Formulation of Pruning Maps with Rhythmic Neural Firing

Feng-Sheng Tsai [1,2], Yi-Li Shih [3], Chin-Tzong Pang [3] and Sheng-Yi Hsu [1,2],*

[1]   Department of Biomedical Imaging and Radiological Science, China Medical University,
      Taichung 40402, Taiwan; fstsai@mail.cmu.edu.tw
[2]   Research Center for Interneural Computing, China Medical University Hospital, Taichung 40447, Taiwan
[3]   Department of Information Management, Yuan Ze University, Chung-Li 32003, Taiwan;
      yilishih@gmail.com (Y.-L.S.); imctpang@saturn.yzu.edu.tw (C.-T.P.)
[*]   Correspondence: syhsu@mail.cmu.edu.tw; Tel.: +886-4-22053366

Abstract: Rhythmic neural firing is thought to underlie the operation of neural function. This triggers the construction of dynamical network models to investigate how the rhythms interact with each other. Recently, an approach concerning neural path pruning has been proposed in a dynamical network system, in which critical neuronal connections are identified and adjusted according to the pruning maps, enabling neurons to produce rhythmic, oscillatory activity in simulation. Here, we construct a sort of homomorphic functions based on different rhythms of neural firing in network dynamics. Armed with the homomorphic functions, the pruning maps can be simply expressed in terms of interactive rhythms of neural firing and allow a concrete analysis of coupling operators to control network dynamics. Such formulation of pruning maps is applied to probe the consolidation of rhythmic patterns between layers of neurons in feedforward neural networks.

Keywords: decirculation process; network dynamics; neural path pruning; pruning maps; rhythmic neural firing

1. Introduction

Rhythms are ubiquitous and have been shown to play an important role in living organisms [1,2]. Lines of research on the roles of rhythms are studied, including synchronous flashing of fireflies [3], pacemaker cells of the heart [4], synchronization of pulse-coupled oscillators [5,6], synchronous neural activity propagating in scale-free networks, random networks, and cortical neural networks [7,8]. The central questions in the above research concern the origin and the control of rhythms, that is, the issue of understanding what causes the rhythms and how the rhythms interact with each other. Especially in neural network dynamics, an important issue is to probe activity-dependent plasticity of neural connections entwined with rhythms [9–11]. Such a kind of plasticity is claimed to be the heart of the organization of behavior [12–15].

Classic models of neural networks have achieved state-of-the-art performances in classification [16–18], object detection [19–21], and instance segmentation [22–24]. One of the key structures involved in network architecture is convolution. Many operators used for edge detection or pattern recognition are convolved with input images to get different types of derivative measurement in the feature-extracting process. Input images are usually structural with geometric features. Thus, convolution can be applied to get features from input images, which are then processed by other parts of network architecture to determine a maximal score for outcome prediction. Here, instead of concerning the geometric feature extraction, our focus is on the neural network models receiving and reactivating with rhythms. Rhythms can be the features of inputs sending to neural

works and also be the features of outputs activating by assemblies of neurons in neural networks. We search for a rule to derive operators from rhythmic firing of neurons. Such operators are acting on the adjustment of neural connections and hence can be convolved with the input rhythm to get features and to propagate features for rhythm formation.

Our point of departure comes from the decirculation process in neural network dynamics [25,26]. The decirculation process describes adaptation of network structure that is crucial for evolutionary neural networks to proceed from one circulating state to another. Two sorts of plasticity operators are derived to modify network structure for decirculation process [27]. One is to measure synchronous neural activity in a circulating state, whereas the other is to measure self-sustaining neural activity in a circulating state. They meet the neurobiological concept called Hebbian synaptic plasticity [12–14]. Such plasticity operation reflects an internal control in neural network dynamics, enabling neurons to produce rhythmic, oscillatory activity [27]. Recently, the concept of decirculation process has been extended to the concept of neural path pruning [28]. There, by setting the beginning and the ending of a desired flow of neural firing states, a pruning map can be conducted to indicate critical neuronal connections which have significant controllability in eliminating the flow of neural firing states. On this basis, a simulation result shows that a mobile robot guided by neural network dynamics can be exploited to change neuronal connections determined critically by neural path pruning with rewards and punishment. The mobile robot may result in a good performance of fault-tolerant behavior of tracking (adaptive matching of sensory inputs to motor outputs) [28].

Neural path pruning paves an alternative way to derive operators convolving with signals and controlling neural network dynamics. This motivates us to formulate pruning maps with rhythmic neural firing, which may induce a framework of feedforward neural networks for rhythm formation.

2. Theoretical Framework

The model description follows. Let $\{0,1\}^n$ be the n-dimensional state space consisting of all vectors $x = (x_1, x_2, ..., x_n)$, with each component x_i being 0 or 1. Consider a dynamical system of n coupled neurons, which is modeled by the equation [25,26]:

$$x(t+1) = H_A(x(t), s(t)), \quad t = 0, 1, \ldots, \tag{1}$$

where $x(t) = (x_1(t), x_2(t), \ldots, x_n(t)) \in \{0,1\}^n$ is a vector of neural firing states at time t, with each component $x_i(t)$ representing the firing state of neuron i at time t; $A = (a_{ij}) \in M_n(\mathbb{R})$ is a coupling weight matrix of the n coupled neurons (be they agents in a recurrent or feedforward neural network), with each entry a_{ij} representing the coupling weight underlying the connection from j to i; $s(t) \subset \{1, 2, \ldots, n\}$ denotes an ensemble of neurons which adjust their states at time t; and $H_A(\cdot, s(t))$ is a transition function whose ith component is defined by

$$[H_A(x, s(t))]_i = \mathbb{1}\left(\sum_{j=1}^{n} a_{ij} x_j - b_i\right) \quad \text{if } i \in s(t), \tag{2}$$

otherwise $[H_A(x, s(t))]_i = x_i$, where $b_i \in \mathbb{R}$ is a firing threshold of neuron i for $i = 1, 2, \ldots, n$ and the function $\mathbb{1}$ is the Heaviside function. Consider a flow $\Omega = [x^0, x^1, \ldots, x^p]$ of neural firing states in $\{0,1\}^n$, where $p > 1$, $x^0, x^1, \ldots, x^p \in \{0,1\}^n$, and $x^0 \neq x^i$ for some $i \in \{1, 2, \ldots, p\}$. Specifically, if $x^0 = x^p$ within Ω, then Ω is said to be a loop of neural firing states in $\{0,1\}^n$. For each $i, j = 1, 2, \ldots, n$, an integer, denoted $c_{ij}(\Omega)$, is assigned according to the rule:

$$c_{ij}(\Omega) = x_j^0(x_i^0 - x_i^1) + x_j^1(x_i^1 - x_i^2) + \cdots + x_j^{p-1}(x_i^{p-1} - x_i^p). \tag{3}$$

The matrix $C(\Omega) = (c_{ij}(\Omega))$ is called the pruning map of Ω. Thus,

$$
C(\Omega) \;=\; \begin{pmatrix}
\sum\limits_{m=1}^{p} x_1^{m-1}(x_1^{m-1} - x_1^m) & \cdots & \sum\limits_{m=1}^{p} x_n^{m-1}(x_1^{m-1} - x_1^m) \\[2mm]
\vdots & \ddots & \vdots \\[2mm]
\sum\limits_{m=1}^{p} x_1^{m-1}(x_n^{m-1} - x_n^m) & \cdots & \sum\limits_{m=1}^{p} x_n^{m-1}(x_n^{m-1} - x_n^m)
\end{pmatrix}. \tag{4}
$$

The pruning map $C(\Omega)$ induces the linear functional $A \longrightarrow \langle A, C(\Omega)\rangle$, which is defined on the Hilbert space $M_n(\mathbb{R})$ of all real $n \times n$ matrices endowed with the Hilbert–Schmidt inner product $\langle \cdot, \cdot \rangle$, where $\langle A, C(\Omega)\rangle = \sum_{i,j} a_{ij} c_{ij}(\Omega)$ for each $A = (a_{ij}) \in M_n(\mathbb{R})$. Let $\langle x, y \rangle$ be the usual inner product of x and y in $\mathbb{R}^n$. With these notions, a basic theorem reveals the determination of whether the neural connections identified by the pruning map are crucial to eliminate the flow of neural firing states in the dynamics of the neural network. It was stated in [27,28] as follows.

Theorem 1. *Let $\Omega = [x^0, x^1, \ldots, x^p]$ be a flow of neural firing states in $\{0,1\}^n$. If $A \in M_n(\mathbb{R})$ and $b \in \mathbb{R}^n$ satisfy*

$$
\langle A, C(\Omega)\rangle > \left\langle b, x^0 - x^p \right\rangle, \tag{5}
$$

then for any initial neural firing state $x(0) \in \{0,1\}^n$ and any updating $s(t) \subset \{1, 2, \ldots, n\}$, $t = 0, 1, \ldots$, the flow $x(t)$ encoded by Equation (1) cannot behave in

$$
x(T) = x^0, x(T+1) = x^1, \ldots, x(T+p) = x^p \tag{6}
$$

for each $T = 0, 1, \ldots$.

Theorem 1 is based on the concept of neural path pruning. It points out a regulatory regime for network formation. To show this, consider a flow $\Omega = [x^0, x^1, \ldots, x^p]$ and a coupling weight matrix $A = (a_{ij}) \in M_n(\mathbb{R})$. If we choose a coupling operator $D = (d_{ij}) \in M_n(\mathbb{R})$ such that $\langle D, C(\Omega)\rangle \geq 0$ (respectively, $\langle D, C(\Omega)\rangle \leq 0$), then

$$
\langle A + D, C(\Omega)\rangle \geq \langle A, C(\Omega)\rangle \tag{7}
$$
$$
(\text{resp.,} \ \langle A + D, C(\Omega)\rangle \leq \langle A, C(\Omega)\rangle). \tag{8}
$$

Since the quantity $\left\langle b, x^0 - x^p \right\rangle$ is determined only by the flow Ω and the threshold b, Theorem 1 coupled with the inequality in Equation (7) (respectively, the inequality in Equation (8)) suggests that, given the dynamical system

$$
x(t+1) = H_{A+D}(x(t), s(t)), \quad t = 0, 1, \ldots, \tag{9}
$$

the change of the coupling weight matrix from A to $A + D$ may enhance (respectively, inhibit) the effect of flow elimination on Ω.

We go further to probe the pruning map with a desired flow of rhythmic neural firing. Each entry in the pruning map may be represented in terms of a quantity related to the firing rhythms of individual neurons. In so doing, we can directly extract information from the pruning map concerning how the rhythms interact with each other. Specifically, for feedforward neural networks, the pruning map provides a recipe to probe the interaction of rhythms between layers, which may reveal a consolidated rule to keep neurons firing in their specific rhythms consistently.

3. Main Results

Assume that neuron i fires in rhythm $r_i \in \mathbb{N}$ for $i = 1, 2, \ldots, n$. Let $p = \mathrm{lcm}(r_1, r_2, \ldots, r_n)$, where $\mathrm{lcm}(r_1, r_2, \ldots, r_n)$ denotes the least common multiple of $r_1, r_2, \ldots, r_n$. For each $i = 1, 2, \ldots, n$, let m_i be a nonnegative integer less than r_i, denoting the initial phase of rhythmic neural firing. Thus, the flow with neurons firing in rhythm $r_1, r_2, \ldots, r_n$ and in phase $m_1, m_2, \ldots, m_n$ can be described as

$$\Omega = [x^0, x^1, \ldots, x^p], \tag{10}$$

where $x_i^m = 1$ if $m = m_i + kr_i$ for $k = 0, 1, \ldots, \lfloor (p - m_i)/r_i \rfloor$ and $i = 1, 2, \ldots, n$; otherwise, $x_i^m = 0$. Here, the notation $\lfloor r \rfloor$ denotes the largest integer less than or equal to $r \in \mathbb{R}$. Theorem 2 represents a matrix decomposition formula of the pruning map underlying rhythmic neural firing.

Theorem 2. *Let $\Omega = [x^0, x^1, \ldots, x^p]$ be a flow of rhythmic neural firing states given in Equation (10). For each $i, j = 1, 2, \ldots, n$, denote by $\mathrm{lcm}(r_i, r_j)$ the least common multiple of r_i and r_j, and $\gcd(r_i, r_j)$ the greatest common divisor of r_i and r_j. Let $[m_i]_{\gcd(r_i, r_j)}$ denote the nonnegative integer less than $\gcd(r_i, r_j)$ such that*

$$[m_i]_{\gcd(r_i, r_j)} \equiv m_i \bmod \gcd(r_i, r_j). \tag{11}$$

Then, the pruning map $C(\Omega)$ can be formulated by $C(\Omega) = \hat{C}(\Omega) + C^\natural(\Omega)$, where $\hat{C}(\Omega) = (\hat{c}_{ij}(\Omega))$ is a symmetric matrix whose non-zero entries are given by

$$\hat{c}_{ij}(\Omega) = \hat{c}_{ji}(\Omega) = -p/\mathrm{lcm}(r_i, r_j) \ \text{ if } \ \gcd(r_i, r_j) = 2 \ \text{ and } \ [m_j + 1]_{\gcd(r_i, r_j)} = [m_i]_{\gcd(r_i, r_j)}, \tag{12}$$

$$\hat{c}_{ij}(\Omega) = \hat{c}_{ji}(\Omega) = p/\mathrm{lcm}(r_i, r_j) \ \text{ if } \ \gcd(r_i, r_j) = 2 \ \text{ and } \ [m_j]_{\gcd(r_i, r_j)} = [m_i]_{\gcd(r_i, r_j)}, \tag{13}$$

and $C^\natural(\Omega) = (c_{ij}^\natural(\Omega))$ is a non-symmetric matrix whose non-zero entries are given by

$$c_{ij}^\natural(\Omega) = -p/\mathrm{lcm}(r_i, r_j) \ \text{ if } \ \gcd(r_i, r_j) > 2 \ \text{ and } \ [m_j + 1]_{\gcd(r_i, r_j)} = [m_i]_{\gcd(r_i, r_j)}, \tag{14}$$

$$c_{ij}^\natural(\Omega) = c_{ji}^\natural(\Omega) = p/\mathrm{lcm}(r_i, r_j) \ \text{ if } \ \gcd(r_i, r_j) > 2 \ \text{ and } \ [m_j]_{\gcd(r_i, r_j)} = [m_i]_{\gcd(r_i, r_j)}. \tag{15}$$

Proof. To prove Theorem 2, note that if $m_j > 0$, then $1 \neq r_j > m_j > 0$ and

$$(\mathrm{lcm}(r_1, r_2, \ldots, r_n) - 1)/r_j \geq (p - m_j)/r_j \geq (\mathrm{lcm}(r_1, r_2, \ldots, r_n) - r_j + 1)/r_j. \tag{16}$$

Thus, we have

$$\lfloor (p - m_j)/r_j \rfloor = \begin{cases} \mathrm{lcm}(r_1, r_2, \ldots, r_n)/r_j - 1 & \text{if } \ m_j \neq 0, \\ \mathrm{lcm}(r_1, r_2, \ldots, r_n)/r_j & \text{if } \ m_j = 0. \end{cases} \tag{17}$$

Denote by $k_j = \mathrm{lcm}(r_1, r_2, \ldots, r_n)/r_j - 1$ for $j = 1, 2, \ldots, n$. Since $x_j^m = 1$ if and only if $m = m_j + kr_j$ for $k = 0, 1, \ldots, \lfloor (p - m_j)/r_j \rfloor$, the ij-entry of the pruning map $C(\Omega)$ is reduced by

$$
\begin{aligned}
c_{ij}(\Omega) &= \sum_{m=1}^{p} x_j^{m-1} \left(x_i^{m-1} - x_i^m \right) \\
&= \sum_{k=0}^{k_j} x_j^{m_j + kr_j} \left(x_i^{m_j + kr_j} - x_i^{m_j + kr_j + 1} \right) \\
&= \sum_{k=0}^{k_j} \left(x_i^{m_j + kr_j} - x_i^{m_j + kr_j + 1} \right).
\end{aligned}
\tag{18}
$$

To compute Equation (18) for each pair of $i, j = 1, 2, \ldots, n$, we estimate the numbers of elements in the intersections

$$\{m_j + kr_j; \ k = 0, 1, \ldots, k_j\} \bigcap \{m_i + kr_i; \ k = 0, 1, \ldots, \lfloor (p - m_i)/r_i \rfloor\} \tag{19}$$

and

$$\{m_j + kr_j + 1; \ k = 0, 1, \ldots, k_j\} \bigcap \{m_i + kr_i; \ k = 0, 1, \ldots, \lfloor (p - m_i)/r_i \rfloor\}. \tag{20}$$

Denote by $g_{ij} = \gcd(r_i, r_j)$. Consider the spanning class $\langle a \rangle_{r_i}$ (respectively, $\langle a \rangle_{g_{ij}}$) consisting of all $a + kr_i$ (respectively, $a + kg_{ij}$), where k runs through all the integers. Let

$$\mathbb{Z}_{r_i} = \{\langle 0 \rangle_{r_i}, \langle 1 \rangle_{r_i}, \ldots, \langle r_i - 1 \rangle_{r_i}\} \tag{21}$$

and

$$\mathbb{Z}_{g_{ij}} = \{\langle 0 \rangle_{g_{ij}}, \langle 1 \rangle_{g_{ij}}, \ldots, \langle g_{ij} - 1 \rangle_{g_{ij}}\}. \tag{22}$$

Define the mapping $\Psi : \mathbb{Z}_{r_i} \longrightarrow \mathbb{Z}_{g_{ij}}$ by

$$\Psi(\langle a \rangle_{r_i}) = \langle a \rangle_{g_{ij}}. \tag{23}$$

Then, Ψ is a homomorphism of $\mathbb{Z}_{r_i}$ onto $\mathbb{Z}_{g_{ij}}$. We claim that the kernel of Ψ, denoted by Ker Ψ, is

$$\text{Ker } \Psi = \left\{\langle kr_j \rangle_{r_i}; \ k \in \mathbb{Z}\right\}. \tag{24}$$

It is readily seen that $kr_j \equiv 0 \mod g_{ij}$ for each $k \in \mathbb{Z}$, so $\langle kr_j \rangle_{g_{ij}} = \langle 0 \rangle_{g_{ij}}$, which implies that

$$\text{Ker } \Psi \supset \left\{\langle kr_j \rangle_{r_i}; \ k \in \mathbb{Z}\right\}. \tag{25}$$

On the other hand, suppose $\langle a \rangle_{r_i} \in \text{Ker } \Psi$. Then, $a = \delta g_{ij}$ for some $\delta \in \mathbb{Z}$. Since g_{ij} is the greatest common divisor of r_i and r_j, there exist $\alpha, \beta \in \mathbb{Z}$ such that

$$g_{ij} = \alpha r_i + \beta r_j. \tag{26}$$

This shows that

$$a = \delta \alpha r_i + \delta \beta r_j, \tag{27}$$

and hence $a \equiv \delta \beta r_j \mod r_i$, proving the Equation (24). Thus, $\mathbb{Z}_{r_i} / \left\{\langle kr_j \rangle_{r_i}; \ k \in \mathbb{Z}\right\}$ is isomorphic to $\mathbb{Z}_{g_{ij}}$. Denote by $l_{ij} = \text{lcm}(r_i, r_j)$. Since

$$l_{ij} = (r_i r_j / g_{ij}) | p \tag{28}$$

and the order of the kernel of Ψ is $|\text{Ker } \Psi| = |\mathbb{Z}_{r_i}| / |\mathbb{Z}_{g_{ij}}| = r_i / g_{ij}$, we may explicitly rewritten Ker Ψ by

$$
\begin{aligned}
\text{Ker } \Psi &= \{\langle 0 \rangle_{r_i}, \langle r_j \rangle_{r_i}, \langle 2r_j \rangle_{r_i}, \ldots, \langle (r_i/g_{ij} - 1)r_j \rangle_{r_i}\} \\
&= \{\langle l_{ij} \rangle_{r_i}, \langle l_{ij} + r_j \rangle_{r_i}, \langle l_{ij} + 2r_j \rangle_{r_i}, \ldots, \langle l_{ij} + (r_i/g_{ij} - 1)r_j \rangle_{r_i}\} \\
&= \cdots \\
&= \{\langle p - l_{ij} \rangle_{r_i}, \langle p - l_{ij} + r_j \rangle_{r_i}, \langle p - l_{ij} + 2r_j \rangle_{r_i}, \ldots, \langle p - l_{ij} + (r_i/g_{ij} - 1)r_j \rangle_{r_i}\}.
\end{aligned}
\tag{29}
$$

On the other hand, since

$$\text{Ker } \Psi = \{\langle 0 \rangle_{r_i}, \langle g_{ij} \rangle_{r_i}, \langle 2g_{ij} \rangle_{r_i}, \ldots, \langle (r_i/g_{ij} - 1)g_{ij} \rangle_{r_i}\}, \tag{30}$$

there exists a bijective mapping σ on $\{0, 1, \ldots, r_i/g_{ij} - 1\}$ with $\sigma(0) = 0$ such that

$$\langle hl_{ij} + kr_j \rangle_{r_i} = \langle \sigma(k)g_{ij} \rangle_{r_i} \tag{31}$$

for each $h = 0, 1, \ldots, p/l_{ij} - 1$ and $k = 0, 1, \ldots, r_i/g_{ij} - 1$. Recall that $m_j = [m_j]_{g_{ij}} + \delta g_{ij}$ for some $\delta \in \mathbb{Z}$. By Equation (30), we have

$$\langle m_j \rangle_{r_i} + \operatorname{Ker} \Psi = \langle [m_j]_{g_{ij}} + \delta g_{ij} \rangle_{r_i} + \operatorname{Ker} \Psi = \langle [m_j]_{g_{ij}} \rangle_{r_i} + \operatorname{Ker} \Psi, \tag{32}$$

and hence by Equation (31), we have

$$\langle m_j + hl_{ij} + kr_j \rangle_{r_i} = \langle [m_j]_{g_{ij}} + \sigma(k)g_{ij} \rangle_{r_i}, \tag{33}$$

which implies that there exists $\lambda \in \mathbb{Z}$ such that

$$m_j + hl_{ij} + kr_j = [m_j]_{g_{ij}} + \sigma(k)g_{ij} + \lambda r_i \tag{34}$$

for each $h = 0, 1, \ldots, p/l_{ij} - 1$ and $k = 0, 1, \ldots, r_i/g_{ij} - 1$. Since $hl_{ij} \le m_j + hl_{ij} + kr_j < (h+1)l_{ij}$ and $0 \le [m_j]_{g_{ij}} + \sigma(k)g_{ij} < r_i$ for each $h = 0, 1, \ldots, p/l_{ij} - 1$ and $k = 0, 1, \ldots, r_i/g_{ij} - 1$, it follows from Equation (34) that

$$(hl_{ij} - r_i)/r_i < (m_j + hl_{ij} + kr_j - [m_j]_{g_{ij}} - \sigma(k)g_{ij})/r_i = \lambda < (h+1)l_{ij}/r_i. \tag{35}$$

Since $\lambda \in \mathbb{Z}$, we may rewrite Equation (35) by

$$hl_{ij}/r_i \le \lambda < (h+1)l_{ij}/r_i. \tag{36}$$

With the inequalities in Equations (34) and (36) established by the construction of Ψ, we now proceed to compute Equations (19) and (20). Let $\tau \in \{0, 1\}$. For the case of $m_i \ne 0$, since

$$\lfloor (p - m_i)/r_i \rfloor = p/r_i - 1 = (p/l_{ij} - 1 + 1)l_{ij}/r_i - 1, \tag{37}$$

we have

$$
\begin{aligned}
&\{ m_i + kr_i; \ k = 0, 1, \ldots, \lfloor (p - m_i)/r_i \rfloor \} \\
= \ &\bigcup_{h=0,1,\ldots,p/l_{ij}-1} \{ m_i + \lambda_h r_i; \ \lambda_h = hl_{ij}/r_i, hl_{ij}/r_i + 1, \ldots, (h+1)l_{ij}/r_i - 1 \}.
\end{aligned}
\tag{38}
$$

Furthermore, since

$$k_j r_j = p - r_j = (p/l_{ij} - 1)l_{ij} + (r_i/g_{ij} - 1)r_j, \tag{39}$$

we have

$$
\begin{aligned}
&\{ m_j + \tau + kr_j; \ k = 0, 1, \ldots, k_j \} \\
= \ &\bigcup_{h=0,1,\ldots,p/l_{ij}-1} \{ m_j + \tau + hl_{ij} + kr_j; \ k = 0, 1, \ldots, r_i/g_{ij} - 1 \}.
\end{aligned}
\tag{40}
$$

Combining Equations (38) and (40) shows that

$$
\begin{aligned}
&\{ m_j + \tau + kr_j; \ k = 0, 1, \ldots, k_j \} \bigcap \{ m_i + kr_i; \ k = 0, 1, \ldots, \lfloor (p - m_i)/r_i \rfloor \} \\
= \ &\left(\bigcup_{h=0,1,\ldots,p/l_{ij}-1} \{ m_j + \tau + hl_{ij} + kr_j; \ k = 0, 1, \ldots, r_i/g_{ij} - 1 \} \right) \\
&\bigcap \left(\bigcup_{h=0,1,\ldots,p/l_{ij}-1} \{ m_i + \lambda_h r_i; \ \lambda_h = hl_{ij}/r_i, hl_{ij}/r_i + 1, \ldots, (h+1)l_{ij}/r_i - 1 \} \right).
\end{aligned}
\tag{41}
$$

Consider a partition of $\{\tau, \tau + 1, \ldots, \tau + p - 1\}$ given by

$$\{\tau, \tau + 1, \ldots, \tau + p - 1\} = \bigcup_{h=0,1,\ldots,p/l_{ij}-1} \{hl_{ij} + \tau, hl_{ij} + 1 + \tau, \ldots, (h+1)l_{ij} - 1 + \tau\}. \tag{42}$$

Fix $h \in \{0, 1, \ldots, p/l_{ij} - 1\}$. Since

$$0 \leq m_j \leq m_j + kr_j \leq m_j + (r_i/g_{ij} - 1)r_j = m_j + l_{ij} - r_j \leq l_{ij} - 1 \tag{43}$$

for each $k = 0, 1, \ldots, r_i/g_{ij} - 1$, we have

$$\{m_j + \tau + hl_{ij} + kr_j; \ k = 0, 1, \ldots, r_i/g_{ij} - 1\} \subset \{hl_{ij} + \tau, hl_{ij} + 1 + \tau, \ldots, (h+1)l_{ij} - 1 + \tau\}. \tag{44}$$

Furthermore, since

$$1 + hl_{ij} \leq m_i + hl_{ij} \leq m_i + \lambda_h r_i \leq m_i + ((h+1)l_{ij}/r_i - 1)r_i = m_i + (h+1)l_{ij} - r_i \leq (h+1)l_{ij} - 1 \tag{45}$$

for each $\lambda_h = hl_{ij}/r_i, hl_{ij}/r_i + 1, \ldots, (h+1)l_{ij}/r_i - 1$, we have

$$\begin{aligned}
&\{m_i + \lambda_h r_i; \ \lambda_h = hl_{ij}/r_i, hl_{ij}/r_i + 1, \ldots, (h+1)l_{ij}/r_i - 1\} \\
\subset \ &\{hl_{ij} + 1, hl_{ij} + 2, \ldots, (h+1)l_{ij} - 1\} \\
= \ &\{hl_{ij}, hl_{ij} + 1, \ldots, (h+1)l_{ij} - 1\} \cap \{hl_{ij} + 1, hl_{ij} + 2, \ldots, (h+1)l_{ij}\} \\
\subset \ &\{hl_{ij} + \tau, hl_{ij} + 1 + \tau, \ldots, (h+1)l_{ij} - 1 + \tau\}.
\end{aligned} \tag{46}$$

By Equations (42), (44) and (46), we may rearrange the intersection of elements in Equation (41) by

$$\begin{aligned}
&\{m_j + \tau + kr_j; \ k = 0, 1, \ldots, k_j\} \bigcap \{m_i + kr_i; \ k = 0, 1, \ldots, \lfloor (p - m_i)/r_i \rfloor\} \\
= \ &\bigcup_{h=0,1,\ldots,p/l_{ij}-1} \Big(\{m_j + \tau + hl_{ij} + kr_j; \ k = 0, 1, \ldots, r_i/g_{ij} - 1\} \\
&\bigcap \{m_i + \lambda_h r_i; \ \lambda_h = hl_{ij}/r_i, hl_{ij}/r_i + 1, \ldots, (h+1)l_{ij}/r_i - 1\} \Big).
\end{aligned} \tag{47}$$

Denote by $\sharp E$ the number of elements in the set E. Since $0 < m_i < r_i$ and $0 \leq [m_j]_{g_{ij}} + \sigma(k)g_{ij} < r_i$ for each $k = 0, 1, \ldots, r_i/g_{ij} - 1$, it follows from Equations (34) and (36) that

$$\begin{aligned}
&\sharp \Big(\{m_j + \tau + hl_{ij} + kr_j; \ k = 0, 1, \ldots, r_i/g_{ij} - 1\} \\
&\bigcap \{m_i + \lambda_h r_i; \ \lambda_h = hl_{ij}/r_i, hl_{ij}/r_i + 1, \ldots, (h+1)l_{ij}/r_i - 1\} \Big) \\
= \ &\sharp \left(\bigcup_{k=0,1,\ldots,r_i/g_{ij}-1} \Big(\{[m_j]_{g_{ij}} + \tau + \sigma(k)g_{ij}\} \cap \{m_i\} \Big) \right) \\
= \ &\sharp \Big(\{[m_j]_{g_{ij}} + \tau, [m_j]_{g_{ij}} + \tau + g_{ij}, \ldots, [m_j]_{g_{ij}} + \tau + r_i - g_{ij}\} \cap \{m_i\} \Big)
\end{aligned} \tag{48}$$

for each $h = 0, 1, \ldots, p/l_{ij} - 1$. It is readily seen that if

$$\sharp \Big(\{[m_j]_{g_{ij}} + \tau, [m_j]_{g_{ij}} + \tau + g_{ij}, \ldots, [m_j]_{g_{ij}} + \tau + r_i - g_{ij}\} \cap \{m_i\} \Big) = 1 \tag{49}$$

then $[m_i]_{g_{ij}} = [[m_j]_{g_{ij}} + \tau]_{g_{ij}} = [m_j + \tau]_{g_{ij}}$. Conversely, if $[m_i]_{g_{ij}} = [m_j + \tau]_{g_{ij}}$, then $m_i = [m_j]_{g_{ij}} + \tau + \delta g_{ij}$ for some $\delta \in \mathbb{Z}$. When $\tau = 0$, it follows that

$$0 \le [m_j]_{g_{ij}} + \tau + \delta g_{ij} < r_i \text{ holds only for } \delta = 0, 1, \ldots, r_i/g_{ij} - 1. \tag{50}$$

When $\tau = 1$, it follows that

$$0 < [m_j]_{g_{ij}} + \tau + \delta g_{ij} \le r_i \text{ holds only for } \delta = 0, 1, \ldots, r_i/g_{ij} - 1. \tag{51}$$

Since $0 < m_i < r_i$, we conclude that there exists only one $\delta \in \{0, 1, \ldots, r_i/g_{ij} - 1\}$ such that $m_i = [m_j]_{g_{ij}} + \tau + \delta g_{ij}$. Hence,

$$\sharp\left(\{[m_j]_{g_{ij}} + \tau, [m_j]_{g_{ij}} + \tau + g_{ij}, \ldots, [m_j]_{g_{ij}} + \tau + r_i - g_{ij}\} \cap \{m_i\}\right) = 1 \text{ if } [m_i]_{g_{ij}} = [m_j + \tau]_{g_{ij}}. \tag{52}$$

Thus, the equality in Equation (48) can be rewritten by

$$\sharp\Bigg(\{m_j + \tau + hl_{ij} + kr_j; \ k = 0, 1, \ldots, r_i/g_{ij} - 1\}$$
$$\bigcap\{m_i + \lambda_h r_i; \ \lambda_h = hl_{ij}/r_i, hl_{ij}/r_i + 1, \ldots, (h+1)l_{ij}/r_i - 1\}\Bigg) \tag{53}$$
$$= \begin{cases} 1 & \text{if } [m_i]_{g_{ij}} = [m_j + \tau]_{g_{ij}}, \\ 0 & \text{otherwise} \end{cases}$$

for each $h = 0, 1, \ldots, p/l_{ij} - 1$. Combining Equations (47) and (53) gives

$$\sharp\left(\{m_j + \tau + kr_j; \ k = 0, 1, \ldots, k_j\} \bigcap\{m_i + kr_i; \ k = 0, 1, \ldots, \lfloor(p - m_i)/r_i\rfloor\}\right) \tag{54}$$
$$= \begin{cases} p/l_{ij} & \text{if } [m_i]_{g_{ij}} = [m_j + \tau]_{g_{ij}}, \\ 0 & \text{otherwise} \end{cases}$$

for the case of $m_i \ne 0$. Now, we turn to the case of $m_i = 0$. Since $\lfloor(p - m_i)/r_i\rfloor = p/r_i$ and $\tau \in \{0, 1\}$, we have

$$\{m_i + kr_i; \ k = 0, 1, \ldots, \lfloor(p - m_i)/r_i\rfloor\} \tag{55}$$
$$= \left(\bigcup_{h=0,1,\ldots,p/l_{ij}-1} \{(\lambda_h + \tau)r_i; \ \lambda_h = hl_{ij}/r_i, hl_{ij}/r_i + 1, \ldots, (h+1)l_{ij}/r_i - 1\}\right) \bigcup\{(1-\tau)p\}$$
$$\subset \left(\bigcup_{h=0,1,\ldots,p/l_{ij}-1} \{hl_{ij} + \tau, hl_{ij} + 1 + \tau, \ldots, (h+1)l_{ij} - 1 + \tau\}\right) \bigcup\{(1-\tau)p\}.$$

Since Equations (40) and (44) hold also for the case $m_i = 0$, they imply that

$$\{m_j + \tau + kr_j; \ k = 0, 1, \ldots, k_j\}$$
$$= \bigcup_{h=0,1,\ldots,p/l_{ij}-1} \{m_j + \tau + hl_{ij} + kr_j; \ k = 0, 1, \ldots, r_i/g_{ij} - 1\} \tag{56}$$
$$\subset \bigcup_{h=0,1,\ldots,p/l_{ij}-1} \{hl_{ij} + \tau, hl_{ij} + 1 + \tau, \ldots, (h+1)l_{ij} - 1 + \tau\}.$$

The partition in Equation (42) and the inclusions in Equations (55) and (56) together imply that

$$
\begin{aligned}
& \{m_j + \tau + kr_j;\ k = 0, 1, \ldots, k_j\} \bigcap \{m_i + kr_i;\ k = 0, 1, \ldots, \lfloor (p - m_i)/r_i \rfloor\} \\
& = \bigcup_{h=0,1,\ldots,p/l_{ij}-1} \Big(\{m_j + \tau + hl_{ij} + kr_j;\ k = 0, 1, \ldots, r_i/g_{ij} - 1\} \\
& \qquad\qquad \bigcap \{(\lambda_h + \tau)r_i;\ \lambda_h = hl_{ij}/r_i, hl_{ij}/r_i + 1, \ldots, (h+1)l_{ij}/r_i - 1\} \Big).
\end{aligned}
\tag{57}
$$

Fix $h \in \{0, 1, \ldots, p/l_{ij} - 1\}$. Since $\tau r_i \in \{0, r_i\}$ and $0 \le [m_j]_{g_{ij}} + \sigma(k)g_{ij} < r_i$ for each $k = 0, 1, \ldots, r_i/g_{ij} - 1$, it follows from Equations (34) and (36) that

$$
\begin{aligned}
& \sharp\Big(\{m_j + \tau + hl_{ij} + kr_j;\ k = 0, 1, \ldots, r_i/g_{ij} - 1\} \\
& \qquad \bigcap \{(\lambda_h + \tau)r_i;\ \lambda_h = hl_{ij}/r_i, hl_{ij}/r_i + 1, \ldots, (h+1)l_{ij}/r_i - 1\} \Big) \\
& = \sharp\Big(\bigcup_{k=0,1,\ldots,r_i/g_{ij}-1} \big(\{[m_j]_{g_{ij}} + \tau + \sigma(k)g_{ij}\} \cap \{\tau r_i\} \big) \Big) \\
& = \sharp\Big(\{[m_j]_{g_{ij}} + \tau, [m_j]_{g_{ij}} + \tau + g_{ij}, \ldots, [m_j]_{g_{ij}} + \tau + r_i - g_{ij}\} \cap \{\tau r_i\} \Big).
\end{aligned}
\tag{58}
$$

It is readily seen that if

$$
\sharp\Big(\{[m_j]_{g_{ij}} + \tau, [m_j]_{g_{ij}} + \tau + g_{ij}, \ldots, [m_j]_{g_{ij}} + \tau + r_i - g_{ij}\} \cap \{\tau r_i\} \Big) = 1
\tag{59}
$$

then $[m_j + \tau]_{g_{ij}} = [[m_j]_{g_{ij}} + \tau]_{g_{ij}} = [\tau r_i]_{g_{ij}} = 0$. Conversely, if $[m_j + \tau]_{g_{ij}} = 0$, then $[m_j]_{g_{ij}} + \tau = \delta g_{ij}$ for some $\delta \in \mathbb{Z}$. When $\tau = 0$, it follows that

$$
0 \le [m_j]_{g_{ij}} = \delta g_{ij} < g_{ij} \text{ holds only for } \delta = 0.
\tag{60}
$$

When $\tau = 1$, it follows that

$$
0 < [m_j]_{g_{ij}} + 1 = \delta g_{ij} \le g_{ij} \text{ holds only for } \delta = 1.
\tag{61}
$$

This implies that $[m_j]_{g_{ij}} + \tau = \tau g_{ij}$. Hence,

$$
\begin{aligned}
& \sharp\Big(\{[m_j]_{g_{ij}} + \tau, [m_j]_{g_{ij}} + \tau + g_{ij}, \ldots, [m_j]_{g_{ij}} + \tau + r_i - g_{ij}\} \cap \{\tau r_i\} \Big) \\
& = \sharp\big(\{\tau g_{ij}, (\tau + 1)g_{ij}, \ldots, r_i + (\tau - 1)g_{ij}\} \cap \{\tau r_i\} \big) \\
& = 1 \text{ if } [m_j + \tau]_{g_{ij}} = 0.
\end{aligned}
\tag{62}
$$

Thus, the equality in Equation (58) can be rewritten by

$$
\begin{aligned}
& \sharp\Big(\{m_j + \tau + hl_{ij} + kr_j;\ k = 0, 1, \ldots, r_i/g_{ij} - 1\} \\
& \qquad \bigcap \{(\lambda_h + \tau)r_i;\ \lambda_h = hl_{ij}/r_i, hl_{ij}/r_i + 1, \ldots, (h+1)l_{ij}/r_i - 1\} \Big) \\
& = \begin{cases} 1 & \text{if } [m_j + \tau]_{g_{ij}} = 0, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}
\tag{63}
$$

Combining Equations (57) and (63) gives

$$\sharp\left(\{m_j + \tau + kr_j; \ k = 0, 1, \ldots, k_j\} \bigcap \{m_i + kr_i; \ k = 0, 1, \ldots, \lfloor (p - m_i)/r_i \rfloor\}\right) \tag{64}$$

$$= \begin{cases} p/l_{ij} & \text{if } [m_i]_{g_{ij}} = [m_j + \tau]_{g_{ij}}, \\ 0 & \text{otherwise} \end{cases}$$

for the case of $m_i = 0$. Now, we split the argument into three cases.

Case 1. $g_{ij} = 1$. Then

$$[m_j]_{g_{ij}} = [m_j + 1]_{g_{ij}} = [m_i + 1]_{g_{ij}} = [m_i]_{g_{ij}}. \tag{65}$$

By Equations (18), (19), (20), (54), *and* (64), *we have*

$$c_{ij}(\Omega) = c_{ji}(\Omega) = p/l_{ij} - p/l_{ij} = 0. \tag{66}$$

Case 2. $g_{ij} = 2$. *Then, exactly one of the following situations holds:*

$$[m_j]_{g_{ij}} = [m_i]_{g_{ij}} \ \text{ or } \ [m_j + 1]_{g_{ij}} = [m_i]_{g_{ij}}. \tag{67}$$

Suppose that $[m_j]_{g_{ij}} = [m_i]_{g_{ij}}$. *Then,*

$$[m_j + 1]_{g_{ij}} \neq [m_i]_{g_{ij}} \ \text{ and } \ [m_i + 1]_{g_{ij}} \neq [m_j]_{g_{ij}}. \tag{68}$$

By Equations (18), (19), (20), (54), *and* (64), *we have*

$$c_{ij}(\Omega) = c_{ji}(\Omega) = p/l_{ij}. \tag{69}$$

Suppose that $[m_j + 1]_{g_{ij}} = [m_i]_{g_{ij}}$. *Then,*

$$[m_j]_{g_{ij}} \neq [m_i]_{g_{ij}} \ \text{ and } \ [m_i + 1]_{g_{ij}} = [m_j]_{g_{ij}}. \tag{70}$$

By Equations (18), (19), (20), (54), *and* (64), *we have*

$$c_{ij}(\Omega) = c_{ji}(\Omega) = -p/l_{ij}. \tag{71}$$

Case 3. $g_{ij} > 2$. *Suppose that* $[m_j]_{g_{ij}} = [m_i]_{g_{ij}}$. *Then,*

$$[m_j + 1]_{g_{ij}} \neq [m_i]_{g_{ij}} \ \text{ and } \ [m_i + 1]_{g_{ij}} \neq [m_j]_{g_{ij}}. \tag{72}$$

By Equations (18), (19), (20), (54), *and* (64), *we have*

$$c_{ij}(\Omega) = c_{ji}(\Omega) = p/l_{ij}. \tag{73}$$

Suppose that $[m_j + 1]_{g_{ij}} = [m_i]_{g_{ij}}$. *Then,*

$$[m_j]_{g_{ij}} \neq [m_i]_{g_{ij}} \ \text{ and } \ [m_i + 1]_{g_{ij}} \neq [m_j]_{g_{ij}}. \tag{74}$$

By Equations (18), (19), (20), (54), *and* (64), *we have*

$$c_{ij}(\Omega) = -p/l_{ij} \ \text{ and } \ c_{ji}(\Omega) = 0. \tag{75}$$

Suppose that $[m_j]_{g_{ij}} \neq [m_i]_{g_{ij}}$ *and* $[m_j + 1]_{g_{ij}} \neq [m_i]_{g_{ij}}$. *Then, by Equations* (18), (19), (20), (54), *and* (64), *we have* $c_{ij}(\Omega) = 0$.

This completes the proof of Theorem 2. □

As an implication, we consider a feedforward neural network consisting of many layers of neurons firing in rhythm. Suppose that there are ℓ_k neurons in layer k for each $k = 1, 2, \ldots \bar{k}$, where layer 1 denotes the input layer and layer $\bar{k}$ denotes the output layer of the feedforward neural network. Thus, the coupling weight underlying the connection from the ℓth neuron in layer $\tilde{k}$ to the ℓ'th neuron in layer $\tilde{k} + 1$ is referred to the entry a_{ij}, where $i = \sum_{k=1,2,\ldots,\tilde{k}} \ell_k + \ell'$ and $j = \sum_{k=1,2,\ldots,\tilde{k}-1} \ell_k + \ell$, in the coupling weight matrix A. Thus, the coupling weight matrix A can be rewritten by

$$
\left(
\begin{array}{ccc|ccc|cccc|ccc}
a_{11} & \cdots & a_{1\ell_1} & 0 & \cdots & 0 & & 0 & & \cdots & & 0 & & 0 & \cdots & 0 \\
 & \ddots & & & \ddots & & & & & \ddots & & & & & \ddots & \\
a_{\ell_1 1} & \cdots & a_{\ell_1 \ell_1} & 0 & \cdots & 0 & & 0 & & \cdots & & 0 & & 0 & \cdots & 0 \\
a_{\ell_1+1,1} & \cdots & a_{\ell_1+1,\ell_1} & 0 & \cdots & 0 & & 0 & & \cdots & & 0 & & 0 & \cdots & 0 \\
 & \ddots & & & \ddots & & & & & \ddots & & & & & \ddots & \\
a_{\ell_1+\ell_2,1} & \cdots & a_{\ell_1+\ell_2,\ell_1} & 0 & \cdots & 0 & & 0 & & \cdots & & 0 & & 0 & \cdots & 0 \\
0 & \cdots & 0 & * & \cdots & * & & 0 & & \cdots & & 0 & & 0 & \cdots & 0 \\
 & \ddots & & & \ddots & & & & & \ddots & & & & & \ddots & \\
0 & \cdots & 0 & * & \cdots & * & & 0 & & \cdots & & 0 & & 0 & \cdots & 0 \\
0 & \cdots & 0 & 0 & \cdots & 0 & \multicolumn{2}{c}{a_{\ell_1+\cdots+\ell_{\bar{k}-1}+1,\ell_1+\cdots+\ell_{\bar{k}-2}+1}} & \cdots & \multicolumn{2}{c}{a_{\ell_1+\cdots+\ell_{\bar{k}-1}+1,\ell_1+\cdots+\ell_{\bar{k}-1}}} & 0 & \cdots & 0 \\
 & \ddots & & & \ddots & & & & & \ddots & & & & & \ddots & \\
0 & \cdots & 0 & 0 & \cdots & 0 & \multicolumn{2}{c}{a_{\ell_1+\cdots+\ell_{\bar{k}},\ell_1+\cdots+\ell_{\bar{k}-2}+1}} & \cdots & \multicolumn{2}{c}{a_{\ell_1+\cdots+\ell_{\bar{k}},\ell_1+\cdots+\ell_{\bar{k}-1}}} & 0 & \cdots & 0 \\
\end{array}
\right).
$$

Specifically, we may equip the input layer with a two-layer substructure, given by

$$
\begin{pmatrix}
a_{11} & \cdots & a_{1\ell_1} \\
 & \ddots & \\
a_{\ell_1 1} & \cdots & a_{\ell_1 \ell_1}
\end{pmatrix}
=
\left(
\begin{array}{c|ccccc}
1 & 0 & 0 & \cdots & 0 & 0 \\
\hline
1 & -1 & 0 & \cdots & 0 & 0 \\
1 & 0 & -1 & \ddots & 0 & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
1 & 0 & 0 & \ddots & -1 & 0 \\
1 & 0 & 0 & \cdots & 0 & -1
\end{array}
\right),
\tag{76}
$$

with the firing threshold satisfying $1 > b_i > 0$ for each neuron $i = 1, 2, \ldots, \ell_1$ and the updating $s(t)$ satisfying $i \in s(t) \cap s(t+1)$ if $t = m_i + \lambda r_i - 1$; otherwise, $i \notin s(t) \cap s(t+1)$ for each $i = 1, 2, \ldots, \ell_1$ and $\lambda = 0, 1, \ldots$. Let $m_1 = 0$ and $r_1 = 1$. Then, neuron i in the input layer can fire in rhythm r_i and in phase m_i for each $i = 1, 2, \ldots, \ell_1$. Neurons in the posterior layer receive signals from neurons in the prior layer and fire correspondingly if the incoming signals are greater than the firing thresholds. Note that since neurons in the prior layer fire in rhythm, the firing thresholds can be adjusted to generate certain rhythmic firing patterns or combination of rhythmic firing patterns in the posterior layer. Figure 1 depicts three layers of such a feedforward neural network, in which the coupling weight matrix A is defined by

$$
A =
\left(
\begin{array}{ccccc|cc|c}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\
\hline
0 & 5 & 3 & -1 & 1 & 0 & 0 & 0 \\
0 & 1 & -2 & 2 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & -1 & 3 & 0
\end{array}
\right),
\tag{77}
$$

the firing threshold is defined by $b = (0.5, 0.5, 0.5, 0.5, 0.5, 5, 2.5, 2.5)$, the updating is defined initially by $s(0) = \{1, 2, 6, 7, 8\}$, and

$$s(12\tau + 1) = \{1, 2, 3, 6, 7, 8\}, \ s(12\tau + 2) = \{1, 3, 4, 6, 7, 8\}, \ s(12\tau + 3) = \{1, 2, 4, 5, 6, 7, 8\}, \quad (78)$$

$$s(12\tau + 4) = \{1, 2, 3, 5, 6, 7, 8\}, \ s(12\tau + 5) = \{1, 3, 6, 7, 8\}, \ s(12\tau + 6) = \{1, 2, 4, 6, 7, 8\}, \quad (79)$$

$$s(12\tau + 7) = \{1, 2, 3, 4, 6, 7, 8\}, \ s(12\tau + 8) = \{1, 3, 6, 7, 8\}, \ s(12\tau + 9) = \{1, 2, 5, 6, 7, 8\}, \quad (80)$$

$$s(12\tau + 10) = \{1, 2, 3, 4, 5, 6, 7, 8\}, \ s(12\tau + 11) = \{1, 3, 4, 6, 7, 8\}, s(12\tau + 12) = \{1, 2, 6, 7, 8\} \quad (81)$$

for each $\tau = 0, 1, \ldots$.
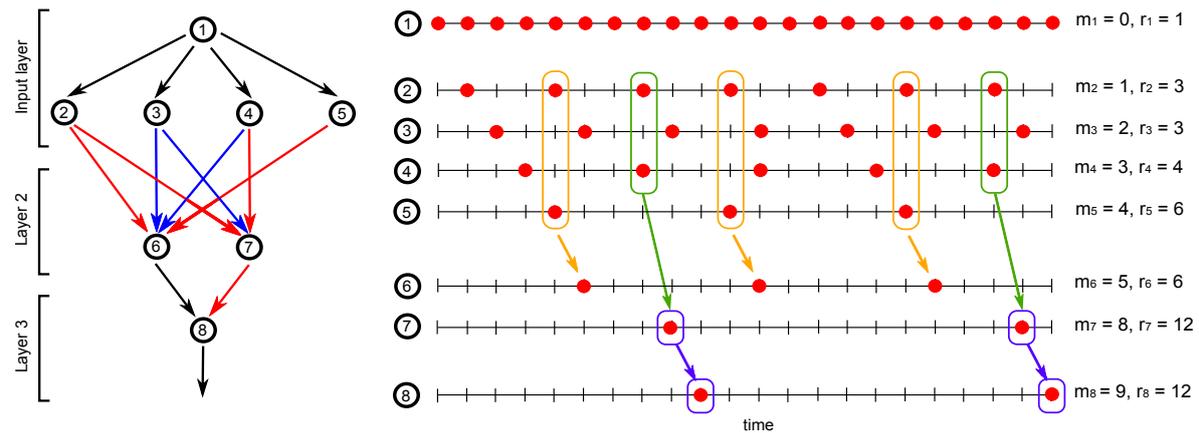
Figure 1. Rhythmic firing in a feedforward neural network.

Given an initial neural firing state $x(0) = (1, 0, 0, 0, 0, 0, 0, 0)$, the dynamical Equation (1) ensures that neurons $1, 2, \ldots, 5$ in the input layer fire in rhythm $r_1 = 1, r_2 = 3, r_3 = 3, r_4 = 4, r_5 = 6$ and in phase $m_1 = 0, m_2 = 1, m_3 = 2, m_4 = 3, m_5 = 4$, respectively. In addition, since

$$x_6(t+1) = \mathbb{1}\left(5x_2(t) + 3x_3(t) - x_4(t) + x_5(t) - 5\right), \quad (82)$$

$$x_7(t+1) = \mathbb{1}\left(x_2(t) - 2x_3(t) + 2x_4(t) - 2.5\right), \quad (83)$$

neurons $6, 7$ in layer 2 fire in rhythm $r_6 = 6, r_7 = 12$ and in phase $m_6 = 5, m_7 = 8$, respectively. This is feasible when the summation of incoming signals from neurons 2, 5 to 6 (respectively, from neurons 2, 4 to 7) is largest, and then the firing threshold of neuron 6 (respectively, neuron 7) can be adjusted to sieve out such a rhythmic firing pattern. Similar results can be found on neuron 8 in layer 3, which fires in rhythm $r_8 = 12$ and in phase $m_8 = 9$ according to

$$x_8(t+1) = \mathbb{1}\left(-x_6(t) + 3x_7(t) - 2.5\right). \quad (84)$$

Denote by Ω the resulting flow of rhythmic neural firing states from time τ to $\tau + 11$. According to the formula of the pruning map proved in Theorem 2, we see that

$$gcd(r_6, r_2) = 3 \ \text{ and } \ [m_2 + 1]_{gcd(r_6, r_2)} = [m_6]_{gcd(r_6, r_2)} = 2, \quad (85)$$

$$gcd(r_6, r_3) = 3 \ \text{ and } \ [m_3]_{gcd(r_6, r_3)} = [m_6]_{gcd(r_6, r_3)} = 2, \quad (86)$$

$$gcd(r_6, r_4) = 2 \ \text{ and } \ [m_4]_{gcd(r_6, r_4)} = [m_6]_{gcd(r_6, r_4)} = 1, \quad (87)$$

$$gcd(r_6, r_5) = 6 \ \text{ and } \ [m_5 + 1]_{gcd(r_6, r_5)} = [m_6]_{gcd(r_6, r_5)} = 5, \quad (88)$$

$$gcd(r_7, r_2) = 3 \ \text{ and } \ [m_2 + 1]_{gcd(r_7, r_2)} = [m_7]_{gcd(r_7, r_2)} = 2, \quad (89)$$

$$gcd(r_7, r_3) = 3 \ \text{ and } \ [m_3]_{gcd(r_7, r_3)} = [m_7]_{gcd(r_7, r_3)} = 2, \quad (90)$$

$$gcd(r_7, r_4) = 4 \ \text{ and } \ [m_4 + 1]_{gcd(r_7, r_4)} = [m_7]_{gcd(r_7, r_4)} = 0, \quad (91)$$

$$gcd(r_8, r_7) = 12 \ \text{ and } \ [m_7 + 1]_{gcd(r_8, r_7)} = [m_8]_{gcd(r_8, r_7)} = 9, \quad (92)$$

and hence

$$c_{62}(\Omega) < 0, \ c_{63}(\Omega) > 0, \ c_{64}(\Omega) > 0, \ c_{65}(\Omega) < 0, \tag{93}$$

$$c_{72}(\Omega) < 0, \ c_{73}(\Omega) > 0, \ c_{74}(\Omega) < 0, \ c_{87}(\Omega) < 0. \tag{94}$$

Then, by neural path pruning described in Equations (7) and (8), the coupling operator can be given by

$$d_{62} > 0, \ d_{63} < 0, \ d_{64} < 0, \ d_{65} > 0 \ \text{ and } \ d_{72} > 0, \ d_{73} < 0, \ d_{74} > 0, \ d_{87} > 0 \tag{95}$$

which maintain the rhythmic firing pattern of neuron 6 (respectively, neuron 7 and neuron 8) by keeping the largest summation of incoming signals from neurons $2, 5$ to 6 (respectively, from neurons $2, 4$ to 7 and from neuron 7 to 8). As illustrated by the orange, green, and purple boxes in Figure 1, the coupling operator in Equation (95) fits well with the neurobiological concept of Hebbian synaptic plasticity, meaning that when neurons $2, 5$ (respectively, neurons $2, 4$ and neuron 7) repeatedly or persistently take part in firing neuron 6 (respectively, neuron 7 and neuron 8), the coupling weights between them are strengthened. Such activity-dependent plasticity between layers of neurons is prone to keep neurons firing in rhythm in the posterior layer.

An experimental setting is defined as follows, showing the performance of rhythm formation in feedforward neural networks. The input rhythm and the layer architecture of feedforward neural networks are specified as in Figure 1. Initially, the coupling weights from prior layers to posterior layers and the firing thresholds of neurons are selected randomly from the interval $[-0.5, 0.5]$. When a neuron in layer 2 or 3 fires dynamically, its firing threshold will be adjusted by a positive value 1, otherwise a negative value -0.1. In addition, the coupling weights are adjusted according to the coupling operators (with intensity 1 or -1) derived from neural path pruning. The transition function of a neuron i is selected to be the Heaviside function defined by Equation (2) or the sigmoid function $\mathcal{S}$ defined by

$$x_i(t+1) = \mathcal{S}\left(\sum_{j=1}^{n} a_{ij} x_j(t) - b_i\right) = \frac{1}{1 + e^{-\alpha\left(\sum_{j=1}^{n} a_{ij} x_j(t) - b_i\right)}}, \quad t = 0, 1, \ldots, \tag{96}$$

where α is a positive real number. For each round of simulation, we say that rhythm formation occurs in the feedforward neural network if the neuron in layer 3 can persistently fire in rhythm for certain time steps. Specifically, a criterion is defined by k times of neural firing (with the firing state $x_8(t) > 0.8$) in rhythm r with $k \cdot r > 500$ during time steps $t = 1001$ to 2000. Figure 2 shows the successful rates for rhythm formation in feedforward neural networks, with neurons activating via the Heaviside functions (red line, $100,000$ rounds of simulation) or the sigmoid functions with a selected $\alpha = 0.01, 0.02, \ldots, 0.99, 1, 2, \ldots, 500$ (blue line, 1000 rounds of simulation per α). It reveals that rhythm formation can robustly occur in feedforward neural networks, even under the change of the transition functions from the Heaviside functions to the sigmoid functions.
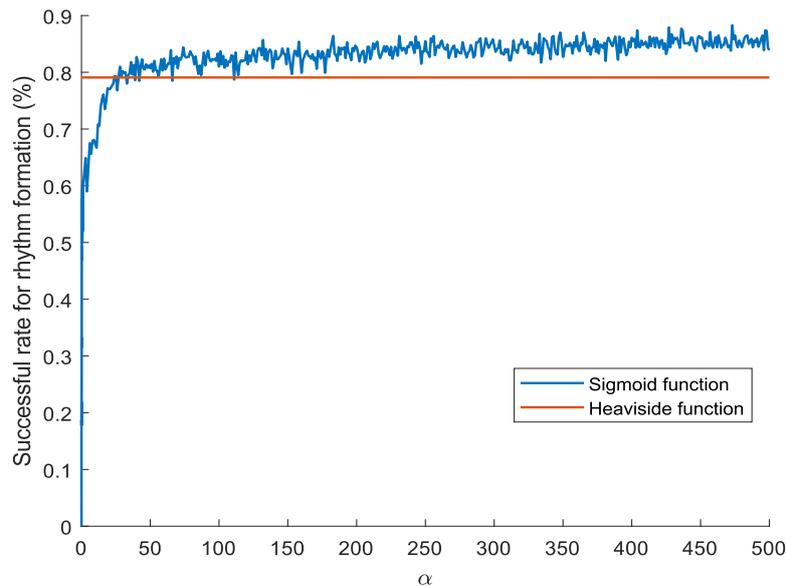
Figure 2. Successful rates for generating rhythmic firing in feedforward neural networks.

4. Conclusions

For neurons firing in rhythm, the pruning map can be formulated in terms of quantities related to rhythmic neural firing. This formulation has the potential to feed crucial control of coupling weights to support the consolidation of rhythms in feedforward neural networks. It reveals a learning rule of feedforward neural networks, which fits well with Hebbian synaptic plasticity for rhythmic pattern formation. Such neural network models can be implemented in relating different kinds of rhythms from input signals to output signals, and hence may be applied to the issues concerning the robot control or pattern recognition with rhythm. We hope that the analysis of pruning maps will stimulate further studies of underlying principles of neural networks concerning how the rhythms interact with each other and organize themselves, providing a theoretical framework mimicking the roles of rhythms in the information processing systems of living organisms.

Author Contributions: Conceptualization, F.-S.T. and S.-Y.H.; Formal analysis, F.-S.T. and S.-Y.H.; and Writing—review and editing, F.-S.T., Y.-L.S., C.-T.P., and S.-Y.H. All authors have read and agreed to the published version of the manuscript.

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

References

1. Strogatz, S.H.; Stewart, I. Coupled oscillators and biological synchronization. *Sci. Am.* **1993**, *269*, 68–75. [CrossRef] [PubMed]
2. Glass, L. Synchronization and rhythmic processes in physiology. *Nature* **2001**, *410*, 277–284. [CrossRef] [PubMed]
3. Smith, H.M. Synchronous flashing of fireflies. *Science* **1935**, *82*, 151–152. [CrossRef] [PubMed]
4. Peskin, C.S. *Mathematical Aspects of Heart Physiology*; Courant Institute of Mathematical Sciences, New York University: New York, NY, USA, 1975.
5. Mirollo, R.E.; Strogatz, S.H. Synchronization of pulse-coupled biological oscillators. *SIAM J. Appl. Math.* **1990**, *50*, 1645–1662. [CrossRef]
6. Golubitsky, M.; Stewart, I.; Török, A. Patterns of synchrony in coupled cell networks with multiple arrows. *SIAM J. Appl. Dyn. Syst.* **2005**, *4*, 78–100. [CrossRef]
7. Diesmann, M.; Gewaltig, M.-O.; Aertsen, A. Stable propagation of synchronous spiking in cortical neural networks. *Nature* **1999**, *402*, 529–533. [CrossRef] [PubMed]

8.   Grinstein, G.; Linsker, R. Synchronous neural activity in scale-free network models versus random network models. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 9948–9953. [CrossRef] [PubMed]

9.   Schechter, B. How the brain gets rhythm. *Science* **1996**, *274*, 339. [CrossRef] [PubMed]

10.   Soto-Treviño, C.; Thoroughman, K.A.; Marder, E.; Abbott, L.F. Activity-dependent modification of inhibitory synapses in models of rhythmic neural networks. *Nat. Neurosci.* **2001**, *4*, 297–303. [CrossRef] [PubMed]

11.   Buzsáki, G. *Rhythms of the Brain*; Oxford University Press: New York, NY, USA, 2006; ISBN 0-19-530106-4.

12.   Hebb, D.O. *The Organization of Behavior*; Wiley: New York, NY, USA, 1949; ISBN 978-0805843002.

13.   Adams, P. Hebb and Darwin. *J. Theoret. Biol.* **1998**, *195*, 419–438. [CrossRef] [PubMed]

14.   Sejnowski, T.J. The book of Hebb. *Neuron* **1999**, *24*, 773–776. [CrossRef]

15.   Sumbre, G.; Muto, A.; Baier, H.; Poo, M. Entrained rhythmic activities of neuronal ensembles as perceptual memory of time interval. *Nature* **2008**, *456*, 102–106. [CrossRef] [PubMed]

16.   LeCun, Y.; Boser, B.; Denker, J.S.; Henderson, D.; Howard, R.E.; Hubbard, W.; Jackel, L.D. Backpropagation applied to handwritten zip code recognition. *Neural Comput.* **1989**, *1*, 541–551. [CrossRef]

17.   He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.

18.   Huang, G.; Liu, Z.; van der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.

19.   Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 23–28 June 2014; pp. 580–587.

20.   Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26 June–1 July 2016; pp. 779–788.

21.   He, K.; Gkioxari, G.; Dollar, P.; Girshick, R. Mask R-CNN. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2961–2969.

22.   Zhang, W.; Li, R.; Deng, H.; Wang, L.; Lin, W.; Ji, S.; Shen, D. Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. *Neuroimage* **2015**, *108*, 214–224. [CrossRef] [PubMed]

23.   Moeskops, P.; Viergever, M.A.; Mendrik, A.M.; de Vries, L.S.; Benders, M.J.; Išgum, I. Automatic segmentation of MR brain images with a convolutional neural network. *IEEE Trans. Med. Imaging* **2016**, *35*, 1252–1261. [CrossRef] [PubMed]

24.   Dolz, J.; Desrosiers, C.; Ayed, I.B. 3D fully convolutional networks for subcortical segmentation in MRI: A large-scale study. *Neuroimage* **2018**, *170*, 456–470. [CrossRef] [PubMed]

25.   Shih, M.-H.; Tsai, F.-S. Growth dynamics of cell assemblies. *SIAM J. Appl. Math.* **2009**, *69*, 1110–1161. [CrossRef]

26.   Shih, M.-H.; Tsai, F.-S. Decirculation process in neural network dynamics. *IEEE Trans. Neural Netw. Learn. Syst.* **2012**, *23*, 1677–1689. [CrossRef] [PubMed]

27.   Shih, M.-H.; Tsai, F.-S. Operator control of interneural computing machines. *IEEE Trans. Neural Netw. Learn. Syst.* **2013**, *24*, 1986–1998. [CrossRef] [PubMed]

28.   Tsai, F.-S.; Hsu, S.-Y.; Shih, M.-H. Adaptive tracking control for robots with an interneural computing scheme. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 832–844. [CrossRef] [PubMed]