

Article

Dynamical and Coupling Structure of Pulse-Coupled Networks in Maximum Entropy Analysis

Zhi-Qin John Xu ^{1,*} , Douglas Zhou ^{2,*} and David Cai ^{1,2,3}¹ NYUAD Institute, New York University Abu Dhabi, Abu Dhabi 129188, UAE; cai@cims.nyu.edu² School of Mathematical Sciences, MOE-LSC and Institute of Natural Sciences, Shanghai Jiao Tong University, Shanghai 200240, China³ Courant Institute of Mathematical Sciences and Center for Neural Science, New York University, New York, NY 10012, USA

* Correspondence: zhiqinxu@nyu.edu (Z.-Q.J.X.); zdz@sjtu.edu.cn (D.Z.)

Received: 5 November 2018; Accepted: 9 January 2019; Published: 16 January 2019

Abstract: Maximum entropy principle (MEP) analysis with few non-zero effective interactions successfully characterizes the distribution of dynamical states of pulse-coupled networks in many fields, e.g., in neuroscience. To better understand the underlying mechanism, we found a relation between the dynamical structure, i.e., effective interactions in MEP analysis, and the anatomical coupling structure of pulse-coupled networks and it helps to understand how a sparse coupling structure could lead to a sparse coding by effective interactions. This relation quantitatively displays how the dynamical structure is closely related to the anatomical coupling structure.

Keywords: maximum entropy; pulse-coupled network; neural network; coupling structure; sparse coding

1. Introduction

Binary-state networks—each node in one sampling time bin is binary-state—arise from many research fields, e.g., gene regulatory modeling and neural dynamics [1–3]. Statistical distributions of network states are essential to encode information [4–8]. For example, with statistical distributions of network states, experimental studies show that rats can perform awake replays of remote experiences in the hippocampus [9]. The number of network states of n binary-state nodes, 2^n , exponentially grows as the node number, n , increases, which creates a challenge of characterizing the probability distribution of network states. Many works effectively characterize the distribution of network states in various systems, e.g., a network of ~ 100 neurons [10], with a *low-order maximum entropy principle* (MEP) analysis [10–18]—a method with few (far less than 2^n) non-zero effective interactions (see a precise definition in Equation (1)) constrained by low-order statistics. In the MEP analysis, the *dynamical structure* of the network, i.e., how nodes interact with each other in the recording of dynamical data, is characterized by effective interactions. This dynamical structure has been used to study the functional connectivity of networks [16,19]. For example, experimental studies show that the second-order effective interaction map of the retina is sparse and dominated by local overlapping effective interaction modules [19]. In this sense, those effective interactions can be regarded as a sparse coding of the information that encoded in the state distribution. Here, the sparseness is defined as the ratio of the number of non-zero effective interactions to 2^n (the number of total effective interactions). Several studies show that high-order effective interactions could be important for characterizing observed distributions of network states [6,10,20]. Although high-order effective interactions are required, the number of required non-zero high-order effective interactions are very small in these experimental studies [10]. What leads to the sparsity of effective interactions remains to be clarified.

We address how a sparse anatomical coupling structure (In the following, for simplicity, we use “coupling structure” instead of “anatomical coupling structure”.) could lead to a sparse coding by effective interactions. Network dynamical structure often closely relates to the underlying coupling structure [21]. For example, when the input of each node is independent from others, (i) high-order (≥ 2) effective interactions are zero in a network of no connections, and (ii) high-order effective interactions are large in a dense and strong connected excitatory network. To efficiently encode information, a realistic system often incorporates a coupling structure with certain features [22,23], e.g., sparsity, small-world, or scale-free. However, it is still unclear how the coupling structure affects the dynamical structure of effective interactions.

In this work, we consider a general class of pulse-coupled networks. The state of each node is binary-state, i.e., active when the node sends pulses to its child nodes, otherwise, silent. We first establish the connection between the coupling structure and the number of non-zero effective interactions in the full-order MEP analysis (constrained by all moments) through an observed *Fact*—which is independent from node dynamics. We then examine the observed *Fact* by numerical simulations. Through our analysis, we can estimate the number of non-zero effective interactions for a given coupling structure when the external input of each node is independent from one another. Our results show that a sparse network could lead to a lot of vanishing high-order effective interactions. For illustration, we verify our results by estimating the number of non-zero effective interactions for each order in a network with *Erdos–Renyi* connection structure, in which our estimated number is much smaller than C_n^k , the number of all possible k th-order effective interactions. Our results establish a connection between the dynamical structure and the network coupling structure. This connection provides an insight into how a sparse coupling structure can lead to a sparse coding scheme. In this work, we would mainly use neural networks as examples for illustration, while our results apply to general binary-state networks.

2. Results

In the following analysis, we use binary vector $V(l) = (\sigma_1, \dots, \sigma_n) \in \{0, 1\}^n$ to represent the state of n nodes within the sampling time bin labeled by l . To obtain correlations up to the m th-order requires to evaluate all $\langle \sigma_{i_1} \dots \sigma_{i_M} \rangle_E$, where $1 \leq i_1 < i_2 < \dots < i_M \leq n$, $1 \leq M \leq m$, and $\langle \cdot \rangle_E$ is defined by $\langle g(l) \rangle_E = \sum_{l=1}^{N_T} g(l) / N_T$ for any function $g(l)$ and N_T is the total number of sampling time bins in the recording. The m th-order MEP analysis is to find the desired probability distribution $P(V)$ for n nodes by maximizing the entropy $S \equiv -\sum_V P(V) \log P(V)$ subject to correlations up to the m th-order ($m \leq n$). To solve this optimization problem [24], one introduces Lagrangian multipliers for each constraint, that is, $-J_{i_1 \dots i_k}$ for the constraint of $\langle \sigma_{i_1} \dots \sigma_{i_k} \rangle_E$ and $(Z - 1)$ for the constraint of $\sum_V P(V) = 1$. The optimization problem is

$$P_m(V) = \arg \max_{P(V)} - \sum_V P(V) \log P(V) - \sum_{k=1}^m \sum_{i_1 < \dots < i_k} F_{i_1 \dots i_k} - F_0,$$

where

$$F_0 = (Z - 1) \left(\sum_V P(V) - 1 \right),$$

$$F_{i_1 \dots i_k} = -J_{i_1 \dots i_k} \left(\sum_V P(V) \prod_{j=1}^k \sigma_{i_j}(V) - \langle \sigma_{i_1} \dots \sigma_{i_k} \rangle_E \right),$$

and $\sigma_{i_j}(V)$ is the state of the i_j th node in the network state V . Since the entropy S is a convex function of $P(V)$, the unique distribution can be solved by taking a derivative of the above objective function with respect to each $P(V)$ as follows:

$$P_m(V) = \frac{1}{Z} \exp\left(\sum_{k=1}^m \sum_{i_1 < \dots < i_k} J_{i_1 \dots i_k} \sigma_{i_1} \dots \sigma_{i_k}\right), \tag{1}$$

where, following the terminology of statistical physics, we call $J_{i_1 \dots i_k}$ a k th-order effective interaction ($1 \leq k \leq m$), the partition function Z is the normalization factor. Equation (1) is referred to as the m th-order MEP distribution. When $m < n$, the constants $J_{i_1 \dots i_k}$ can be determined by all constraints through a commonly used iteration method [13]. When $m = n$, all order moments of the experimentally observed distribution are used as constraints; thus, the above optimization problem has only one feasible solution, that is, the experimentally observed distribution [25,26].

First, we discuss the relationship between effective interactions and the statistical distribution of network states. By taking the logarithm of both sides of Equation (1) for $P_n(V)$, we can get a set of linear equations of all-order effective interactions for all states V . Since P_n is the same as the experimentally observed distribution [25], we can obtain the effective interactions in P_n in terms of the experimentally observed distribution [27]. For example, $n = 3$, we can obtain $J_1 = \log(P_{100}/P_{000})$ and $J_{12} = \log(P_{110}/P_{010}) - J_1$, where $P_{\sigma_1 \sigma_2 \sigma_3}$ represents the probability of the network state $(\sigma_1, \sigma_2, \sigma_3)$. By applying $P(\sigma_1, \sigma_2, \sigma_3) = P(\sigma_1|\sigma_2, \sigma_3)P(\sigma_2, \sigma_3)$, we have $J_1 = \log \frac{P(\sigma_1=1|\sigma_2=0, \sigma_3=0)}{P(\sigma_1=0|\sigma_2=0, \sigma_3=0)}$ and $J_{12} = \log \frac{P(\sigma_1=1|\sigma_2=1, \sigma_3=0)}{P(\sigma_1=0|\sigma_2=1, \sigma_3=0)} - J_1 \triangleq J_1^1 - J_1$. Our earlier study has shown a recursive structure among effective interactions, that is, the $(k + 1)$ st-order effective interaction $J_{123 \dots (k+1)}$ can be obtained as follows [27]: first, we switch the state of the $(k + 1)$ st node in $J_{123 \dots k}$ from silent to active to obtain a new term $J_{123 \dots k}^1$ e.g., from J_1 to J_1^1 ; then, we subtract $J_{123 \dots k}$ from the new term to obtain $J_{123 \dots (k+1)}$, i.e.,

$$J_{123 \dots (k+1)} = J_{123 \dots k}^1 - J_{123 \dots k}. \tag{2}$$

Without loss of generality, we randomly select two nodes labeled by 1 and 2. By the recursive relation, any k th-order effective interaction that includes node 1 and 2 can be expressed as the summation of terms with the following *basic form*:

$$J_{12}^b(\sigma_3, \dots, \sigma_n) = \log \frac{P(\sigma_1=1|\sigma_2=1, \sigma_3, \dots, \sigma_n)}{P(\sigma_1=0|\sigma_2=1, \sigma_3, \dots, \sigma_n)} - \log \frac{P(\sigma_1=1|\sigma_2=0, \sigma_3, \dots, \sigma_n)}{P(\sigma_1=0|\sigma_2=0, \sigma_3, \dots, \sigma_n)}. \tag{3}$$

Note that we use superscript b to denote “basic”. For any i and j ($i \neq j$), the states of node l ($l \neq i, j$) in J_{ij} is silent. J_{ij}^b is a function of states of nodes σ_l 's ($l \neq i, j$) by replacing the silent state of every node l ($l \neq i, j$) in J_{ij} with variable $\sigma_l \in \{0, 1\}$. For example, $J_{123} = J_{12}^b(1, 0, \dots, 0) - J_{12}^b(0, 0, \dots, 0)$ and $J_{1234} = [J_{12}^b(1, 1, 0, \dots, 0) - J_{12}^b(0, 1, 0, \dots, 0)] - J_{123}$. We can observe that if nodes 1 and 2 are independent conditioned on all other nodes, i.e., $P(\sigma_1|\sigma_2 = 1, \sigma_3, \dots, \sigma_n) = P(\sigma_1|\sigma_2 = 0, \sigma_3, \dots, \sigma_n)$, any effective interaction containing these two nodes is zero.

Next, we would show what kind of coupling structure could entail the conditional independence of two nodes. Here, we define some notations. In any sampling time bin $[0, \Delta)$ with state $V = (\sigma_1, \dots, \sigma_n)$, $\forall t \in [0, \Delta)$, we denote $I_{i,t}$ as node i 's input from the outside of the network, denote $w_{ij}(t)$ as the input from the node i to node j , denote $C(i)$ as the set of all child nodes of node i , denote $U_i = C(i) \cup \{i\}$, denote $P(e)$ as the probability of event e , denote $U_0 = \{1, 2, \dots, n\}$.

Fact 1. For n pulse-coupled nodes with binary-state dynamics on a network with a coupling structure G_0 , in any sampling time bin $[0, \Delta)$, $\forall t \in [0, \Delta)$, $\forall i_1, j_1 \in U_0$, we assume that: (a) the external inputs of each node are independent from one another, i.e., $P(I_{i_1,t}, I_{j_1,t}) = P(I_{i_1,t})P(I_{j_1,t})$; (b) whether a parent node sends spikes to its child nodes only depends on its state, i.e., $P(w_{i_1 j_1}(t), V) = W(\sigma_{i_1}, i_1, j_1, t)$, where $W(\cdot, \cdot, \cdot, \cdot)$ is a real function. $\forall i, j \in U_0$, if they neither are connected nor share any common child node, i.e., $U_i \cap U_j = \emptyset$, then, node i and j are independent conditioned on the state of all other nodes, i.e.,

$$P(\sigma_i, \sigma_j|H) = P(\sigma_i|H)P(\sigma_j|H), \tag{4}$$

where H is a possible state of nodes in $U_0 \setminus \{i, j\}$.

We justify our two assumptions as follows: to avoid the influence of correlation in external inputs when we are studying the relation between the dynamical structure and the coupling structure, we assume that the external input of each node is independent from one another, i.e., assumption (a). The second assumption implicates a Markov-like property; that is, for a connected pair of pulse-coupled nodes in an equilibrium state, the pulse from the parent node to the child node only depends on the state of the parent node but is independent from inputs imposed on the parent node. For example, in neural networks, a neuron sends out spikes only when this neuron is active, regardless of what inputs are imposed on the neuron.

The argument for the conclusion in Equation (4) is as follows. By assumption (a), node i and node j can be dependent only through the coupling structure G_0 . When we are considering how node i and node j affect each other by changing their states through the coupling structure G_0 , we can consider a simplified coupling structure, G_1 , which ignores those connections that are independent from states of node i and node j , i.e., σ_i and σ_j . $\forall k \in U_0 \setminus \{i, j\}$, i.e., any other node k , its state σ_k is fixed when we are considering the conditional probability in Equation (4). By assumption (b), for node k 's any child node l , the input from node k to node l is independent from σ_i and σ_j . Thus, the connections started from those nodes in $U_0 \setminus \{i, j\}$ are fixed for different states of σ_i and σ_j . Therefore, G_1 is a simplified coupling structure that only keeps those connections that originated from node i and node j in G_0 . In G_1 , any connection only exists in either sub-network U_i or sub-network U_j . Under the condition $U_i \cap U_j = \phi$, i.e., they neither are connected nor share any common child node, sub-network U_i and sub-network U_j are two isolated sub-networks. σ_i and σ_j cannot affect each other by changing their states through the coupling structure G_1 , that is, node i and j are independent conditioned on the states of all other nodes.

Figure 1 displays an example to illustrate our observed Fact. The coupling structure G_0 is shown in Figure 1a. We focus on node 1 and node 2, where they neither are connected nor share any child node. When the state of other nodes (black) are fixed, all outputs from black nodes can be ignored in the simplified coupling structure G_1 , as shown in Figure 1b. Node 1 and node 2 respectively belong to two separate sub-networks. Therefore, nodes 1 and 2 are independent conditioned on the state of all other nodes.

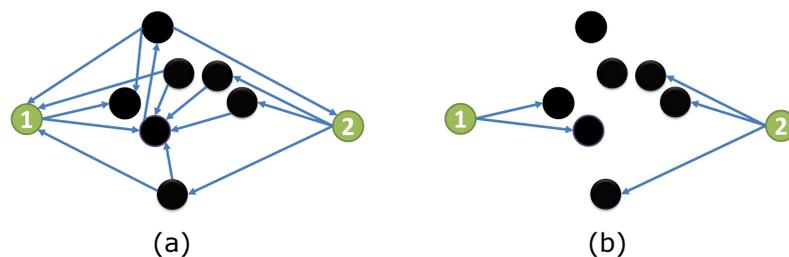


Figure 1. Structure vs. simplified structure.

Based on the recursive structure of effective interactions and the observed Fact, we reach the following conclusion: with the two assumptions in the observed Fact, for a group of nodes $\{i_1, i_2, \dots, i_k\}$, if there exists at least one pair of nodes that neither are connected nor share any child node, effective interaction J_{i_1, i_2, \dots, i_k} is zero.

The system we would use to examine our conclusion is an integrate-and-fire (I&F) network, a general pulse-coupled network, with both excitatory and inhibitory nodes [21]. For the i th node, the dynamics of its state variable x_i with time scales τ are governed by

$$\dot{x}_i = -\frac{x_i}{\tau} - (g_i^{\text{bg}} + g^{\text{ex}})(x_i - x_{\text{ex}}) - g_i^{\text{in}}(x_i - x_{\text{in}}), \tag{5}$$

where x_{ex} and x_{in} are the reversal values of excitation (ex) and inhibition (in), respectively. $g_i^{\text{bg}} = f \sum_k \mathbb{H}(t - T_{i,k}^F) \exp[-(t - T_{i,k}^F)/\sigma^{\text{ex}}]$ is the background input with magnitude f and time scale σ^{ex} , $T_{i,k}^F$ is a Poisson process with rate μ , $\mathbb{H}(\cdot)$ is the Heaviside function, $g_i^{\text{ex}} = \sum_j \sum_k S_{ij}^{\text{ex}} \mathbb{H}(t - T_{j,k}^{\text{ex}}) \exp[-(t - T_{j,k}^{\text{ex}})/\sigma^{\text{ex}}]$ is the excitatory input from other j th excitatory nodes, and $g_i^{\text{in}} = \sum_j \sum_k S_{ij}^{\text{in}} \mathbb{H}(t - T_{j,k}^{\text{in}}) \exp[-(t - T_{j,k}^{\text{in}})/\sigma^{\text{in}}]$ is the inhibitory input from other j th inhibitory nodes. The j th excitatory (inhibitory) node x_j evolves continuously according to Equation (5) until it reaches a firing threshold x_{th} . That moment in time is referred to as a firing event (say, the k th spike) and denoted by $T_{j,k}^{\text{ex}}$ ($T_{j,k}^{\text{in}}$). Then, x_j is reset to the reset value x_r ($x_{\text{in}} < x_r < x_{\text{th}} < x_{\text{ex}}$) and held x_r for an absolute refractory period of τ_{ref} . Each spike emerging from the j th excitatory (inhibitory) node causes an instantaneous increase S_{ij}^{ex} (S_{ij}^{in}) in g_i^{ex} (g_i^{in}), where S_{ij}^{ex} and S_{ij}^{in} are the excitatory and inhibitory coupling strengths, respectively. The model (5) describes a general class of physical networks [1,3,21,28,29]. To be intuitive, the I&F model described in Equation (5) can be understood through a resistor–capacitor circuit. Each neuron is a leaky circuit, which consists of a capacitor with dimensionless capacitance as 1 in parallel with two resistors. x_i is the voltage. \dot{x}_i on the left-hand side is the current which passes through the capacitor. On the right-hand side, $-x_i/\tau$ is the leaky current. The first resistor has a reversal potential x_{ex} with conductance ($g_i^{\text{bg}} + g^{\text{ex}}$). Therefore, the second term on the right-hand side obtained by the Ohm’s law is the current that passes through the first resistor. Note that the conductance is affected by inputs. It is similar for the third term on the right-hand side, i.e., the current that passes through the second resistor.

In the first example, two excitatory and two inhibitory I&F nodes form a ring coupling structure (Figure 2a). For any pair of nodes, say, node i and j , we compute $\Delta_{ij}(H) = |P(\sigma_i = 1 | \sigma_j = 1, H) - P(\sigma_i = 1 | \sigma_j = 0, H)|$, where H is a state vector of other two nodes. By our observed Fact, the conditional independent pairs are (neuron 1, neuron 3) and (neuron 2, neuron 4), and other pairs are categorized as dependent pairs. In Figure 2b, the strengths of $\Delta_{ij}(H)$ of independent pairs (green) are almost two orders of magnitude smaller than those of dependent pairs (red). We then shuffle spike trains of each node. We similarly compute $\Delta_{ij}(H)$ for 10 different pieces of shuffled data. Blue dots and cyan dots in Figure 2b are results of all shuffled data of dependent pairs and independent pairs, respectively. The strength of $\Delta_{ij}(H)$ of independent pairs (green)—computed from the observed data—are within the statistical error of shuffled data. We then solve effective interactions in the full-order MEP analysis P_n for this ring network. As shown in Figure 2c, the effective interaction strengths of independent pairs (J_{24} and J_{13}) are within the statistical error of shuffled results (red). Since every high-order (≥ 3) effective interaction includes at least one independent pair of nodes, as predicted, the strengths of all high-order effective interactions are within the statistical error of shuffled results as shown in Figure 2d.

In the second example in the second row in Figure 2, results are similar in that dependent pairs and independent pairs can be identified through our observed Fact, and the strength of any effective interaction that includes the independent pair of nodes (node 1 and node 3) is within the statistical error of shuffled data. In this example, as shown in Figure 2h, J_{124} is very small, i.e., within the statistical error of shuffled results. However, J_{124} does not include the independent pair of nodes (node 1 and node 3); thus, the theoretical estimation of zero-strength effective interactions misses J_{124} . This example indicates that our theoretical estimation result may give rise to an upper bound of the number of non-zero effective interactions. In contrast, for a network of all excitatory nodes with the same coupling structure as the one in Figure 1e, J_{124} is significantly larger than zero (not shown). Since the estimation of the strength of high-order effective interactions involves the estimation of high dimensional joint probability distribution, a very long recording due to the curse of dimensionality constrains us from examining $\Delta_{ij}(H)$ for a large network.

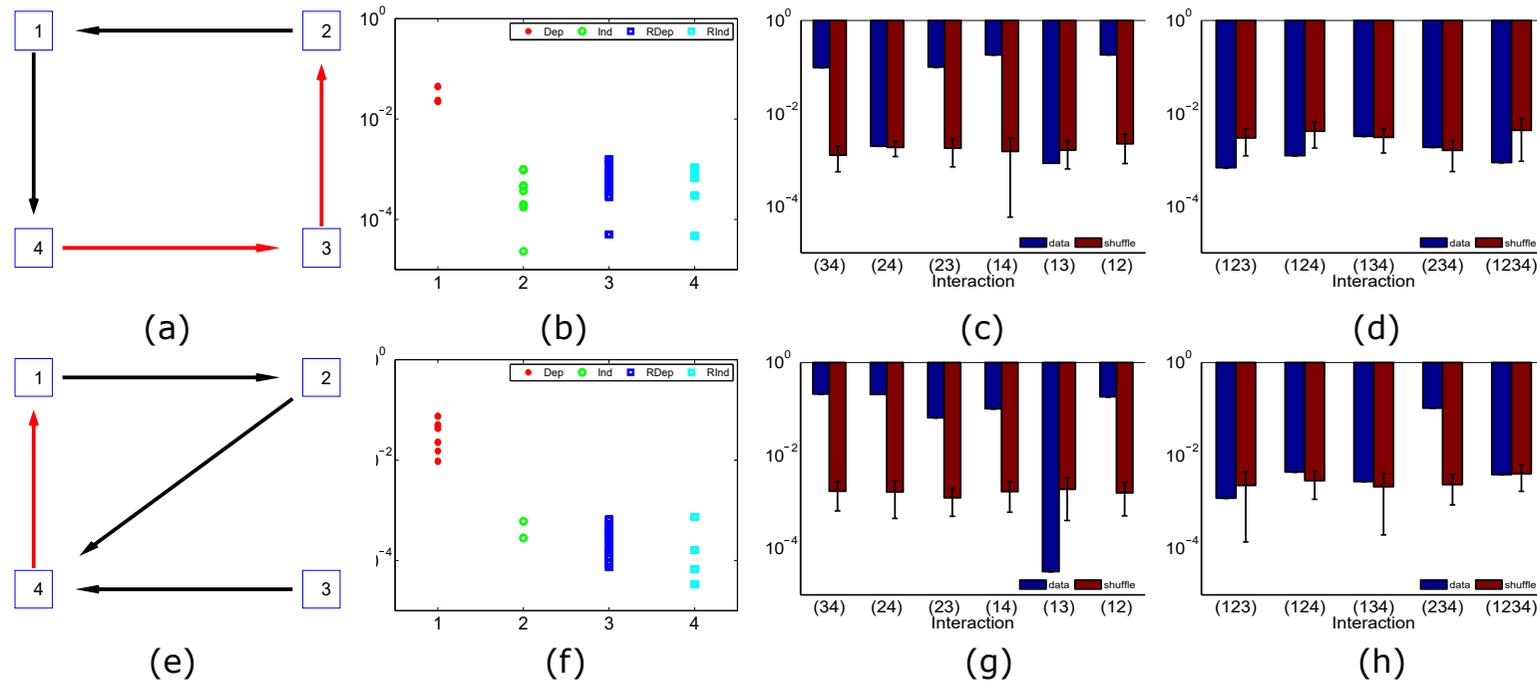


Figure 2. Anatomical structure vs. effective interactions of integrate-and-fire networks. Each row shows a numerical case. In the first column, black arrows and red arrows represent excitatory and inhibitory connections, respectively. In the second column, red and green dots are the strengths of $\Delta_{ij}(H)$ of dependent and independent pairs, respectively. Blue dots and cyan dots are the strengths of $\Delta_{ij}(H)$ of dependent and independent pairs from ten shuffled spike trains, respectively. Each dot is for one $\Delta_{ij}(H)$. The third and fourth columns display absolute effective interaction strengths (blue bars). The corresponding node indexes for each effective interaction are shown in the abscissa. The mean and standard deviation of absolute strengths of each effective interaction of ten shuffled spike trains are also displayed by garnet bars. The simulation time for each network is 1.2×10^8 ms. The time bin size for analysis is 10 ms [12,13]. Independent Poisson inputs for each network are $\mu = 0.1 \text{ ms}^{-1}$ and $f = 0.1 \text{ ms}^{-1}$. The firing rate of each node is about 50 Hz. Parameters are chosen [28] as $x_{\text{ex}} = 14/3$, $x_{\text{in}} = -2/3$, $\sigma^{\text{ex}} = 2 \text{ ms}$, $\sigma^{\text{in}} = 5 \text{ ms}$, $\tau = 20 \text{ ms}$, $x_{\text{th}} = 1$, $x_r = 0$, and $\tau_{\text{ref}} = 2 \text{ ms}$, $S_{ij}^{\text{ex}} = S_{ij}^{\text{in}} = 0.02$.

Based on the relation between the coupling structure and effective interactions, the number of non-zero high-order effective interactions can be small in a sparsely connected network compared with C_n^k , which is the number of all possible k th-order interactions. For example, we estimate the number of each-order non-zero effective interactions in a network with an *Erdos–Renyi* connection structure. We generate 1000 Erdos-Renyi random networks of 100 nodes (the same connection probability but different random samples). In each network, for any pair of node i and node j , we assign the connection from node i to node j by the following rule: we generate a random number from the uniform distribution on $[0, 1]$; if the number is smaller than 0.05, we assign a connection from node i to node j . As shown in Figure 3, the number of non-zero k th-order ($k > 1$) effective interactions is much smaller than C_{100}^k (too large to be shown). The number of high-order effective interactions (order higher than 11th) almost vanishes (order higher than 20th not shown). In this example, the sparseness, i.e., the ratio of the number of non-zero effective interactions to 2^n , is less than 10^{-24} .

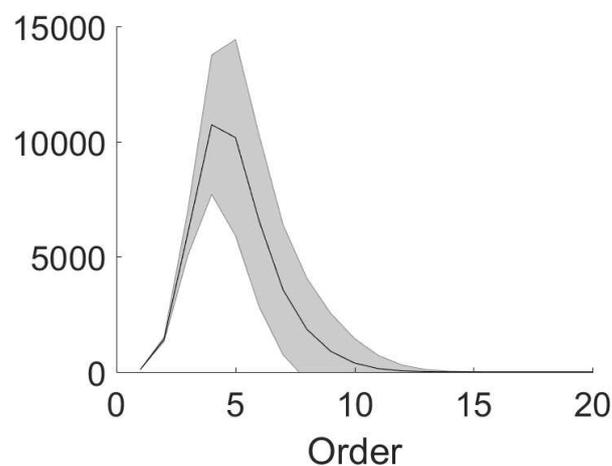


Figure 3. Non-zero effective interactions in the Erdos-Renyi random networks. We generate 1000 Erdos-Renyi random networks of 100 nodes (the same connection probability but different random samples). The connection probability between two nodes is 0.05. The number of non-zero effective interaction is plotted against effective interaction order. The mean and standard deviation are respectively shown by the black line and shaded area.

3. Conclusions and Discussion

In summary, we have established a relation between effective interactions in MEP analysis and the anatomical coupling structure (“coupling structure” for simplicity in the following) of pulse-coupled networks to understand how a sparse coupling structure could lead to a sparse coding by effective interactions. For example, the sparseness of the case in Figure 3 is less than 10^{-24} . Since effective interactions characterize how nodes interact with each other in the recording of dynamical data, i.e., network dynamical structure, our study quantitatively displays how the dynamical structure closely relates to the coupling structure.

Even though high-order effective interactions are often much smaller compared with low-order ones [27], it is still unclear why small high-order effective interactions do not accumulate to have a significant effect in a large network [10,30]. For example, it has been shown that MEP distribution with sparse low-order effective interactions—non-zero effective interactions are sparse and vanish when the order is higher than the eighth-order—can well capture the state distribution of 99 ganglion cells in the salamander retina responding to a natural movie clip or natural pixel [10]. In this study, we show that a large amount of effective interactions vanish in a sparse coupling structure, thus rationalizing the absence of the accumulation of high-order interactions for a large network.

Finally, we point out that some important issues remain to be elucidated in the future. First, we have ignored correlations in external inputs when estimating the number of non-zero effective

interactions. Correlated inputs can induce non-zero high-order effective interactions [31]. It still needs to be considered how the statistics of inputs affect the sparsity of effective interactions. Second, current algorithms for estimating non-zero effective interactions (not limited to the second-order) for a large network (e.g., ~ 100 nodes) are very slow, e.g., Monte Carlo based methods [30,32]. Our undergoing work is developing a fast algorithm that exploits the sparsity of effective interactions. We have seen an indication that the algorithm can work well for an I&F network with sparse coupling structure; however, that work has yet to be fully verified to be conclusive.

Author Contributions: Conceptualization, Z.-Q.J.X., D.Z. and D.C.; Methodology, Z.-Q.J.X.; software, Z.-Q.J.X.; Validation, Z.-Q.J.X., D.Z.; Formal Analysis, Z.-Q.J.X.; Investigation, Z.-Q.J.X.; Resources, Z.-Q.J.X., D.Z., D.C.; Data Curation, Z.-Q.J.X.; Writing—Original Draft Preparation, Z.-Q.J.X.; Writing—Review and Editing, Z.-Q.J.X. and D.Z.; Visualization, Z.-Q.J.X. and D.Z.; Supervision, D.Z. and D.C.; Project Administration, D.Z. and D.C.; Funding Acquisition, D.Z. and D.C.

Funding: This work was supported by National Science Foundation in China with Grant Nos. 11671259, 11722107, and 91630208 (D.Z.); by NSFC-31571071 (D.C.); by SJTU-UM Collaborative Research Program (D.C. and D.Z.); and by the NYU Abu Dhabi Institute G1301 (Z.-Q.J.X., D.Z., and D.C.).

Acknowledgments: The authors thank David W. McLaughlin for helpful discussions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mirollo, R.E.; Strogatz, S.H. Synchronization of pulse-coupled biological oscillators. *SIAM J. Appl. Math.* **1990**, *50*, 1645–1662. [[CrossRef](#)]
2. Stricker, J.; Cookson, S.; Bennett, M.R.; Mather, W.H.; Tsimring, L.S.; Hasty, J. A fast, robust and tunable synthetic gene oscillator. *Nature* **2008**, *456*, 516–519. [[CrossRef](#)] [[PubMed](#)]
3. Wang, Z.; Ma, Y.; Cheng, F.; Yang, L. Review of pulse-coupled neural networks. *Image Vis. Comput.* **2010**, *28*, 5–13. [[CrossRef](#)]
4. Dan, Y.; Alonso, J.M.; Usrey, W.M.; Reid, R.C. Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. *Nat. Neurosci.* **1998**, *1*, 501–507. [[CrossRef](#)] [[PubMed](#)]
5. Vinje, W.E.; Gallant, J.L. Sparse coding and decorrelation in primary visual cortex during natural vision. *Science* **2000**, *287*, 1273–1276. [[CrossRef](#)]
6. Ohiorhenuan, I.E.; Mechler, F.; Purpura, K.P.; Schmid, A.M.; Hu, Q.; Victor, J.D. Sparse coding and high-order correlations in fine-scale cortical networks. *Nature* **2010**, *466*, 617–621. [[CrossRef](#)]
7. Shemesh, Y.; Sztainberg, Y.; Forkosh, O.; Shlapobersky, T.; Chen, A.; Schneidman, E. High-order social interactions in groups of mice. *eLife* **2013**, *2*, e00759. [[CrossRef](#)] [[PubMed](#)]
8. Knill, D.C.; Pouget, A. The Bayesian brain: The role of uncertainty in neural coding and computation. *TRENDS Neurosci.* **2004**, *27*, 712–719. [[CrossRef](#)]
9. Karlsson, M.P.; Frank, L.M. Awake replay of remote experiences in the hippocampus. *Nat. Neurosci.* **2009**, *12*, 913. [[CrossRef](#)]
10. Ganmor, E.; Segev, R.; Schneidman, E. Sparse low-order interaction network underlies a highly correlated and learnable neural population code. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 9679–9684. [[CrossRef](#)]
11. Schneidman, E.; Berry, M.J.; Segev, R.; Bialek, W. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* **2006**, *440*, 1007–1012. [[CrossRef](#)] [[PubMed](#)]
12. Shlens, J.; Field, G.D.; Gauthier, J.L.; Grivich, M.I.; Petrusca, D.; Sher, A.; Litke, A.M.; Chichilnisky, E. The structure of multi-neuron firing patterns in primate retina. *J. Neurosci.* **2006**, *26*, 8254–8266. [[CrossRef](#)] [[PubMed](#)]
13. Tang, A.; Jackson, D.; Hobbs, J.; Chen, W.; Smith, J.L.; Patel, H.; Prieto, A.; Petrusca, D.; Grivich, M.I.; Sher, A.; et al. A maximum entropy model applied to spatial and temporal correlations from cortical networks in vitro. *J. Neurosci.* **2008**, *28*, 505–518. [[CrossRef](#)]
14. Marre, O.; El Boustani, S.; Frégnac, Y.; Destexhe, A. Prediction of spatiotemporal patterns of neural activity from pairwise correlations. *Phys. Rev. Lett.* **2009**, *102*, 138101. [[CrossRef](#)] [[PubMed](#)]
15. Bury, T. Statistical pairwise interaction model of stock market. *Eur. Phys. J. B* **2013**, *86*, 89. [[CrossRef](#)]

16. Watanabe, T.; Hirose, S.; Wada, H.; Imai, Y.; Machida, T.; Shirouzu, I.; Konishi, S.; Miyashita, Y.; Masuda, N. A pairwise maximum entropy model accurately describes resting-state human brain networks. *Nat. Commun.* **2013**, *4*, 1370. [[CrossRef](#)]
17. Barreiro, A.K.; Gjorgjieva, J.; Rieke, F.; Shea-Brown, E. When do microcircuits produce beyond-pairwise correlations? *Front. Comput. Neurosci.* **2014**, *8*, 10. [[CrossRef](#)] [[PubMed](#)]
18. Martin, E.A.; Hlinka, J.; Davidsen, J. Pairwise network information and nonlinear correlations. *Phys. Rev. E* **2016**, *94*, 040301. [[CrossRef](#)]
19. Ganmor, E.; Segev, R.; Schneidman, E. The architecture of functional interaction networks in the retina. *J. Neurosci.* **2011**, *31*, 3044–3054. [[CrossRef](#)]
20. Yu, S.; Yang, H.; Nakahara, H.; Santos, G.S.; Nikolić, D.; Plenz, D. Higher-order interactions characterized in cortical activity. *J. Neurosci.* **2011**, *31*, 17514–17526. [[CrossRef](#)]
21. Zhou, D.; Xiao, Y.; Zhang, Y.; Xu, Z.; Cai, D. Causal and structural connectivity of pulse-coupled nonlinear networks. *Phys. Rev. Lett.* **2013**, *111*, 054102. [[CrossRef](#)]
22. Newman, M.E. The structure and function of complex networks. *SIAM Rev.* **2003**, *45*, 167–256. [[CrossRef](#)]
23. Bullmore, E.; Sporns, O. Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **2009**, *10*, 186–198. [[CrossRef](#)] [[PubMed](#)]
24. Jaynes, E.T. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620. [[CrossRef](#)]
25. Amari, S.I. Information geometry on hierarchy of probability distributions. *IEEE Trans. Inf. Theory* **2001**, *47*, 1701–1711. [[CrossRef](#)]
26. Xu, Z.Q.J.; Crodelle, J.; Zhou, D.; Cai, D. Maximum Entropy Principle Analysis in Network Systems with Short-time Recordings. *arXiv* **2018**, arXiv:1808.10506.
27. Xu, Z.Q.J.; Bi, G.; Zhou, D.; Cai, D. A dynamical state underlying the second order maximum entropy principle in neuronal networks. *Commun. Math. Sci.* **2017**, *15*, 665–692. [[CrossRef](#)]
28. Gerstner, W.; Kistler, W.M. *Spiking Neuron Models: Single Neurons, Populations, Plasticity*; Cambridge University Press: Cambridge, UK, 2002.
29. Cai, D.; Rangan, A.V.; McLaughlin, D.W. Architectural and synaptic mechanisms underlying coherent spontaneous activity in V1. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5868–5873. [[CrossRef](#)]
30. Shlens, J.; Field, G.D.; Gauthier, J.L.; Greschner, M.; Sher, A.; Litke, A.M.; Chichilnisky, E. The structure of large-scale synchronized firing in primate retina. *J. Neurosci.* **2009**, *29*, 5022–5031. [[CrossRef](#)]
31. Macke, J.H.; Opper, M.; Bethge, M. Common input explains higher-order correlations and entropy in a simple model of neural population activity. *Phys. Rev. Lett.* **2011**, *106*, 208102. [[CrossRef](#)]
32. Nasser, H.; Marre, O.; Cessac, B. Spatio-temporal spike train analysis for large scale networks using the maximum entropy principle and Monte Carlo method. *J. Stat. Mech. Theory Exp.* **2013**, *2013*, P03006. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).