsible for the increase in the number of steady states than g_{ie}. This tracks with the presence of a multiple cusp bifurcation. As the large network simulations showed, for each inhibitory neuron with all-to-all inhibitory connections added, there is an additional cusp bifurcation. This doubles the number of unique combinations of states the network can perform.

One area our research can be extended to is examining the effects of gap junctions on the network. Some have found that mutual inhibition and gap junctions tend to coexist in some circuits [34]. It has been shown that gap junctions have a tendency to average out the winner-take-all aspect of the decisions [58].

In addition, in our study we have mostly overlooked plastic synapses. Several studies have shown that this

is an important component in network functionality [7, 12, 44, 45, 47]. Moreover, recent work have shown that synaptic depression can desegregate different cusp bifurcations by weakening the strong synaptic effects between mutual excitatory loops [40, 59]. Aside from that, by using plastic synapses, one can construct networks that are both balanced and near bifurcations [21]. In this paper different functions require different synaptic weights. It is possible that a motif circuit can switch between functions based on learning or plasticity induced weight changes. This can be a future direction of this study.

A final direction for future work is to examine the best way to control networks with mutual inhibition. Several studies have suggested that networks near bifurcations are computationally useful [12, 21, 49, 50]. Some studies use balanced inhibition to tune the network to be useful [21]. Here, we wonder if this process can be used to control networks with mutual inhibition. If mutual inhibition is removed, are these networks still computationally useful?

To conclude, we have shown that mutual inhibition is key in expanding the functionality of the network. This is possible because mutual inhibition expands the number of cusp bifurcations in a network. Ultimately this helps elucidate why neural circuits in the brain are so flexible in their operation and are capable of such rich dynamics.

Methods

Spiking neural model and synaptic model

All simulations in this study are ran on flysim [39]. Each individual neuron was modeled as a leaky integrate-and-fire neuron:

$$C_m \frac{dV}{dt} = -\frac{C_m}{\tau}(V - E_L) + \sum I_{j,r}^{syn}(t) \tag{1}$$

$$\text{When } V > V_{Th} \text{ then } V \to V_{reset},\ t \to t + t_{reset} \tag{2}$$

where C_m is the capacitance, V the membrane potential, τ is the time constant, E_L the reversal potential, and I_j^{syn} the current from the jth presynaptic neuron, and r represents the receptor. When the membrane potential reaches V_{Th}, the neuron fires, and is promptly reset to V_{reset} after an amount of time t_{reset}. The parameters can be found in Table 1.

Excitation is modeled as an exponentially decaying AMPA receptor, while inhibition is modeled as and

exponentially decaying GABA receptor. More precisely, we have have excitatory and inhibitory current given by

$$I_r(t) = g_r s_r(t)(V - E_r)$$

$$\tau_r \frac{ds_r}{dt} = -s + \sum_{t_{spike}} \delta(t_{spike} - t)$$

Here, r represents the receptor ($r \in \{AMPA, GABA\}$), I is the current of the receptor, V is the membrane potential, E is the reverse potential, g is the conductance of the receptor, s is the gating variable for the receptor ion channel. s is governed by a linear ODE with time constant τ, and is forced by a sum of dirac delta impulses $\sum_{t_{spike}} \delta(t_{spike} - t)$, where t_{spike} is the timing of presynaptic spikes. It is important to note that $g_{ee} = g_{AMPA}$ when both pre and post synaptic neurons are excitatory and $g_{ei} = g_{AMPA}$ when presynaptic neuron is excitatiory and postsynaptic neuron is inhibitory. Likewise, $g_{ii} = g_{GABA}$ when both pre and post synaptic neurons are inhibitory and $g_{ie} = g_{GABA}$ when presynaptic neuron is inhibitory and postsynaptic neuron is exicatitory.

Spiking Neural Model Parameters			
Parameter	Value	Parameter	Value
C_m	0.5 (nF)	τ_{AMPA}	2 (ms)
τ	20 (ms)	V_{AMPA}	0 (mV)
E_L	-70 (mV)	τ_{GABA}	5 (ms)
V_{reset}	-55 (mV)	V_{GABA}	-90 (mV)
V_{Th}	-50 (mV)		

Table 1: Parameters for neuronal model and synaptic model.

Equilibrium points in microcircuits

It is conceivable that the functional repertoire of a motif is correlated to the abundance of its dynamical states. Here, we employ a statistical approach that provides us an estimate of the number of equilibrium points in relevant parameter space. We compared motifs that are specifically chosen to highlight the presence or absence recurrent loops, along with some randomly chosen feedforward motifs (Fig. 1b). After sweeping through the parameter space, we then estimated the number of equilibrium points in each trial by counting the number of clusters (S1a).

Parameter Space

For a four neuron motif, there are 12 possible synaptic connections. To lower dimensionality in parameters, we constrained synaptic weights with the same type of pre-synaptic and post-synaptic neuron to have the same value, except for threshold based filtering (Fig. 1a). Thus we have four types of synaptic weights: excitatory-to-excitatory (g_{ee}), excitatory-to-inhibitory (g_{ei}), inhibitory-to-inhibitory (g_{ii}) and inhibitory-to-excitatory (g_{ie}). For each motif, we ran simulations over 18900 trials, where each trial has a different synaptic weight parameter set g_{ee}, g_{ie}, g_{ii} (Table S1).

Aside from synaptic weights, we also iterated through different underlying bias currents in the neurons. The bias currents adjust the threshold for firing, allowing inhibitory neurons to fire in the absent of excitatory input (Table S1). This ensures that we capture the dynamic states produced by the inhibitory neurons. For each trial, a brief injected current goes into different pairs of neurons, for example one trial would stimulate E_1 only, and another would stimulate E_1 and I_1 simultaneously. This is ran through all possible pairs. The amplitude of the stimulus changes as well. Both stimulating different pairs as well as varying the stimulus amplitude ensures a wider exploration of the parameter space, thereby allowing us to find more equilibrium points. Each brief pulse of positive stimulation is separated by a negative reset signal. These are shown in S1b. Note that this method may not capture every single equilibrium point present, nor will it be able to detect other stable structures such as limit cycles. Nevertheless, the parameter sweeping still provides valuable information regarding the complexity of the dynamical states of the system.

Clustering

To analyze the amount of equilibrium points present in each trial, we look at the firing rate trace of the motif across the four-dimension space (e_1, e_2, i_1, i_2). We used the elbow method to obtain the amount of clusters in the firing rate trace of the trial [60]. The elbow method detects the number of clusters by iterating through possible number of clusters k and performing K-means clustering, and then evaluating each guess k with its standard deviation, which is $SSE = \sum_{i=1}^{k} \sqrt{\frac{\sum_{j=1}^{N_i}(r_{ij}-\bar{r_i})^2}{N_i-1}}$. By plotting SSE as a function of k, we see that sometimes the plot has a sharp drop in SSE followed by a shallow decrease when k becomes large. The sharp drop is cleverly termed "the elbow" (Fig. S1a). Such a sharp drop usually indicates that the number of clusters assigned has hit the correct amount. If no elbows are present, then the data set most likely only has one cluster. A criteria is needed to determine "how sharp" a drop can be considered an elbow, and this criteria is set so that the slope of the "arm" must be seven times the slope of the "forearm".

We purposely neglected the transient activities 100ms after a positive stimulation pulse or a negative reset

signal in order to neglect "fake" clusters (i.e. clusters that are present, but are not true equilibrium points). Finally, the total number of equilibrium points in all possible parameter sets, which total up to 18900 trials, are calculated for each motif for comparison.

CPG in microcircuits

Another important computation for small motifs is the ability to generate oscillations under constant input. Similar to how the equilibrium points were counted, for CPGs, the parameter space was scanned across suitable values of g_{ie} and g_{ei}. For each trial, we then determine whether the motif under said parameter set was capable of oscillation. The ISI for each trial was calculated, and rounded to the nearest 0.001 (s). Since there is no noise in the system, the periodic firing of a neuron being driven by a constant current has a near perfect period (where the slight differences were accounted for by the rounding procedure), and thus the ISI would consist of one single value. The choice of rounding, 0.001 (s), is hence chosen so that when receiving a constant bias current with no noise, the ISI will be 1 single value, as it should be. For any rounding numbers smaller than that, for example 0.0005 (s), such trials will yield 2 or more ISI values, which is due to time step choices of the simulator and does not really reflect whether the motif is in a CPG mode. For CPGs, the ISI would contain two or more values. The counts for trials with CPGs were calculated for each motif.

Decision Making and Bistability

To further illustrate how feedback and mutual inhibition differs in terms of functions, we focused on two functions in particular: decision making and bistability. We swept through the relevant parameter space, and observed whether motifs with or without mutual inhibition could perform these two functions. Once again, to lower the dimensionality of the parameter space, we constrained synaptic weights with the same type of pre-synaptic and post-synaptic neuron to have the same value. For bistability, since it happens within the excitatory subsystem, the most crucial synaptic weight is g_{ee}. For decision making in feedback inhibitory networks, the important synaptic weight is g_{ie}, since strong g_{ie} allows the circuit to turn off the unfavorable excitatory neuron. For decision making in mutual inhibitory networks, the key synaptic weight is g_{ii}. Therefore, for the purpose of this analysis, we will hold g_{ei} constant and sweep through relevant parameter regimes of g_{ee}, g_{ie} and g_{ii}. The parameters are listed in Table S2.

For each point (g_{ee}, g_{ie}, g_{ii}) in the parameter space, two trials with different input protocols are given to the network. The bistable protocol gives a 100 ms positive pulse to one of the excitatory neurons, resets the circuit using a strong negative pulse, and repeats this process for all of the excitatory neurons present within the

network (Fig.S2). If the network is capable of bistability, then the pulse to one of the excitatory neurons should be able to turn the network to the ON state and remain there until the resetting pulse comes in. Therefore, if the circuit stays on after E_1 is excited, and also stays on after E_2 is excited, then it is classified as being able to perform the switch function.

The decision protocol iterates through all possible pairs of excitatory neurons, where for each time period each pair is given an asymmetrical input (Fig. S2). If the network decides, then there will be a difference in firing rate between a pair of neurons at that time period. The condition for decide is that the "winning" neuron (if one neuron's time-averaged firing rate is 10 Hz larger than the other, it is considered a winner) must switch as the relative input strength switches. For instance, if E_1 wins when input 1 is larger than input 2, then E_2 must win when input 2 is larger than input 1. For motif B, decision is made between the excitatory neurons, hence the aforementioned analysis is performed on the two excitatory neurons. The decision is made between the inhibitory neurons for motif A, hence the decision analysis here is performed on I_1 and I_2. Each point (g_{ee}, g_{ie}, g_{ii}) will be capable of either bistability, quasi-decision, decision, both, or none. The results are then drawn into a heat map, where broader regions of "both functions present" means that it is more possible for the two functions to co-exist.

Dynamical analysis

For further understanding of the statistical results, we reduced the CRIREL system into a two dimensional model given by the equations

$$\tau \frac{dE}{dt} = k_e^* - a_e E + E^3 - \epsilon a_{ie} I \tag{3}$$

$$\tau \frac{dI}{dt} = k_i^* - a_i I - I^3 - \epsilon a_{ei} E \tag{4}$$

where, a_e, a_i, k_e^*, and k_i^* are the parameters of the cusp bifurcations. ϵa_{ie} and ϵa_{ie} are the coupling strengths between each subsystem. The full derivation of the reduced model is given in the Supplementary Note. In this paper we only consider small ϵ, and consider large ϵ out of the scope of this paper. The reduced model allows us to conceptually understand what is occurring dynamically in the spiking neural network. Furthermore, the intuition gained from the reduced model is invaluable, as any function found in the reduced model can also be found in the spiking model for an appropriate set of parameters. Note that this does *not* work in the converse direction, meaning there could be functions in the spiking model that are not in the reduced model. This is especially apparent when ϵ is not small.

As a broad overview, we first begin by showing that the mutual excitatory loop and the mutual inhibitory loop each have a cusp bifurcation. To contrast this, we also showed that feedback inhibition cannot undergo a cusp bifurcation. Next, we coupled the two circuits together, and show that both cusps are maintained. Finally, we worked out the reduced model's simple cubic form.

CRIREL function simulations

The reduced model allowed us to predict the different functions the CRIREL circuit may be capable of performing. These functions are actualized by simulation. While the reduced model is general for the whole class of CRIREL circuits, the results we have shown are simulated on one particular circuit (motif A in Fig. 1b), with the exception of threshold-based filtering, which is simulated on motif H (Fig. 6a). The parameters for simulation are listed in Table S3, and the stimulation are shown in the results section. Parameter searches are illustrated in detail below.

Switch, Toggle and Anti-Toggle. The robustness of the functions determines how realistically it can be implemented in noisy biological systems. For switch, toggle and anti-toggle, we swept through appropriate regimes to showcase its working parameter range (Fig. 3c, f, Fig. 4f). A pulse of a specific amplitude and duration was given to the circuit, in the same manner as displayed in Fig. 3a, d and Fig. 4f respectively.

If the circuit is turned on – defined as mean firing rate larger than zero in the designated time period (200-400 ms) – and off – defined as mean firing rate equal to zero after 500 ms, then the circuit is deemed capable of performing switch/toggle within said parameter set. Toggling is furthermore sensitive to the precise timing of the second pulse near the transition to toggling functionality, which means that within certain parameter sets, the behavior of toggling is not robust when the onset or offset pulse timing varies. Therefore, the onset of the pulse is varied from 500-505 ms, and the circuit is only labeled as capable of performing toggling if it toggles for all 6 onset times.

Similarly, a parameter set is defined as performing anti-toggle if the designated neuron fires after the duration of the input ends, and the other gradually returns to silence (evaluated by seeing if the designated neuron has a higher firing rate than the other).

sCPG and aCPG. For central pattern generators, the ability to adjust its firing period is essential (Fig. 3i, Fig. 4i). This can be achieved by modifying its input into the inhibitory neurons. Again, the protocol mimics those of Fig. 3g and Fig. 4g respectively. The ISI of the inhibitory neurons is calculated by averaging all the ISIs within one simulation trial, while disregarding the initial ISI.

Decision making. The psychometric function of the decision making circuit is shown to quantify its

accuracy and reaction time (Fig. 4c). The input difference $2c$ (%) is defined similar to those of [6], where the input to one population (in our case, one inhibitory neuron) is $1 + c$, and the input for the other neuron is $1 - c$, and c ranges from 0 to 1. Essentially, the larger the input difference, the easier the decision is to make. The input lasts for 100 ms. The accuracy is defined as the number of correct trials over total trials, and in cases where no decision is reached, the winning population is randomly assigned. The reaction time is defined as the time from which the stimulus begins to the time where one population reaches the threshold firing rate of 200 Hz. In cases with no decisions, the reaction time is set at 2 s. The mean reaction time is then calculated by averaging over all trials.

Threshold-based filtering. By tuning the bias current of I_1, the threshold of the filter can be adjusted. The stimulation used here is identical to those of Fig. 6b. The input "passes" the threshold when the maximum firing rate of E_2 reaches 100 Hz within 100 ms after stimulation. Note that the maximum passed input amplitude appears to be discrete because the input used in Fig. 6b is discrete itself.

Coincidence and Anti-coincidence detectors. Similar to how we tested the working parameter range for switch and toggle, we swept through different pulse amplitudes and timing differences for the coincidence and anti-coincidence detectors. A working parameter set is defined as the circuit firing according to its truth table (1 for firing rate larger than zero, and 0 for firing rate equal to zero).

Large random network

To determine whether the flexibility of mutual inhibition is preserved in larger networks, we built a larger network and investigated a specific function, memory. To construct the large network, we used 100 LIF neurons with the same membrane parameters as the microcircuits. The excitatory population had 75 neurons, while the inhibitory population had 25 neurons. The parameters for connection probability and synaptic weights are given in Table S4.

Each excitatory neuron was given a random 500 ms pulse with an amplitude drawn from an normal distribution with mean $\mu = 3$ and variance $\sigma^2 = 3$ in addition to an underlying bias current of -5 nA. After the pulse ends, we waited for an extra 1 s before giving a reset signal of -20 nA to the excitatory neurons. Through this protocol, if the network is capable of sustaining memory, then it will settle into a random equilibrium point after the positive pulse is removed, and before the reset signal arrives. Three different types of entropy are calculated to analyze the variety of memory states in the network.

Entropy

For a discrete random variable X, where there are n potential outcomes $x_i, ...x_n$ with associated probability $P(x_1), ...P(x_n)$, its entropy is

$$H(X) = -\Sigma_{i=1}^{n} P(x_i) log P(x_i) \tag{5}$$

Entropy of words, H_{word}

After the pulse is turned off and before the reset signal arrives, each neuron i will have a different time-averaged firing rate $\bar{r_i} = \langle r_i \rangle_t$, which represents the "state" the neuron is in for that particular equilibrium point. Together, these firing rates can form a sequence, i.e. $r_1, r_2, r_3, ...r_100$, which we call a "word". Each trial will produce one such word. If the memory network has multiple equilibrium points, then the variation of words would be large, and the entropy resulting from such a distribution of words, H_{word}, would be large.

When calculating the entropy, what we actually used is the quotient of $\bar{r_i}$ over 20. This is similar to grouping the firing rate into different bins, where the firing rates within the same bin is assumed to carry the same amount of information. The choice of having a 20 Hz bin is arbitrary – qualitatively similar results can be obtained by other bins, such as 30 Hz and 40 Hz as well.

Entropy of symbols, H_{symbol}

A word $r_1, r_2, r_3, ...r_{100}$ itself is a distribution of firing rates. The entropy of this distribution is called H_{symbol}, which reflects how widely distributed these firing rates are. For the word produced by each trial, its entropy of symbols is calculated, and then averaged across trials.

Entropy of binary symbols, H_{binary}

Instead of representing each neuron by its firing rate, we instead observe whether the neuron fires or not. If it fires, the neuron is given a label $l_i = 1$, and $l_i = 0$ otherwise. This again forms a distribution across the neuron population, $l_1, l_2, l_3, ...l_{100}$, and its associated entropy is called H_{binary}.

Acknowledgements

We thank Gabrielle Gutierrez and for discussion and feedback on the figures, Youngmin Park for useful discussions on the mathematical proofs presented in the study, and Fred Rieke and Peter J. Thomas for feedback on the manuscript. This work is partially supported by the Featured Areas Research Center Program within the framework of the Higher Education Sprout Project, a join fund from the Ministry of Education (MOE) and the

Ministry of Science and Technology (MOST) in Taiwan. This work was also partially supported by Ministry of Science and Technology grant number 109-2218-E-007-021- in Taiwan.

References

[1] Yang, G. R., Joglekar, M. R., Song, H. F., Newsome, W. T. & Wang, X.-J. Task representations in neural networks trained to perform many cognitive tasks. *Nature Neuroscience* **22**, 297–306 (2019). URL https://www.nature.com/articles/s41593-018-0310-2.

[2] Morton, D. W. & Chiel, H. J. Neural architectures for adaptive behavior. *Trends in Neurosciences* **17**, 413–420 (1994). URL https://www.sciencedirect.com/science/article/pii/0166223694900159.

[3] Machens, C. K., Romo, R. & Brody, C. D. Flexible Control of Mutual Inhibition: A Neural Model of Two-Interval Discrimination. *Science* **307**, 1121–1124 (2005). URL https://science.sciencemag.org/content/307/5712/1121.

[4] Lyttle, D. N., Gill, J. P., Shaw, K. M., Thomas, P. J. & Chiel, H. J. Robustness, flexibility, and sensitivity in a multifunctional motor control model. *Biological Cybernetics* **111**, 25–47 (2017). URL https://doi.org/10.1007/s00422-016-0704-8.

[5] Shenoy, K. V., Kaufman, M. T., Sahani, M. & Churchland, M. M. A dynamical systems view of motor preparation: Implications for neural prosthetic system design. *Progress in brain research* **192**, 33–58 (2011). URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3665515/.

[6] Wang, X.-J. Decision Making in Recurrent Neuronal Circuits. *Neuron* **60**, 215–234 (2008). URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2710297/.

[7] Destexhe, A. & Marder, E. Plasticity in single neuron and circuit computations. *Nature* **431**, 789–795 (2004). URL https://www.nature.com/articles/nature03011.

[8] Leonard, J. L. Network Architectures and Circuit Function: Testing Alternative Hypotheses in Multifunctional Networks. *Brain, Behavior and Evolution* **55**, 248–255 (2000). URL https://www.karger.com/Article/FullText/6659.

[9] Marder, E. & Thirumalai, V. Cellular, synaptic and network effects of neuromodulation. *Neural Netw* **15**, 479–493 (2002).

[10] Harris-Warrick, R. M. & Marder, E. Modulation of neural networks for behavior. *Annual Review of Neuroscience* **14**, 39–57 (1991).

[11] Bargmann, C. I. Beyond the connectome: how neuromodulators shape neural circuits. *Bioessays* **34**, 458–465 (2012).

[12] Hoppensteadt, F. C. & Izhikevich, E. M. *Weakly Connected Neural Networks*. Applied Mathematical Sciences (Springer-Verlag, New York, 1997). URL `https://www.springer.com/gp/book/9780387949482`.

[13] Beer, R. D. On the Dynamics of Small Continuous-Time Recurrent Neural Networks. *Adaptive Behavior* (1995). URL `https://journals.sagepub.com/doi/10.1177/105971239500300405`.

[14] Song, H. F., Yang, G. R. & Wang, X.-J. Training Excitatory-Inhibitory Recurrent Neural Networks for Cognitive Tasks: A Simple and Flexible Framework. *PLOS Computational Biology* **12**, e1004792 (2016). URL `https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004792`.

[15] Webster-Wood, V. A., Gill, J. P., Thomas, P. J. & Chiel, H. J. Control for multifunctionality: bioinspired control based on feeding in Aplysia californica. *Biological Cybernetics* **114**, 557–588 (2020). URL `https://doi.org/10.1007/s00422-020-00851-9`.

[16] Neustadter, D. M., Herman, R. L., Drushel, R. F., Chestek, D. W. & Chiel, H. J. The kinematics of multifunctionality: comparisons of biting and swallowing in Aplysia californica. *Journal of Experimental Biology* **210**, 238–260 (2007). URL `https://jeb.biologists.org/content/210/2/238`.

[17] Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Ryu, S. I. & Shenoy, K. V. Cortical preparatory activity: representation of movement or first cog in a dynamical machine? *Neuron* **68**, 387–400 (2010).

[18] Marder, E. & Taylor, A. L. Multiple models to capture the variability in biological neurons and networks. *Nat. Neurosci.* **14**, 133–138 (2011).

[19] Shew, W. L., Yang, H., Yu, S., Roy, R. & Plenz, D. Information Capacity and Transmission Are Maximized in Balanced Cortical Networks with Neuronal Avalanches. *Journal of Neuroscience* **31**, 55–63 (2011). URL `https://www.jneurosci.org/content/31/1/55`.

[20] Kim, R. & Sejnowski, T. J. Strong inhibitory signaling underlies stable temporal dynamics and working memory in spiking neural networks. *Nature Neuroscience* **24**, 129–139 (2021). URL `https://www.nature.com/articles/s41593-020-00753-w`.

[21] Hennequin, G., Vogels, T. & Gerstner, W. Optimal Control of Transient Dynamics in Balanced Networks Supports Generation of Complex Movements. *Neuron* **82**, 1394–1406 (2014). URL http://www.sciencedirect.com/science/article/pii/S0896627314003602.

[22] Litwin-Kumar, A. & Doiron, B. Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nature Neuroscience* **15**, 1498–1505 (2012). URL https://www.nature.com/articles/nn.3220.

[23] Peng, Y., Barreda Tomás, F. J., Klisch, C., Vida, I. & Geiger, J. R. P. Layer-specific organization of local excitatory and inhibitory synaptic connectivity in the rat presubiculum. *Cerebral Cortex* **27**, 2435–2452 (2017). URL https://academic.oup.com/cercor/article/27/4/2435/3052756.

[24] Seeman, S. C. *et al.* Sparse recurrent excitatory connectivity in the microcircuit of the adult mouse and human cortex. *eLife* **7**, e37349 (2018). URL https://doi.org/10.7554/eLife.37349.

[25] Holmgren, C., Harkany, T., Svennenfors, B. & Zilberter, Y. Pyramidal cell communication within local networks in layer 2/3 of rat neocortex. *The Journal of Physiology* **551**, 139–153 (2003).

[26] Douglas, R. J., Koch, C., Mahowald, M., Martin, K. A. & Suarez, H. H. Recurrent excitation in neocortical circuits. *Science* **269**, 981–985 (1995). URL https://science.sciencemag.org/content/269/5226/981.

[27] Yu, J., Hu, H., Agmon, A. & Svoboda, K. Recruitment of GABAergic interneurons in the barrel cortex during active tactile behavior. *Neuron* **104**, 412–427.e4 (2019). URL http://www.sciencedirect.com/science/article/pii/S0896627319306531.

[28] Durstewitz, D., Seamans, J. K. & Sejnowski, T. J. Neurocomputational models of working memory. *Nature Neuroscience* **3**, 1184–1191 (2000). URL https://www.nature.com/articles/nn1100$_1$184.

[29] Ferguson, K. A. & Cardin, J. A. Mechanisms underlying gain modulation in the cortex. *Nature Reviews Neuroscience* **21**, 80–92 (2020). URL https://www.nature.com/articles/s41583-019-0253-y.

[30] Herstel, L. J. & Wierenga, C. J. Network control through coordinated inhibition. *Current Opinion in Neurobiology* **67**, 34–41 (2021). URL http://www.sciencedirect.com/science/article/pii/S0959438820301124.

[31] Sadeh, S. & Clopath, C. Inhibitory stabilization and cortical computation. *Nature Reviews Neuroscience* **22**, 21–37 (2021). URL https://www.nature.com/articles/s41583-020-00390-z.

[32] Kogo, N. *et al.* Dynamics of a mutual inhibition between pyramidal neurons compared to human perceptual competition. *bioRxiv* 2020.05.26.113324 (2020). URL https://www.biorxiv.org/content/10.1101/2020.05.26.113324v2.

[33] Koyama, M. & Pujala, A. Mutual inhibition of lateral inhibition: a network motif for an elementary computation in the brain. *Current Opinion in Neurobiology* **49**, 69–74 (2018). URL http://www.sciencedirect.com/science/article/pii/S095943881730291X.

[34] Espinoza, C., Guzman, S. J., Zhang, X. & Jonas, P. Parvalbumin + interneurons obey unique connectivity rules and establish a powerful lateral-inhibition microcircuit in dentate gyrus. *Nature Communications* **9**, 4605 (2018). URL https://www.nature.com/articles/s41467-018-06899-3.

[35] Haroush, N. & Marom, S. Inhibition increases response variability and reduces stimulus discrimination in random networks of cortical neurons. *Scientific Reports* **9**, 4969 (2019). URL https://www.nature.com/articles/s41598-019-41220-2.

[36] Buzsáki, G. & Wang, X.-J. Mechanisms of Gamma Oscillations. *Annual Review of Neuroscience* **35**, 203–225 (2012). URL https://www.annualreviews.org/doi/10.1146/annurev-neuro-062111-150444.

[37] Ben-Ari, Y. Excitatory actions of gaba during development: the nature of the nurture. *Nature Reviews Neuroscience* **3**, 728–739 (2002). URL https://www.nature.com/articles/nrn920.

[38] Feller, M. B. Spontaneous Correlated Activity in Developing Neural Circuits. *Neuron* **22**, 653–656 (1999). URL http://www.sciencedirect.com/science/article/pii/S0896627300807242.

[39] Huang, Y.-C. *et al.* A Single-Cell Level and Connectome-Derived Computational Model of the Drosophila Brain. *Frontiers* (2018).

[40] Mongillo, G., Rumpel, S. & Loewenstein, Y. Inhibitory connectivity defines the realm of excitatory plasticity. *Nature Neuroscience* **21**, 1463–1470 (2018). URL https://www.nature.com/articles/s41593-018-0226-x.

[41] Wang, C.-T., Lee, C.-T., Wang, X.-J. & Lo, C.-C. Top-down modulation on perceptual decision with balanced inhibition through feedforward and feedback inhibitory neurons. *PLoS ONE* **8**, e62379 (2013).

[42] Lo, C.-C. & Wang, X.-J. Cortico–basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nature Neuroscience* **9**, 956–963 (2006). URL https://www.nature.com/articles/nn1722.

[43] Wong, K.-F. & Wang, X.-J. A Recurrent Network Mechanism of Time Integration in Perceptual Decisions. *J. Neurosci.* **26**, 1314–1328 (2006). URL http://www.jneurosci.org/content/26/4/1314.

[44] Marder, E., Goeritz, M. L. & Otopalik, A. G. Robust circuit rhythms in small circuits arise from variable circuit components and mechanisms. *Curr Opin Neurobiol* **31**, 156–163 (2015). URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4375070/.

[45] Hopfield, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences* **79**, 2554–2558 (1982). URL https://www.pnas.org/content/79/8/2554.

[46] Couey, J. J. *et al.* Recurrent inhibitory circuitry as a mechanism for grid formation. *Nat. Neurosci.* **16**, 318–324 (2013).

[47] Chen, B. & Miller, P. Attractor-state itinerancy in neural circuits with synaptic depression. *The Journal of Mathematical Neuroscience* **10**, 15 (2020). URL https://doi.org/10.1186/s13408-020-00093-w.

[48] Izhikevich, E. M. *Dynamical Systems in Neuroscience* (MIT Press, 2007).

[49] Buendía, V., di Santo, S., Villegas, P., Burioni, R. & Muñoz, M. A. Self-organized bistability and its possible relevance for brain dynamics. *Physical Review Research* **2**, 013318 (2020). URL https://link.aps.org/doi/10.1103/PhysRevResearch.2.013318.

[50] Toups, J. V., Fellous, J.-M., Thomas, P. J., Sejnowski, T. J. & Tiesinga, P. H. Multiple Spike Time Patterns Occur at Bifurcation Points of Membrane Potential Dynamics. *PLOS Computational Biology* **8**, e1002615 (2012). URL https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1002615.

[51] Najafi, F. *et al.* Excitatory and Inhibitory Subnetworks Are Equally Selective during Decision-Making and Emerge Simultaneously during Learning. *Neuron* **105**, 165–179.e8 (2020). URL http://www.sciencedirect.com/science/article/pii/S0896627319308487.

[52] Chien, V. S. C., Maess, B. & Knösche, T. R. A generic deviance detection principle for cortical On/Off responses, omission response, and mismatch negativity. *Biol Cybern* **113**, 475–494 (2019). URL https://doi.org/10.1007/s00422-019-00804-x.

[53] Wilson, H. R. & Cowan, J. D. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* **13**, 55–80 (1973). URL https://doi.org/10.1007/BF00288786.

[54] Wang, X.-J. Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological reviews* **90**, 1195–1268 (2010). URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2923921/.

[55] Shih, C.-T. *et al.* Connectomics-based analysis of information flow in the Drosophila brain. *Current biology: CB* **25**, 1249–1258 (2015).

[56] Shih, C.-T. *et al.* Diverse Community Structures in the Neuronal-Level Connectome of the Drosophila Brain. *Neuroinformatics* **18**, 267–281 (2020). URL https://doi.org/10.1007/s12021-019-09443-w.

[57] Chiang, A.-S. *et al.* Three-Dimensional Reconstruction of Brain-wide Wiring Networks in Drosophila at Single-Cell Resolution. *Current Biology* **21**, 1–11 (2011). URL https://www.sciencedirect.com/science/article/pii/S0960982210015228.

[58] Lau, T., Gage, G. J., Berke, J. D. & Zochowski, M. Local dynamics of gap-junction-coupled interneuron networks. *Physical Biology* **7**, 16015 (2010).

[59] Miller, P., Brody, C. D., Romo, R. & Wang, X.-J. A recurrent network model of somatosensory parametric working memory in the prefrontal cortex. *Cereb. Cortex* **13**, 1208–1218 (2003).

[60] Syakur, M., Khotimah, B., Rohman, E. & Dwi Satoto, B. Integration K-Means Clustering Method and Elbow Method For Identification of The Best Customer Profile Cluster. *IOP Conference Series: Materials Science and Engineering* **336**, 012017 (2018).