Preprints of the 4th IFAC Conference on Analysis and Control of Chaotic Systems, Tokyo, Japan, August 26-28, 2015

Scaling of Hodge-Kodaira decomposition distinguishes learning rules of neural networks ⋆

Keiji Miura * **Takaaki Aoki** **

* *Kwansei Gakuin University, Sanda, Hyogo 669-1337 JAPAN (e-mail: miura@kwansei.ac.jp)*
** *Kagawa University, Takamatsu, Kagawa 760-8522 JAPAN (e-mail: aoki@ed.kagawa-u.ac.jp)*

Abstract: In analyzing a complex network in the real world, it is ideally of great help to recognize its universality class. While biological networks, in particular, grow under various 'learning rules', their impacts on scaling have not yet been characterized enough. Here we applied the Hodge-Kodaira decomposition, a topological method to count global loops, to neural networks with different learning rules and edge densities. Interestingly, the networks which evolved under different learning rules showed different scalings with edge densities. The causal learning rule scaled similarly to its underlying graph (i.e. Erdős-Rényi random graph, in this study), on which a network can grow, while the Hebbian-like rule did not.

Keywords: chaos, oscillators, synchronization, networks, learning algorithms, neural dynamics, Hodge decomposition

1. INTRODUCTION

In analyzing the complex networks in the real world, recognizing their universality classes in some sense can be, ideally, of great help. In particular, co-evolving networks such as neural networks are intriguing because not only their dynamical variables on vertices but also the network structures themselves evolve over time (Aoki and Aoyagi (2007, 2009, 2011, 2012); Takahashi et al. (2009); E. M. Izhikevich (2004); Morrison et al. (2007); Buonomano (2005); Liu and Buonomano (2009); Magnasco et al. (2009)). Although biological networks can have various 'learning rules' for structural evolution, their roles on scalings have not yet been characterized enough.

The scaling of the number of global loops in a network with edge densities has been analytically obtained for the Erdős-Rényi random graph (Kahle (2009); Kahle and Meckes (2013)) by using topological methods (Fulton (1995); Hatcher (2002); Ghrist (2014); Curto et al. (2013); Chen et al. (2014); Miura and Nakada (2014)). Recent advances of the field of computational topology made it possible to compute topological invariants such as the number of "holes" in any networks in a computationally accessible way (Edelsbrunner and Harer (2009); Kaczynski et al. (2010); Arai et al. (2009); Gameiro et al. (2012); Hirata et al. (2013)).

Here we applied the Hodge-Kodaira decomposition, a topological method to count global loops (Jiang et al. (2011); Miura and Aoki (2015); Hodge (1941); Kodaira (1949); de Rham (1984); Warner (1983); Dodziuk (1974); Polthier and Preuß (2003); Bossavit (1997); Anne and Torki-Hamza (2014)), to neural network models with different biological learning rules (Aoki and Aoyagi (2009)) in order to elucidate their dependencies on edge densities. We found that the networks which evolved according to different learning rules scales differently with edge densities. Especially the spike timing dependent plasticity (STDP) rule scaled similarly to its underlying base network (i.e. Erdős-Rényi random graph, in this study), on which a network can evolve, while the Hebbian-like rule did not.

In Sec. 2, we explain the evolving neural network model we simulated. We also show the method of Hodge-Kodaira decomposition. In Sec. 3, we show the results of Hodge-Kodaira decomposition applied to the evolving neural networks. Finally, Sec. 4 presents a summary and discussions.

2. MATERIALS AND METHODS

2.1 Simulations

We simulated the following model of $N(=100)$ phase oscillators whose couplings evolve over time (Aoki and Aoyagi (2009)):

$$\frac{d\phi_i}{dt} = \omega - \frac{1}{N}\sum_{j=1}^{N} k_{ij}\sin(\phi_i - \phi_j + \alpha\pi)$$
$$\frac{dk_{ij}}{dt} = -\epsilon\sin(\phi_i - \phi_j + \beta\pi), \quad (1)$$

where ϕ_i and k_{ij} denote the phase of i-th neuron and the coupling strength from j-th to i-th neuron and we solely use $\omega = 1$ and $\epsilon = 0.005$. The learning scheme can be controlled by β: Hebb-rule for $\beta \sim -0.5$, STDP rule for $\beta \sim 0$, and anti-Hebb-rule for $\beta \sim 0.5$ (Figs. 1 and 2). We

⋆ This work was supported by JSPS KAKENHI Grant Numbers 24120701, 24120708, 24700301, 24740266, 26520202 and 26520206.

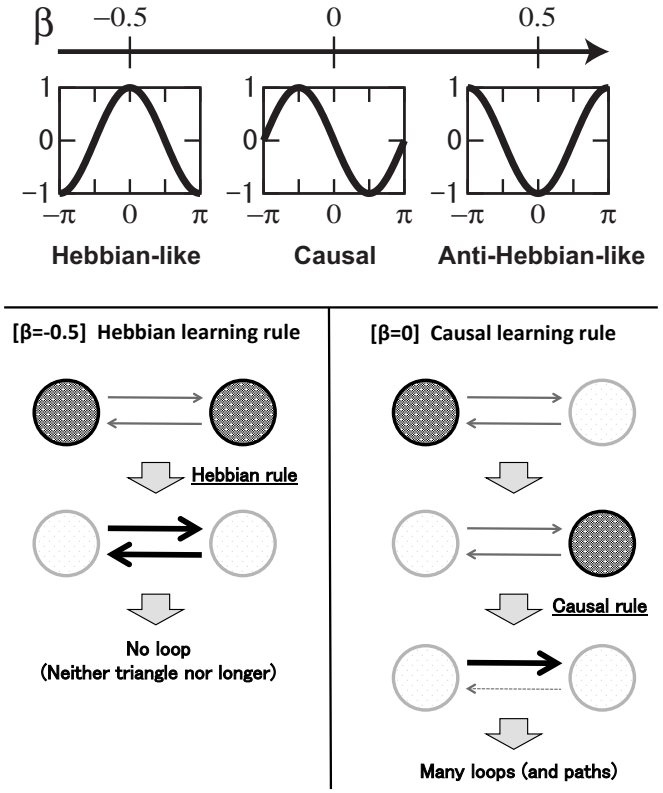

Fig. 1. Learning rule of the model: $\frac{dk_{ij}(\Delta\phi)}{dt}$ in Eq. 1 (top). The parameter β can control the learning rule. The bottom left figure illustrates the Hebbian learning rule in which the neurons that fire together wire together. The bottom right figure illustrates the causal or STDP learning rule in which the neurons that fire sequentially wire in the causal direction. Note that the phase oscillator neurons continue to fire almost periodically and the mutual timings matters in learning rules.

entirely used $\alpha = 0.3$, although the result did not change qualitatively when we used $\alpha = 0.1$.

The network can grow, i.e. the coupling strength k_{ij} can be non-zero, only on its underlying base network. The base network, which is undirected and connected, was randomly generated in the following way so that only limited coupling strengths (k_{ij}) can take non-zero (Bollobas (2001), Fig. 5 might also help). To construct an undirected graph, pairs of nodes were randomly connected with probability p (=0.05, 0.1, 0.15, 0.2, or 0.25). Note that our interest in this paper is to see how a network under a certain biological learning rule scales with this edge density p. When two nodes of the base network are connected, both directions of couplings are allowed. As we avoid self-loops by design, the base network has 100 nodes and $10000p$ directed edges (=$5000p$ undirected edges).

For mathematical convenience, the dimension of the antisymmetric "flow" was computed as the number of non-zero couplings after the time evolution where the initial coupling strengths and phases are randomized uniformly. Note that we interchangeably use flows and directed couplings in this paper. That is, the adjacency matrix consisting of the coupling strengths $K = \{k_{ij}\}$ is antisymmetrized as

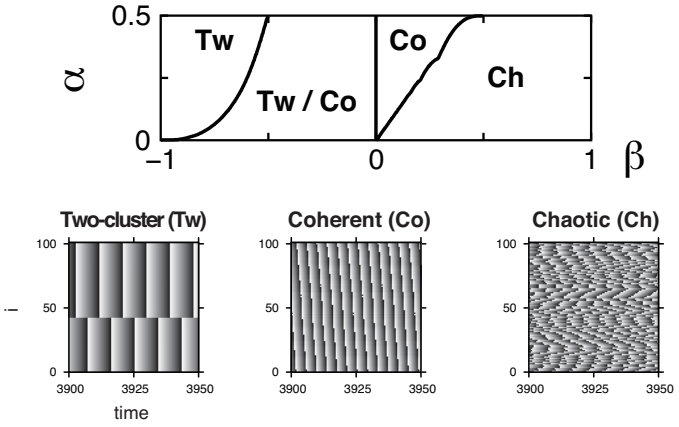

Fig. 2. Bifurcation diagram of the model consisting of three attractor states: two-cluster (Tw), coherent (Co) and chaotic (Ch) states (top). The dynamics of 100 neurons in each attractor state (bottom). The phase of each neuron in $[0, 2\pi]$ is represented by gradation. The parameter sets (α, β) are $(0.3, -0.6)$ (left), $(0.3, 0)$ (middle) and $(0.3, 0.6)$ (right). Note that although we simulated the random network with 100 neurons ($N = 100$), the above phase diagram for reference was analytically obtained under the mean field approximation for the all-to-all coupled network with $N = \infty$ (Aoki and Aoyagi (2011)).

$$A = \frac{K - K^t}{2}, \tag{2}$$

where ^{t} represents the transpose of a matrix. Then the dimension, that is, the number of non-zero components divided by two, of the antisymmetrized matrix, which represents the directed components of connections, is computed. We judged a coupling strength as non-zero when its absolute values is larger than the threshold(= 0.05). The result did not change qualitatively when we used 0.2 for the threshold although we had smaller dimensions. Similarly, for the dimension of the symmetric flows, we computed the half number of non-zero component of the symmetrized matrix:

$$S = \frac{K + K^t}{2}. \tag{3}$$

Note that any matrix can be decomposed into the symmetric and antisymmetric matrices:

$$K = A + S. \tag{4}$$

In the next subsection, we show that the antisymmetric matrix can be further decomposed into three matrices uniquely by the Hodge-Kodaira decomposition.

2.2 Hodge-Kodaira decomposition of a flow on a graph

The Hodge-Kodaira decomposition uniquely decomposes a flow on a graph into orthogonal three components: gradient flows, harmonic flows and curl flows (Fig. 3) (Jiang et al. (2011)). The gradient flow can be represented as a difference of a potential function on nodes. The potential is shown as the gray numbers in the figure for the gradient flow. The other flows are cyclic. The curl flows have only cycles of length three while the harmonic flows

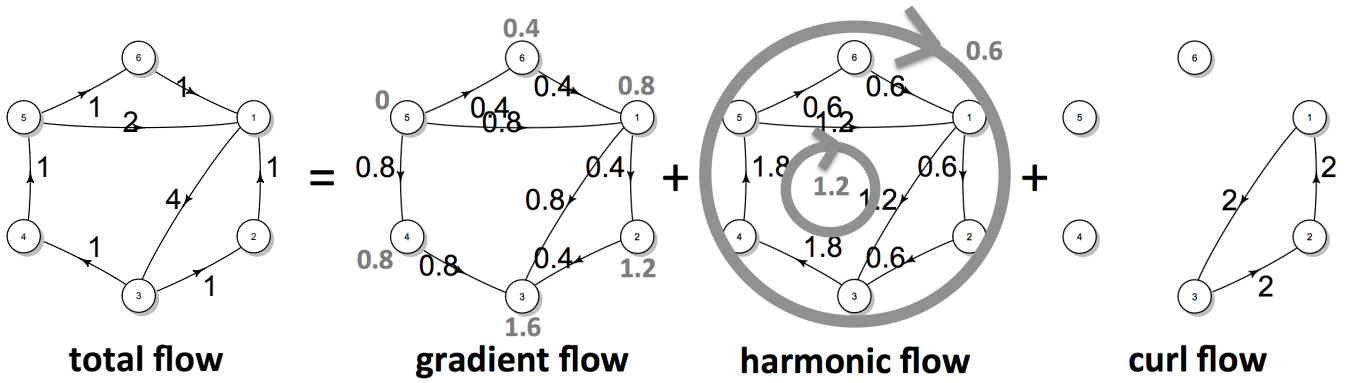

Fig. 3. An example of Hodge-Kodaira decomposition of a flow on a graph. The black numbers denote the volumes of flows on each edge. The gray numbers denote the potentials whose differences give the gradient flow. The two gray circles with numbers denote the globally cyclic flows constituting the harmonic flow.

have longer cycles. Note that the sum of the three flows on the right hand side of Fig. 3 reproduces the flow on the left hand side.

Mathematically, the space of all possible flows $\{\overrightarrow{a}\}$ (=antisymmetric components of couplings) can be represented as the space of all possible antisymmetric matrices $\{A\}$. We would like to have orthogonal direct sum decomposition of this antisymmetric matrix space into three subspaces guided by the two operators of the vector calculus.

The gradient operator is defined as an operator that returns the difference of the potentials of connected nodes on each edge. Note that the gradient operator is a linear operator as the difference can be implemented by the linear combination of potentials on nodes. The curl operator is defined as a linear operator that returns the sum of flows for each triangle. For example, the curl for the triplet of nodes 1, 2 and 3 in Fig. 3 is $(-1)+(-1)+(-4) = -6$. The curl operator is also linear as it can be implemented by the linear summation of flows on edges. Here we regard a flow(=antisymmetric matrix) a vector, $\overrightarrow{a}$, just by aligning its components. To be concrete, we show operators for the example graph in Fig. 3:

$$\begin{pmatrix} a_{12} \\ a_{13} \\ a_{23} \\ a_{34} \\ a_{15} \\ a_{45} \\ a_{16} \\ a_{56} \end{pmatrix} = \mathrm{grad}(\overrightarrow{s}) := \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \end{pmatrix},$$

$$\begin{pmatrix} C_{123} \\ C_{156} \end{pmatrix} = \mathrm{curl}(\overrightarrow{a}) := \begin{pmatrix} 1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} a_{12} \\ a_{13} \\ a_{23} \\ a_{34} \\ a_{15} \\ a_{45} \\ a_{16} \\ a_{56} \end{pmatrix} \quad (5)$$

where s_i represents a (hypothetical) potential function evaluated at the i-th node.

Based on the gradient operator, any (antisymmetric) flows can be decomposed into gradient (=difference of a potential) and non-gradient flows (=cycles). To be precise, any antisymmetric flow is composed of a linear space that can be represented by the gradient of a potential function and the orthogonal complement consisting of cycles. Similarly, based on the curl operator, you can orthogonally decompose an antisymmetric flow into curl and curl-free flows. Therefore it seems that there are four categories of flows: gradient-curl, gradient-curlfree, nongradient-curl and nongradient-curlfree. However a gradient flow is always curl-free, a known fact of vector calculus. Therefore we have only three possibilities: gradient, curl and harmonic(=nongradient-curlfree).

For numerical computations of the Hodge-Kodaira decomposition, we entirely used least square methods. Everything can be done by linear algebra as the both operators are linear. The gradient flow is obtained by finding the best potential that explains as much flow components as possible. Mathematically, this is equivalent to the linear regression in statistics that squeezes as much variations in data as possible by finding the optimal slope and intercept. Similarly, as much curl flow as possible is squeezed from the original flow by the least square method. Then the residual flow represents the harmonic flow. As the triplets of nodes are considered for the curl operator, the computation of the Hodge-Kodaira decomposition costs at least $O(N^3)$. This is why we used the network consisting of $(N =)100$ nodes, which is large enough but still computationally tractable.

3. RESULTS

Here we consider an evolving neural network model (Aoki and Aoyagi (2009)), which has a parameter β to control its learning rule for structural evolution (Materials and Methods). Therefore, we can compare the dynamics of the network structure for various learning rules such as Hebbian, STDP and anti-Hebbian as in Fig. 1. For example, it is expected that STDP rule generates more paths coincident with causal firing orders. In fact, as in Fig. 2, three attractor states have been observed respectively (Aoki and Aoyagi (2009)): two-cluster states for the Hebbian rule, coherent states for the STDP rule

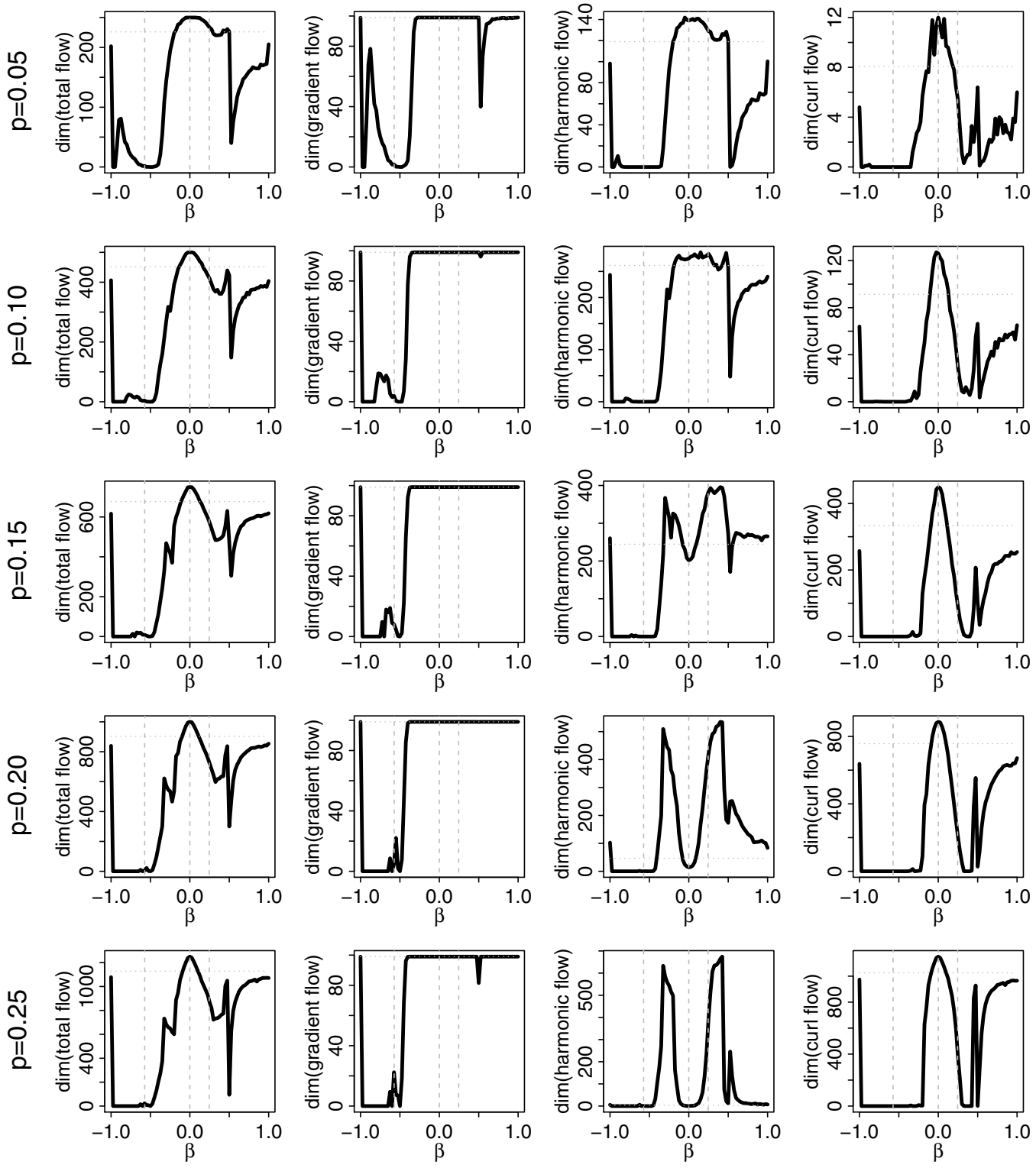

Fig. 4. Three flow components of the (antisymmetrized) simulated coupling strengths for $\alpha = 0.3$ and various β and p. The vertical dashed gray lines represent the bifurcation points in Fig. 2. The horizontal dotted gray lines represent the average initial values at $t = 0$, which are independent of β. The leftmost cluster region ($\beta << -0.5$) has no "antisymmetric" flow in most plots as the neurons fire altogether and the couplings get symmetric following the Hebbian rule. For $\beta \approx 0$ or a model with a STDP-rule that tends to form paths coincident with causal firing orders, dim(curl flow) or the number of local loops of length three is peaked. Within the rightmost chaotic region, the freedoms of harmonic and curl flows keep changing, suggesting the inhomogeneity of the chaotic region.

and chaotic states for anti-Hebbian rule. We simulated this model for a sufficiently long period ($0 \leq t \leq 80$) and analyzed its network structure. To characterize the network structure topologically, we counted the number of global loops by the Hodge-Kodaira decomposition of the connectivity matrix of the simulated model. Compared with our previous study (Miura and Aoki (2015)), here we investigate wider parameter regions: finer step sizes for β ($\Delta\beta$=0.025) as well as diverse edge densities p(=0.05, 0.1, 0.15, 0.2, or 0.25).

In Fig. 4, we plotted the degrees of freedoms for the three decomposed components of flows (gradient, harmonic and curl flows) because it is not necessarily helpful to simply draw the entire network (Miura and Aoki (2015)) and the flow dimensions are graph invariants, which is unaffected by relabeling of nodes and robust against noises on flows. The dimension (degrees of freedoms) for each flow was computed from the matrix ranks of the operators, which reflected the network structure. We plotted the dimensions at the final state ($t = 80$) averaged over 10 trials with random initial conditions. Note that the dynamics, after initial transience, went to the steady states where temporal variations were fairly small (Miura and Aoki (2015)). This is especially true for large p as the number of degrees of freedoms or edges are large. Although we randomized base graphs and initial conditions for 10 trials for each fixed parameter set (β, p), all the curves are fairly smooth. This indicates that the base structure- and initial value-dependencies are fairly small even if the network size is finite ($N = 100$). Note that the plots are periodic in β as obvious from the dynamics (Eq. 1), i.e. $\beta = 1$ and $\beta = -1$ represent the same model.

In general, the dimension of the gradient flow is equal to N(=100) - (the number of connected components). It is 99 and 0 if all the neurons are connected and disconnected, respectively. The dimension of the harmonic flow represents the number of loops whose lengths are larger than three. The dimension of the curl flow represents the number of triangles or loops whose lengths are three.

The leftmost cluster region ($\beta < -0.5$) in Fig. 4 has no flow (=no antisymmetric edge in the network) in most parameters as the neurons fire altogether and the couplings got symmetric following the Hebbian rule. Note that symmetric couplings have no (antisymmetric) flow in our definition. The fact that there is no anti-symmetric flow in the two-cluster state can be directly derived from Eq. 1. The phase difference between two neurons in the two-cluster state can be either 0 (in the same cluster) or π (in the different clusters). According to Eq. 1, k_{ij} converges to 1 or 0, respectively. Thus only symmetric couplings $k_{ij} = k_{ji}$ (= 1 or 0) are allowed in the two-cluster state. Although nonzero gradient flow dimensions, indicatives of disconnectedness, are observed especially for small p, most trials of 10 repeats showed convergence to zero and only minorities of initial conditions showed a plateau. The discrepancy from our phase diagram (dotted gray line), which was analytically computed for $N = \infty$, might be due to a finite size effect.

In the middle coherent region beyond the bifurcation point ($-0.5 < \beta$), all the flows get to have significant freedoms. Especially the gradient flow dimension, that reflects the number of connected clusters, gets full (=99), suggesting that all the neurons are connected in an oriented manner in the coherent state (sequential firings). Importantly, for a model with a STDP-rule ($\beta \approx 0$), which tends to form paths coincident with causal firing orders, we observed the most curl flows or local loops of length three. Interestingly, the harmonic flows showed strong dependency on the edge density p. Therefore we will come back to this point in Fig. 5 and replot it as a function of p to see the scaling.

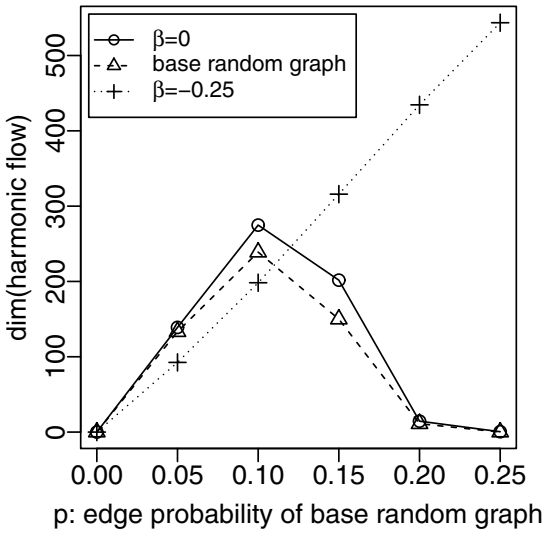

Fig. 5. Scaling of the harmonic flow dimension as a function of edge density of underlying base graph, on which coupling strengths can be non-zero. The STDP rule ($\beta = 0$) scaled similarly to the underlying Erdős-Rényi random graph (Kahle and Meckes (2013)) while the Hebbian-like rule ($\beta = -0.25$) did not.

In the rightmost chaotic region, surprisingly, the freedoms of harmonic and cyclic flows kept changing, suggesting the inhomogeneity of the chaotic region. As we used a finer step size for β ($\Delta\beta = 0.025$) than in the previous paper ($\Delta\beta = 0.1$, Miura and Aoki (2015)), we can observe the sharp transition in the left side of $\beta \approx 0.6$ and smooth relaxation in the right side.

Fig. 5 plots the scaling of the harmonic flow dimension with edge density. The same values as in Fig. 4 were replotted, but, as a functions of edge density of the base graph, on which the coupling strengths can be non-zero. The STDP rule ($\beta = 0$) reflected and scaled similarly to the underlying base graph (Erdős-Rényi random graph, in this study) while the Hebbin-like rule ($\beta = -0.25$) did not. It is known that the harmonic flow dimension for the Erdős-Rényi random graph peaks at $p = 0.1$ (Kahle (2009); Kahle and Meckes (2013)), because networks with $p \ll 0.1$ get disconnected, while networks with $0.1 \ll p$ have loops of only lengths 3 due to overcrowding.

4. SUMMARY AND DISCUSSIONS

We applied the Hodge-Kodaira decomposition, a topological method, to an evolving neural network model in order to characterize its loop structure. The Hodge-Kodaira decomposition decomposes a graph flow into three components (gradient, curl and harmonic flows), and allows us

to characterize global loop structures of a directed graph topologically. We found that the networks which evolved according to different learning rules scales differently with edge densities. Especially the causal learning rule scaled similarly to its underlying base network (i.e. Erdős-Rényi random graph, in this study), on which a network can grow, while the Hebbian-like rule did not.

In analyzing real world networks, any detailed modeling can, in principle, suffer from the lack of observational data. Instead, it may be advantageous to be aware of the universality classes for the purpose of obtaining clues on coarse network properties. The Hodge-Kodaira decomposition can be one of the useful tools to characterize and, ideally, distinguish various classes of networks.

REFERENCES

Anne, C. and Torki-Hamza, N. (2014). The gauss-bonnet operator of an infinite graph. *Anal. Math. Phys.*

Aoki, T. and Aoyagi, T. (2007). Synchrony-induced switching behavior of spike pattern attractors created by spike-timing-dependent plasticity. *Neural Comput.*, 19, 2720–38.

Aoki, T. and Aoyagi, T. (2009). Co-evolution of phases and connection strengths in a network of phase oscillators. *Phys. Rev. Lett.*, 102, 034101.

Aoki, T. and Aoyagi, T. (2011). Self-organized network of phase oscillators coupled by activity-dependent interactions. *Phys. Rev. E*, 84, 066109.

Aoki, T. and Aoyagi, T. (2012). Scale-free structures emerging from co-evolution of a network and the distribution of a diffusive resource on it. *Phys. Rev. Lett.*, 109, 208702.

Arai, Z., Kokubu, H., and Pilarczyk, P. (2009). Recent development in rigorous computational methods in dynamical systems. *Japan J. Indust. Appl. Math.*, 26, 393–417.

Bollobas, B. (2001). *Random Graphs*. Cambridge University Press, Cambridge.

Bossavit, A. (1997). *Computational Electromagnetism: Variational Formulations, Complementarity, Edge Elements*. Academic Press, Boston.

Buonomano, D.V. (2005). A learning rule for the emergence of stable dynamics and timing in recurrent networks. *J Neurophysiol.*, 94, 2275–83.

Chen, Z., Gomperts, S.N., Yamamoto, J., and Wilson, M.A. (2014). Neural representation of spatial topology in the rodent hippocampus. *Neural Comput.*, 26, 1–39.

Curto, C., Itskov, V., Veliz-Cuba, A., and Youngs, N. (2013). The neural ring: an algebraic tool for analyzing the intrinsic structure of neural codes. *Bull. Math. Biol.*, 75, 1571–1611.

de Rham, G. (1984). *Differentiable Manifolds: Forms, Currents, Harmonic Forms*. Springer-Verlag, New York.

Dodziuk, J. (1974). Combinatorial and continuous hodge theories. *Bull. Amer. Math. Soc.*, 80, 1014–1016.

E. M. Izhikevich, J. A. Gally, G.M.E. (2004). Spike-timing dynamics of neuronal groups. *Cereb. Cortex*, 14, 933–44.

Edelsbrunner, H. and Harer, J.L. (2009). *Computational Topology*. American Mathematical Society, Providence.

Fulton, W. (1995). *Algebraic Topology: A First Course*. Springer-Verlag, New York.

Gameiro, M., Hiraoka, Y., Izumi, S., Kramar, M., Mischaikow, K., and Nanda, V. (2012). Topological measurement of protein compressibility via persistence diagrams. *MI Preprint Series*, 6, 1–10.

Ghrist, R. (2014). *Elementary Applied Topology*. CreateSpace.

Hatcher, A. (2002). *Algebraic Topology*. Cambridge University Press, Cambridge.

Hirata, A., Kang, L.J., Fujita, T., Klumov, B., Matsue, K., Kotani, M., Yavari, A.R., and Chen, M.W. (2013). Geometric frustration of icosahedron in metallic glasses. *Science*, 341, 376–9.

Hodge, W.V.D. (1941). *The Theory and Applications of Harmonic Integrals*. Cambridge University Press, New York.

Jiang, X., Lim, L.H., Yao, Y., and Ye, Y. (2011). Statistical ranking and combinatorial hodge theory. *Math. Program., Ser. B*, 127, 203–244.

Kaczynski, T., Mischaikow, K., and Mrozek, M. (2010). *Computational Homology*. Springer-Verlag, New York.

Kahle, M. (2009). Topology of random clique complexes. *Discrete Mathematics*, 309, 1658–1671.

Kahle, M. and Meckes, E. (2013). Limit theorems for betti numbers of random simplicial complexes. *Homology, Homotopy and Applications*, 15, 343–374.

Kodaira, K. (1949). Harmonic fields in riemannian manifolds (generalized potential theory). *Annals of Mathematics*, 50, 587–665.

Liu, J.K. and Buonomano, D.V. (2009). Embedding multiple trajectories in simulated recurrent neural networks in a self-organizing manner. *J Neurosci.*, 29, 13172–81.

Magnasco, M.O., Piro, O., and Cecchi, G.A. (2009). Self-tuned critical anti-hebbian networks. *Phys. Rev. Lett.*, 102, 258102.

Miura, K. and Aoki, T. (2015). Hodge-kodaira decomposition of evolving neural networks. *Neural Netw.*, 62, 20–24.

Miura, K. and Nakada, K. (2014). Neural implementation of shape-invariant touch counter based on euler calculus. *IEEE Access*, 2, 960–970.

Morrison, A., Aertsen, A., and Diesmann, M. (2007). Spike-timing-dependent plasticity in balanced random networks. *Neural Comput.*, 19, 1437–67.

Polthier, K. and Preuß, E. (2003). Identifying vector field singularities using a discrete hodge decomposition. *Visualization and Mathematics III*, 113–134.

Takahashi, Y.K., Kori, H., and Masuda, N. (2009). Self-organization of feed-forward structure and entrainment in excitatory neural networks with spike-timing-dependent plasticity. *Phys. Rev. E*, 79, 051904.

Warner, F.W. (1983). *Foundations of Differentiable Manifolds and Lie Groups*. Springer-Verlag, New York.