Chaos

Inferring causal relationship in coordinated flight of pigeon flocks

Cite as: Chaos **29**, 113118 (2019); https://doi.org/10.1063/1.5120787
Submitted: 21 July 2019 . Accepted: 23 October 2019 . Published Online: 19 November 2019

Duxin Chen, Yuchen Wang, Ge Wu, Mingyu Kang, Yongzheng Sun, and Wenwu Yu

COLLECTIONS

Paper published as part of the special topic on When Machine Learning Meets Complex Systems: Networks, Chaos and Nonlinear Dynamics

Note: This paper is part of the Focus Issue, "When Machine Learning Meets Complex Systems: Networks, Chaos and Nonlinear Dynamics."

This paper was selected as Featured

Chaos ARTICLE scitation.org/journal/cha

Inferring causal relationship in coordinated flight of pigeon flocks

Cite as: Chaos 29, 113118 (2019); doi: 10.1063/1.5120787
Submitted: 21 July 2019 · Accepted: 23 October 2019 ·
Published Online: 19 November 2019

View Online Export Citation CrossMark

Duxin Chen,[1,2] Yuchen Wang,[2] Ge Wu,[2] Mingyu Kang,[2] Yongzheng Sun,[1] and Wenwu Yu[2,a)]

AFFILIATIONS

[1]School of Mathematics, China University of Mining and Technology, Xuzhou 221008, People's Republic of China
[2]School of Mathematics, Southeast University, Nanjing 210096, People's Republic of China

Note: This paper is part of the Focus Issue, "When Machine Learning Meets Complex Systems: Networks, Chaos and Nonlinear Dynamics."
[a)]**Electronic mail:** wwyu@seu.edu.cn

ABSTRACT

Collective phenomenon of natural animal groups will be attributed to individual intelligence and interagent interactions, where a long-standing challenge is to reveal the causal relationship among individuals. In this study, we propose a causal inference method based on information theory. More precisely, we calculate mutual information by using a data mining algorithm named "k-nearest neighbor" and subsequently induce the transfer entropy to obtain the causality entropy quantifying the causal dependence of one individual on another subject to a condition set consisting of other neighboring ones. Accordingly, we analyze the high-resolution GPS data of three pigeon flocks to extract the hidden interaction mechanism governing the coordinated free flight. The comparison of spatial distribution between causal neighbors and all other remainders validates that no bias exists for the causal inference. We identify the causal relationships to establish the interaction network and observe that the revealed causal relationship follows a local interaction mode. Interestingly, the individuals closer to the mass center and the average velocity direction are more influential than others.

One of the most appealing phenomena in natural gregarious animal groups is coordinated behavior. The information flow pattern, interaction mechanism, and decision-making strategy therein have attracted wide attention from various research fields. In this study, we propose an information theory based method to reveal the causal relationship among individuals. In particular, we use a data mining algorithm named *k*-nearest neighbor to quantify the mutual information and subsequently induce the transfer entropy to obtain the causal dependence of one individual on another subject to a condition set. We further analyze the high-resolution GPS data of three pigeon flocks to extract the hidden interaction mechanism by using the proposed method. We observe that the revealed causal relationship obeys a local interaction mode, and the most influential individuals in the network are the ones closer to the mass center and the average velocity direction. We suggest that the causal inference method can assist the research on collective behavior and complex systems to understand deeper the underlying interaction mechanisms.

I. INTRODUCTION

Coordinated behavior is commonly encountered in biological systems ranging from microscopic to macroscopic levels, such as bacterial colonies,[1] insect swarms,[2] fish schools,[3] bird flocks,[4] etc. To understand the underlying coordination mechanism of these appealing phenomena, related investigations have been extensively conducted in recent years, which has attracted considerable attention from biologists, physicists, and computer scientists.[5] Dating back a few decades, simple interaction rules for self-propelling particles have been proposed, which are sufficiently realistic to reproduce numerous observed phenomena. A well known work was completed by Reynolds,[6] where the author simulated the group behaviors of bird flocks, fish schools, and other coordinated biological systems according to natural observations and further proposed the famous three rules: avoid collision, velocity alignment, and cohesion to center. Vicsek *et al.* further proposed the analytical model, where the moving direction of each particle is determined by the average of its neighbors within a metric range. The so-called Vicsek model[7] captures the emergence phenomenon of highly ordered equilibrium from an arbitrary group state. Following the research line, a large variety of variations of the Vicsek model have been proposed to explain the emergence of coordination. Afterward, Couzin *et al.*[8] proposed a three-level interaction model with the consideration of a blind visual range that yields three typical coordination patterns of fish schools, i.e., swarm, torus, and consensus migration. These celebrated works inspired additional theoretical and empirical research studies on collective behaviors.[9–14] For instance, Zhang *et al.*[15,16] proposed the predictive mechanism to achieve fast flocking. Tian *et al.*[17] found the optimal view angle in collective dynamics of self-propelled agents based on the Vicsek model.

Although numerical studies provide sufficiently intuitive explanations to the interaction mechanism and emergence phenomenon of coordinated behavior, the proposed simple interaction rules still lack the support of empirical evidence. Thanks to the fast development of data acquisition techniques, such as image processing, GPS sampling, and other sensor based approaches, abundant data of various kinds of animal groups and even active matters have been collected. Generally speaking, there exist two main analytical ways to handle these massive data, i.e., correlation analysis and causal inference. The former is to analyze the signs that are indeed relevant in the sampling space. It is a process of describing the closeness of the relationship between objectives by using appropriate statistical indicators. Most of the recent works extract the correlated relationship of indices by designing appropriate correlation functions. Ballerini *et al.*[18] proposed an interaction rule of a limited fixed number of neighbors in huge flocks consisting of thousands of starlings, where they suggested that each individual interacted with around seven topological neighbors, challenging the well-accepted assumption of metric interaction range. Additionally, Nagy *et al.*[19] obtained high-resolution GPS data of pigeon flocks. They designed the velocity correlation function with time delay and found a hierarchical leadership network, where each individual plays the role as a leader or a follower. To determine whether pigeon flocks obey a despotic or egalitarian interaction pattern, Zhang *et al.*[20] suggested that pigeons tend to follow the average of their neighbors while moving smoothly, but they switch to follow the leaders during sudden turns or zigzags. Moreover, Chen *et al.*[21] reanalyzed the data of pigeon flocks[19] and indicated that the leadership is quite fixed during the long-term homing flight, while the leadership switches to a temporary individual moving in front of others during sudden turns. Other previous studies[22,23] also obtained important results of interaction mechanism of collective flights of pigeon flocks. Although the aforementioned studies reveal important correlation relationships at a sufficient confidence level, we cannot conclude that the key reasons triggering coordination among the complex systems are those correlated factors.

Causality refers to a property that produces an output response only when the input signal is excited. It quantifies the mutual dependence and real influential relationship of two variables or events. For collective behavior investigation, it goes further than only correlation analysis in establishing the causal network by determining the parent set of child set or a focal individual. In this study, we use the name of parent and child just for a single directed edge, where we call the initiation point as parent and the acceptance point as child. Thus, a node may play a role as parent of child simultaneously. One of the most classical causal inference methods is the Granger causality test[24] that facilitates extracting coupling network dynamics in climate change, large-scale human crisis, policy diffusion processes, coarse-grained molecular dynamics, and neuronal spiking systems.[25–28] This method assumes that the information about the predictions is contained in the time series of these variables and requires an estimate according to regression analysis. It is worth noting that the conclusion of the Granger causality test is only a kind of prediction, a causality in the statistical sense, not the causal relationship in the true sense. It cannot be used as a basis for affirming or negating causality. Other existing ways to asset interactions in complex systems mainly include cointegration,[29] transfer entropy,[30] convergent cross-mapping,[31] and S maps.[32] Collective oscillations or other kinds of behaviors can be regarded as coupling time series, where the effects of dynamical noise, autocorrelation, and high dimensionality are highly influential to the reconstruction of causal relationships.[33–35] From a Bayesian viewpoint, Chen *et al.* proposed a differentially variable component analysis strategy to estimate event-related signals.[36] They treated residual ongoing activity as a piecewise stationary stochastic process and used an adaptive multivariate autoregressive modeling strategy to analyze them, after subtracting out the event-related signals from recorded single trial time series. By considering whether the causal inference methods are applicable in linear and nonlinear complex systems, Stavroglou *et al.*[37] identified an unexplored form of interaction named dark causality, which is beyond the distinction between positive and negative interdependence in complex financial systems.

To investigate the strategy governing self-organized systems, Ferrante *et al.*[38] developed an active-elastic-sheet method, which explains the emergence of the order of both natural and artificial swarms. Attanasi *et al.*[39] obtained high-resolution data of thousands of starlings based on stereo imaging and proposed a realistic flocking dynamics model concerning spontaneous symmetry breaking and conservation rules. Further, Mora *et al.*[40] introduced a dynamical inference technique based on maximum entropy to infer the strength and range of alignment forces. From another view of data mining, Chen *et al.*[41] adopted a newly developed machine learning method[42] to extract interagent dependence among pigeon flocks. To quantify the causal dependence of one variable on another subject to a set of conditions, Sun *et al.* developed the transfer entropy method[43] and further proposed the optimal causation entropy (CE) strategy in dealing with multiple coupling time series. All of the aforementioned research studies provide an insight into the causal inference of coordinated behaviors.

The objective of this study is to reveal the underlying interagent dependence in collective motions with a high degree of coupling strength. To this end, we focus on circular movements of small-sized pigeon flocks and investigate the high-resolution GPS datasets consisting of 30 releases of three pigeon flocks, each of which consists of ten individuals. We use the data of three pigeon flocks (flocks A, B, and C, respectively) with more than 30 releases from Ref. 44 (a sampling rate of 10 Hz), each of which lasts for several minutes. The GPS logger was randomly allocated and affixed to the back of each pigeon before every flight. The typical format of the data consists of 3D positions and 3D velocities generated by the time difference of positions. To avoid the inherent system error of the GPS device in the z-axis and the average standard deviation of altitude of the pigeon flocks being sufficiently small (5.22 ± 1.27 m), it suffices to use only the x-axis and y-axis data for investigation. Thus, the challenge is to detect the

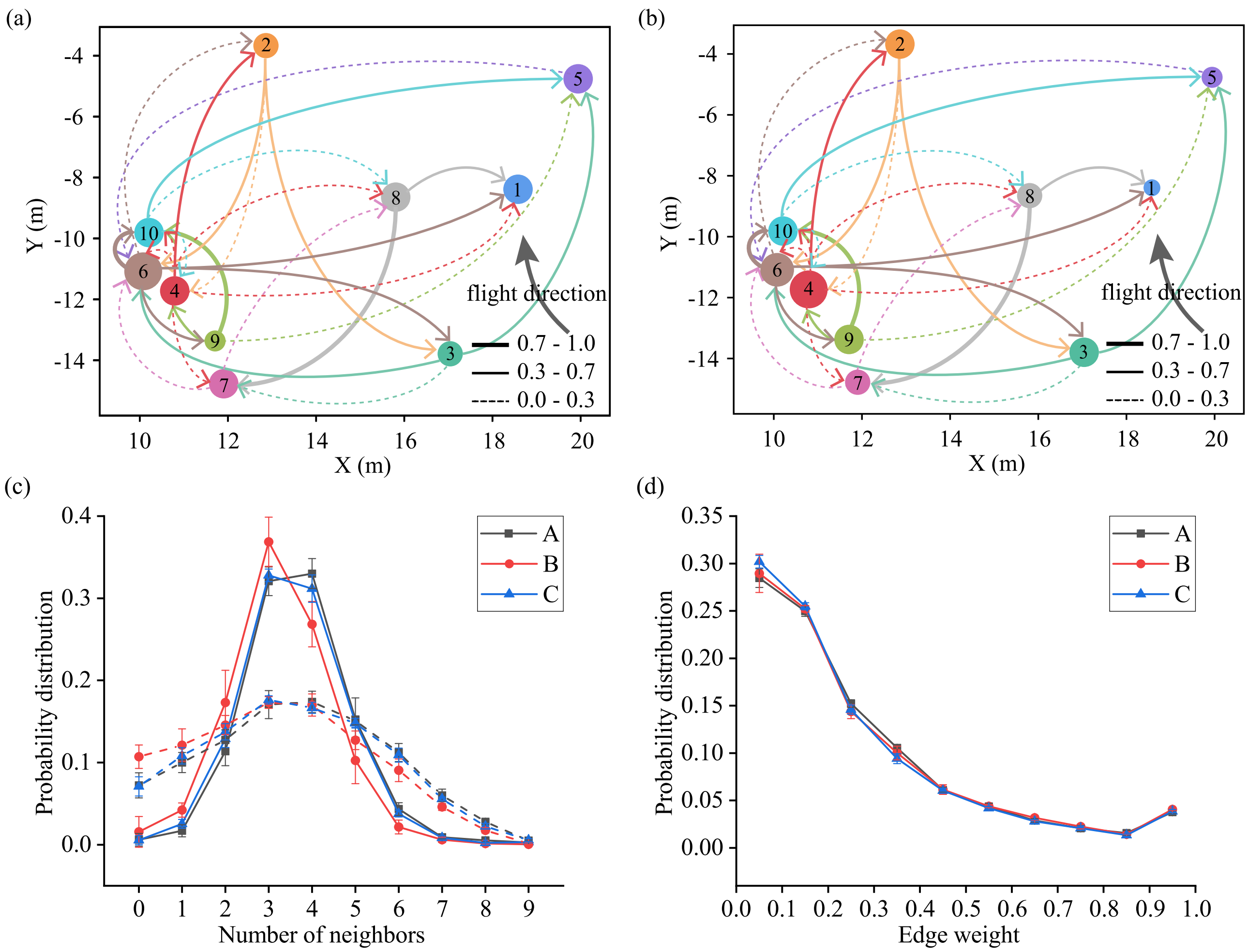

FIG. 1. Causal network construction. Different colors indicate different individuals, and all the links starting from a node will be set in the same color of the node. The size of nodes indicates the number of parents in (a) and the number of children in (b). The degree of causality is sorted into three ranges. (c) Probability distribution of the number of parents and the number of children. (d) Probability distribution of the edge weight.

dependent relationship from time series recording the highly ordered coordinated flight.

II. METHODS

A. oCEP algorithm

The causal relationship of two individuals from single time series is not difficult to extract based on the existing inference and text methods. However, in highly ordered coordinated movements, the dependent relationship becomes more complicated, since the behavior of an individual may be induced by more than one neighbors, and there may exist a condition set of individuals, which is the basic reference for the generation of causal relationships. Thus, we should consider a quantity that can detect direct information dependence and uncertainty. We use the definition of causation entropy (CE) as stated in Ref. 43 to quantify the amount of information about X obtained from Y given the condition of Z, where X, Y, and Z are node sets. In this study, every individual in the flock is regarded as a node in the causal network. Thus, CE is defined as

$$C_{Y\to X|Z} = I\left(X^{t+\tau}; Y^t | Z^t\right), \tag{1}$$

where I means the mutual information. We apply a causal inference algorithm based on the optimal causation entropy principle (oCEP)[43] to mine the causal network structure. In detail, oCEP refers to that for a given node in the network, its causal parents form the minimal set of nodes that maximizes causation entropy. According to this principle, oCEP can be used as an algorithmic approach to efficiently learn the underlying interaction network structure of pigeon flock flights. In this study, $\tau = 0.1$ s is selected based on the sampling rate of 10 Hz of the dataset. We can also choose other suitable values of τ that match the behavior of the pigeon, which will not change the main results.

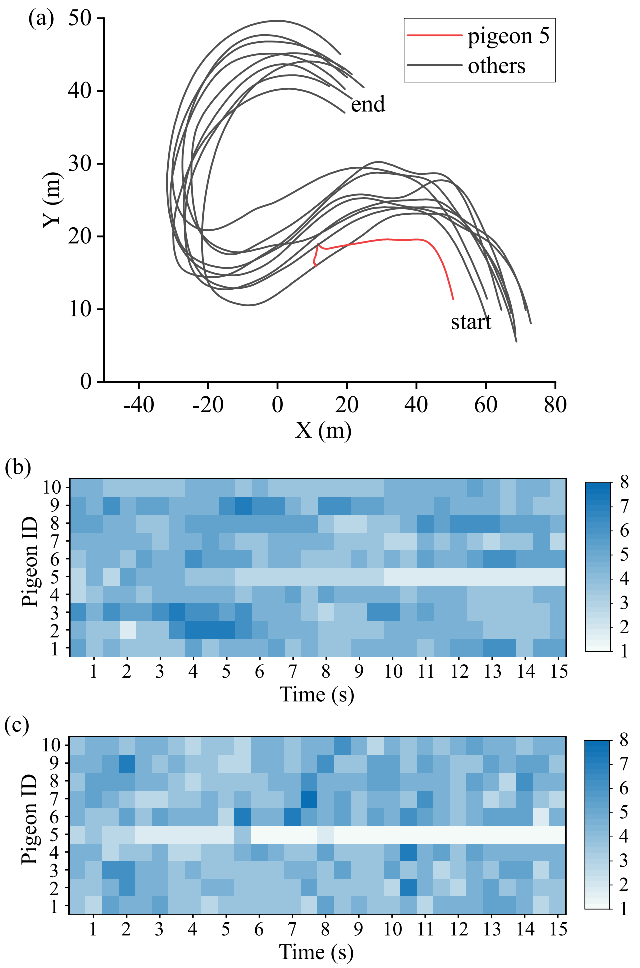

FIG. 2. (a) A segment of trajectory lasting for 15 s. (b) and (c) The evolution of the number of parents and the number of children, respectively, for every individual belonging to the trajectory.

The implementation process of the algorithm is as follows. First, for any target node x in the sampling space G (the set of all nodes), calculate the causal entropy $C_{y\to x|Z}$ of other nodes in G, where Z at first is an empty set. Second, add all possible nodes which have positive causal effects on x (both direct and indirect) to the set of Z. Third, test all nodes in Z to remove the indirectly affecting nodes that do not contribute to the value of CE and obtain the set of causal parents (a set of nodes with direct causal effects on x) Pa_x of x. Finally, by iterating to get the set of causal parents of all nodes in G, construct the inferred causal network.

The strategy of adding and removing nodes for the learning problem of the causal network construction not only reduces the computational complexity but also improves the accuracy of the inference.

The algorithm can be divided into the detailed following three aspects:

Aggregation—Select $x \in G$ as the target node randomly. The algorithm starts with an initial set of condition $Z = x$, since x at time t has a causal effect on its state at $t + \tau$. Then calculate the causal entropy $C_{y\to x|Z}, \forall y \in \{G - Z\}$ of all other nodes which are not in Z, respectively, compare the numerical values, and put the node that maximizes the causal entropy into Z. Repeat the above step until no nodes in $\{G - Z\}$ can make $C_{y\to x|Z} > 0$. After stopping adding operations, the aggregation stage ends and obtains Z, which includes all nodes that may have a causal effect on x.

Removal—Select any node except x in Z as y, calculate $C_{x\to y|Z\setminus\{y\}}, \forall y \in \{Z - x\}$. If the CE is greater than 0, then leave y in Z, otherwise delete node Y from Z; repeat the above step until all nodes except x in Z are checked. At this time, $Z - x$ is the set of causal parents Pa_x of the target node x.

Traversal—For each node $\{x_1, x_2, \ldots, x_N\}$ in the space G with N nodes, iterate over the above aggregation and removal algorithms to obtain the sets of causal parents Pa_{x_i} corresponding to x_i. From this, we can efficiently learn the underlying causal network structure.

B. KSG estimator

Essentially, CE is equal to conditional mutual information with a time delay. A straightforward method for estimating conditional mutual information is to estimate the probability density $p(x)$ by binning (histogram) methods and then calculate the conditional mutual information by the expression of entropy. The drawback of this approach is its slow convergence speed, especially for high-dimensional networks.

By estimating the density of a point by the distance based on a data mining method named k-nearest neighbor, Kraskov *et al.* proposed a KNN-based mutual information estimation method,[45] KSG estimator (named after the three inventors: Kraskov, Stögbauer, and Grassberger), to estimate the mutual information between two variables X and Y,

$$I(X;Y) = \psi(k) + \psi(N) - \langle \psi(n_x + 1) + \psi\left(n_y + 1\right)\rangle. \quad (2)$$

Here, $\langle\cdot\rangle$ means the average of all samples; parameter k indicates that k-nearest neighbor is selected; $\psi(k) = \Gamma'(t)/\Gamma(t)$ is a digamma function; N denotes the sample size, n_x is a shorthand for $n_x(i)$, which represents the number of points whose distance from the ith point $w^{(i)} \equiv \left(x^{(i)}, y^{(i)}\right)$ in the projected space of x is less than $\varepsilon(i)$, where $\varepsilon(i)$ denotes the distance from $w^{(i)}$ to its kth nearest neighbor; and n_y represents the number of points whose distance from the ith point in the projected space of y is less than $\varepsilon(i)$.

The advantage of the KSG estimator is that relatively reliable estimates can be achieved with small sample sizes.

Furthermore, Tsimpiris *et al.*[46] generalized the KSG estimator to make it suitable for estimating conditional mutual information,

$$I(X;Y|Z) = \psi(k) - \langle \psi(n_{xz} + 1) + \psi\left(n_{yz} + 1\right) - \psi(n_z + 1)\rangle, \quad (3)$$

where Z is a variable representing the condition set.

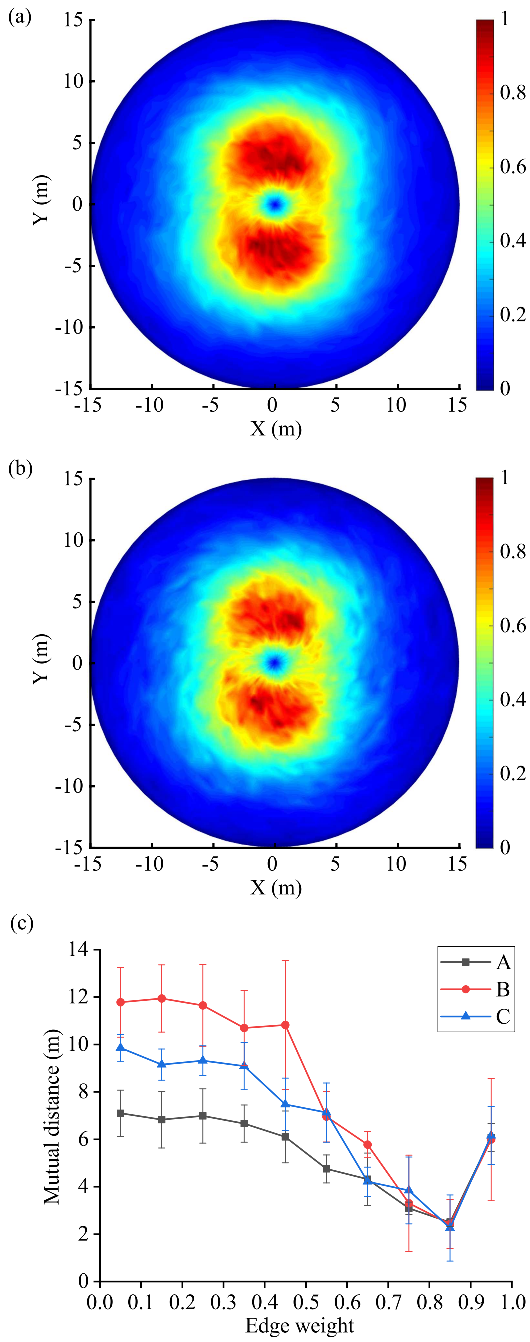

FIG. 3. (a) The average spatial distribution of the remainders for every individual in flock *A*. The average velocity direction is on the positive *x*-axis. (b) The average spatial distribution of the causal neighbors including both parent and children individuals for everyone in flock *A*. It is observed that the distribution of causal neighbors is consistent with the remainders, which suggests that the causal inference method is not random but well founded. (c) The relationship between the degree of causality and mutual distance. It indicates a strong correlation between causality and mutual distance of coordinated individuals.

When $Z = \{Z_1, Z_2, \ldots, Z_r\}$ is a multidimensional random variable, it can be generalized as

$$I\left(X; Y|Z_1, Z_2, \ldots, Z_r\right) = \psi(k) - \langle \psi(n_{xz} + 1) + \psi\left(n_{yz} + 1\right) - \psi(n_z + 1)\rangle. \quad (4)$$

Therein, $\mathbf{z} = (z_1, z_2, \ldots z_r)^{\mathrm{T}}$ and n_{xz} represents the number of points satisfying KNN in the $(r+1)$-dimensional point set $(X, Z_1, Z_2, \ldots, Z_r)$. For a given value k, let $\varepsilon(i)$ denote the distance from $w^{(i)} \equiv \left(x^{(i)}, y^{(i)}, \mathbf{z}^{(i)}\right)$ to its k-nearest neighbor. The distance is measured as the max-norm in the joint space that $\left\|w^{(i)} - w^{(j)}\right\| = \max\left\{\left\|x^{(i)} - x^{(j)}\right\|, \left\|y^{(i)} - y^{(j)}\right\|, \left\|\mathbf{z}^{(i)} - \mathbf{z}^{(j)}\right\|\right\}$. More precisely,

- $n_{xz}(i)$ is the number of points $w^{(j)}$ satisfying $\max\left\{\left\|x^{(i)} - x^{(j)}\right\|, \left\|\mathbf{z}^{(i)} - \mathbf{z}^{(j)}\right\|\right\} < \varepsilon(i), i \neq j$;
- $n_{yz}(i)$ is the number of points $w^{(j)}$ satisfying $\max\left\{\left\|y^{(i)} - y^{(j)}\right\|, \left\|\mathbf{z}^{(i)} - \mathbf{z}^{(j)}\right\|\right\} < \varepsilon(i), i \neq j$; and
- $n_z(i)$ is the number of points $w^{(j)}$ satisfying $\left\|\mathbf{z}^{(i)} - \mathbf{z}^{(j)}\right\| < \varepsilon(i), i \neq j$.

Note that the core idea of the KNN algorithm is that if the majority of the k most neighboring samples in a feature space belong to a certain category, the sample also belongs to this category and has the characteristics of the samples on this category. The method determines the category to which the sample to be classified belongs based on only the category of the nearest one or several samples in determining the classification decision. In this study, the value of k in the KNN algorithm is selected as $k = 4$ as suggested by the previous study,[47] and we implement the oCEP algorithm and the KSG estimator with the input of two-dimensional acceleration data of pigeon flocks. It mainly due to the fact that acceleration shows less autocorrelation than position and velocity, which contains more independent decisions. To obtain the causal relationship of a moment, we input the data of the last 5 s.

III. RESULTS

Based on the aforementioned causal inference method, we construct the evolutionary networks of the three pigeon flocks. To give a typical example, we display the directed causal network for a single moment in Fig. 1. If an individual is dependent on another parent, then it is a child which will be pointed by an arrow from the parent node. Therein, each individual is represented as a node, where the size indicates the in-degree (number of parents) in Fig. 1(a) and out-degree (number of children) in Fig. 1(b) of each node, respectively. Note that any individual itself is not included in the set of parent nodes of the set of child nodes. Their flight direction is uniquely indicated by the arrow on the left upper corner. The link weights indicate the values of CE for each individual which are proportionally normalized in a region of 0–1 and sorted into three ranges. The distribution of the link weight is exhibited in Fig. 1(d). It is observed that the front individuals have a larger number of both children and parents. Although intuitively the rear ones are more likely to be strongly influenced by the front ones, they also have causal influence on the front ones from our observation. Generally, an individual has

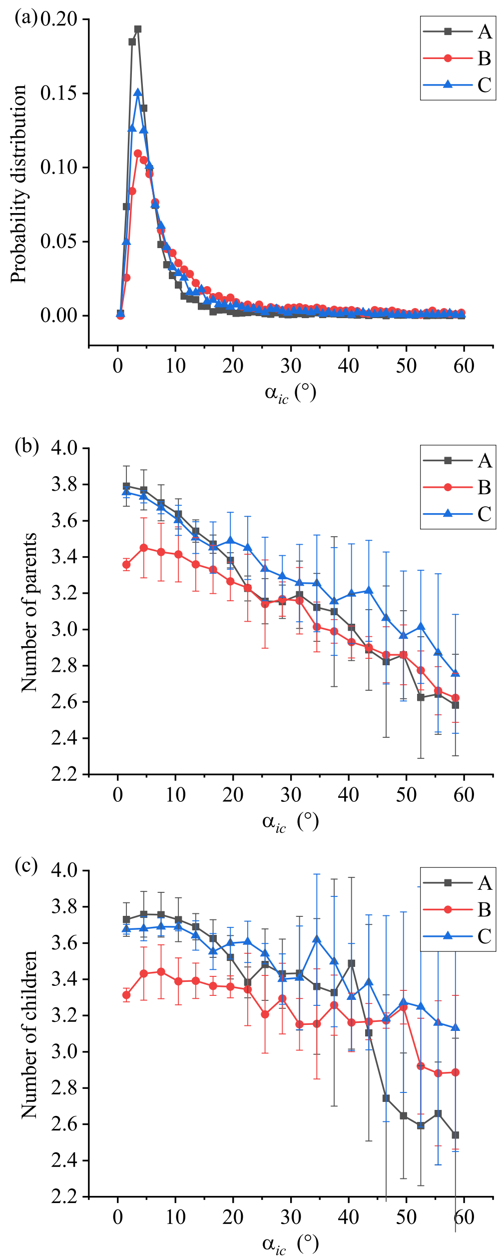

FIG. 4. (a) Probability distribution of the angle between an individual and the average velocity direction, α_{ic}. (b) The relationship between the number of parents and α_{ic}. (c) The relationship between the number of children and α_{ic}. It indicates a strong correlation between the number of causal neighbors and α_{ic}.

a set of parent nodes and a set of child nodes, which are the ones influence itself and influenced by itself. We show the probability distributions of the number of parents and children in Fig. 1(c). We find that an individual usually has three or four causal neighbors, and the probability of the case that an individual has no parent nodes is close to zero, whereas the case that an individual has no child nodes seldom occurs. The different distributions of the parent number and child number indicate that the directed causal network is not balanced or symmetric. The observation suggests that the spatial organization of the members is influential to the balance of the causal network.

To be more specific, we pick a segment of trajectory lasting for 15 s in flock A. The trajectory shown in Fig. 2(a) includes a change of rotational direction, which is common to encountered in the free flight of pigeon flocks. It can be observed that the ten individuals form a highly ordered coordinated flight, while individual 5 becomes an outlier in the later process. Based on the oCEP algorithm and the KSG estimator, we extract the causal relationships and show the evolution of parent number and child number in Figs. 2(b) and 2(c), respectively. We exhibit the number of parents and children for the first time instant in every five consecutive instants (0.5 s). It is consistent with the distribution of parent number and child number shown in Fig. 1 that the most typical case is that an individual has around four causal neighbors. Obviously, individual 5 almost has no interaction with the others in the later flight process, as it becomes an outlier then. Clearly, the statistical results of the number of causal neighbors in flock A indicate that an individual flying extremly behind has less influence on the others, while it may still be affected by others before it totally becomes an outlier.

For pigeon flocks, the flight competition is intensive, which means that the individuals are not likely to establish fixed causal connections with others. The interacting neighbors are easy to change when angular velocity is larger, since the constraints to keep cohesive exist.[48] Thus, the question how do pigeons choose interaction neighbors attracts considerable attention. A debate has been existing in previous studies that bird flocks choose neighbors in terms of metric distances,[7] and conversely birds interact with a fixed number of topological neighbors.[18] By implementing the causal inference algorithm, we show the average probability distribution of interagent metric distances of interacting neighbors corresponding to flock A in Fig. 3. We fix the position of the focal individual onto the original point in the Cartesian coordinates and show the 360° average position distribution of the nine remainders including both interacting and noninteracting ones in Fig. 3(a) and the average position distribution of causal neighbors in Fig. 3(b). We set the velocity direction on the positive x-axis for every time instant. Comparing them suggests that our causal inference method has no bias in extracting causal neighbors, since the spatial distributions seem to be similar. It can be observed that other individuals locate averagely on the left and right sides of a focal one, which indicates that the entire flock moves nearly shoulder to shoulder in the flight competition. The interaction is anisotropic in space, where the maximum probability of interaction occurs within the range of 3–4 m. However, the low causal interaction probability in the close space is due to the fact that, within a short pairwise distance, individuals are more likely to avoid collision instead of alignment with its neighbor. Compared with the results of starling flocks,[18] it is observed that pigeons in small size of flocks tend to interact symmetrically with neighbors located right or left but not along the velocity direction. Inspired by the finding, we suggest that the anisotropic interaction can be attributed to the physiological structure of visual fields of pigeons.[49,50] Analogous results can be obtained for flock B and flock C, which are no longer elaborated here.

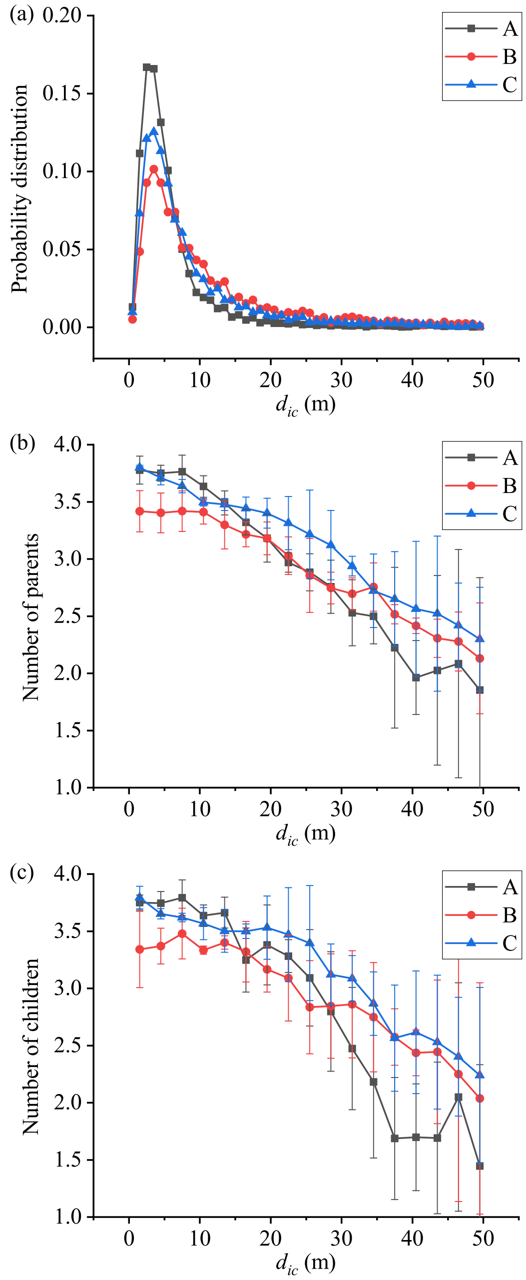

FIG. 5. (a) Probability distribution of the distance between an individual and the mass center, d_{ic}. (b) The relationship between the number of parents and d_{ic}. (c) The relationship between the number of children and d_{ic}. It indicates a strong correlation between the number of causal neighbors and d_{ic}.

Among the abundant collective motion patterns of pigeon flocks, elaborate and efficient interactions are required for the emergence of coordination. Such pairwise interactions are directed and contain the causal information of individuals. Thus, a natural question arises: what are the factors affecting the causal relationships.

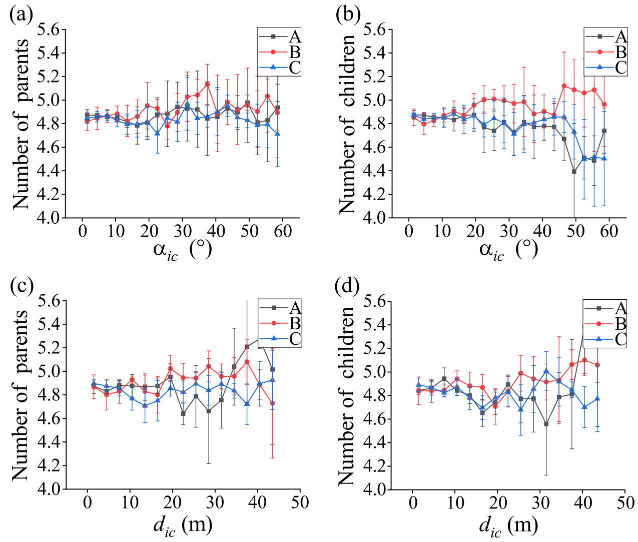

FIG. 6. Null model results. We randomly generate acceleration data with the same mean and standard deviation values with the original data for comparison. It can be observed that the number of neighbors and the angle between velocity directions α_{ic} show almost no correlation. Still, the same case is observed for correlation between the number of neighbors and the distance from the mass center d_{ic}.

To answer the appealing question, we further investigate the angle between the velocity direction of an individual and the average flight direction, namely, here α_{ic}. As shown in Fig. 4(a), the distribution of the angle between individual velocity and average velocity has a peak value around 5°. In the coordinated flight with extremely high order, the velocity directions of the individuals have little difference with the average flight direction. However, it is interesting to observe in Figs. 4(b) and 4(c) that with the increase of α_{ic}, both the numbers of parent and child decrease gradually. It is intuitive to consider that if the velocity of an individual is greatly different from the flock, it may fail to establish connection with the others. Thus, the monotonic decreasing curves of causal number of neighbors vs the intersection angle α_{ic} supports the fact that the closer to the average of the moving direction, the more causal interactions.

Another factor under consideration is the metric distance between the position of an individual to the mass center, namely, here d_{ic}. Analogously, we exhibit the distribution of distance in Fig. 5(a) and observe that the curve has a peak around $d_{ic} = 4$ m, and it decreases fast afterwards. It is consistent with the spatial distribution of remainders shown in Fig. 3(a). Interestingly, with an increase of the distance to the mass center, both the numbers of parent and child decrease gradually. It means that an individual flying far away from the flock has fewer causal neighbors. On the other hand, an individual close to the mass center may have a larger betweenness centrality in the causal network, since it may have more interacting neighbors.

For comparison, we construct null models by randomly generating acceleration data with the same mean and standard deviation values with the original data, see Fig. 6. It can be observed that the number of neighbors and the angle between velocity directions, α_{ic},

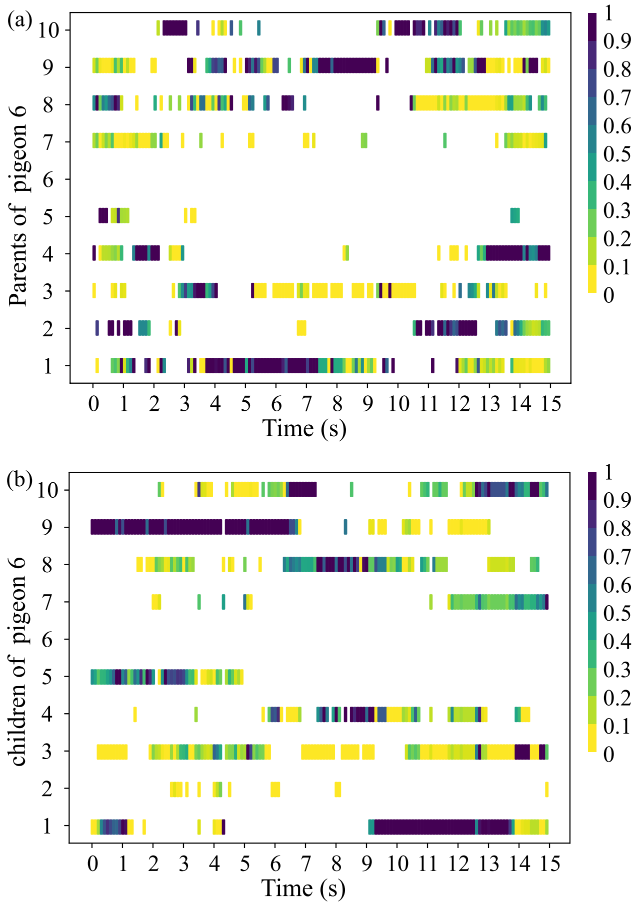

FIG. 7. Evolution of the value of causation entropy of the parents and children of individual 6. It can be observed that although the configuration of causal links can and does change in time, they are stable and reliable.

show almost no correlation. Still, the same case is observed for correlation between the number of neighbors and the distance from the mass center, d_{ic}. This shows that our algorithm has no bias in extracting the causal network characteristics.

Note that the casual networks change less frequently and have less links, see Fig. 7. In this example, we focus on individual 6 and show the value of causation entropy of the parents and children of it. Although the configuration of causal links can and does change in time, the stability and reliability of the causal network can be guaranteed. Although there is occasionally some dropout of links from one instant to the next, the overall results of the algorithm are clearly stable and more-or-less continuous in time. Long stretches of symbols or blank space demonstrate that the oCEP algorithm is robust to changes in the spatial configuration and kinematics of the flock.

IV. CONCLUSIONS

This study focuses on the causal inference for coordinated behaviors. We have taken free flight of pigeon flocks as an example, since the alignment of velocity may carry significant causal information. To quantify the causal dependence of one individual on another subject to a condition set consisting of other neighboring ones, requires obtaining the mutual information at first. By regarding the complex coordinated movement as a strongly coupling time series, we have calculated mutual information by using the data mining algorithm k-nearest neighbor according to the KSG estimator proposed in the previous study.[45] Furthermore, we can induce the transfer entropy, which describes the causal dependence of an individual on another. However, to exactly extract the causality in coordinated behavior, we should consider a condition set of events that is essential for the causal relationship. Thus, we have introduced the oCEP method, by which we can obtain the causality entropy based on the calculation of mutual information. The implementation process of the oCEP algorithm includes three steps: aggregation, deletion, and integration. The strategy of adding and removing nodes for the learning problem of causal network structure reduces the computational complexity and improves the accuracy of the inference. The comparison of spatial distribution between causal neighbors and the remainders including both causal and noncausal individuals validates that no bias exists for the causal inference. Accordingly, we reveal the hidden interaction mechanism governing the coordinated free flight for three pigeon flocks. We identify the causal relationships to establish the directed network. Importantly, we observe that the revealed causal relationship follows a local interaction mode, and if the individual is closer to the mass center and the average velocity direction, it is more influential to the group members.

We suggest that the proposed causal inference method helps discover the interaction mechanism from the perspective of a dynamical analysis in collective behavior. In addition, the proposed causal inference method can provide some insights into the research of various kinds of complex systems to deeper understand the underlying interaction mechanisms. Furthermore, it will be necessary to scale up from the current small-sized bird flocks to larger-sized animal groups, e.g., European starlings, and to check the interspecies issues. How can the adopted method be extended to other types of coordinated behavior? What are the causality and interaction mechanism in other gregarious animal groups? These fascinating questions merit further investigations in the future.

ACKNOWLEDGMENTS

We would like to express our sincere gratitude to Professor Jie Sun for useful discussion and guidance. This work is supported by the Fundamental Research Funds for the Central Universities (Grant Nos. 2015XKMS076 and 2242019K40111), the National Natural Science Foundation of China (NNSFC) (Grant Nos. 61903079 and 61673107), and the (Grant No. BM2017002).

REFERENCES

[1] A. Sokolov, I. S. Aranson, J. O. Kessler, and R. E. Goldstein, Phys. Rev. Lett. **98**, 158102 (2007).
[2] J. Buhl, D. J. Sumpter, I. D. Couzin, J. J. Hale, E. Despland, E. R. Miller, and S. J. Simpson, Science **312**, 1402 (2006).
[3] D. J. Hoare, I. D. Couzin, J. G. Godin, and J. Krause, Anim. Behav. **67**, 155 (2004).
[4] I. L. Bajec and F. H. Heppner, Anim. Behav. **78**, 777 (2009).
[5] T. Vicsek and A. Zafeiris, Phys. Rep. **517**, 71 (2012).
[6] C. W. Reynolds, ACM SIGGRAPH Comput. Graph. **21**, 25 (1987).

[7] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, Phys. Rev. Lett. **75**, 1226 (1995).
[8] I. D. Couzin, J. Krause, R. James, G. D. Ruxton, and N. R. Franks, J. Theor. Biol. **218**, 1 (2002).
[9] R. Olfati-Saber, IEEE Trans. Automat. Contr. **51**, 401 (2006).
[10] W. Yu, G. Chen, M. Cao, and J. Kurths, IEEE Trans. Syst. Man Cybern. B **40**, 881 (2009).
[11] N. W. Bode, D. W. Franks, and A. J. Wood, J. R. Soc. Interface **8**, 301 (2010).
[12] J. Gao, S. Havlin, X. Xu, and H. E. Stanley, Phys. Rev. E **84**, 046115 (2011).
[13] Z. Cheng, Z. Chen, T. Vicsek, D. Chen, and H. T. Zhang, New J. Phys. **18**, 103005 (2016).
[14] X. Mo, Z. Chen, and H. T. Zhang, Automatica **103**, 254 (2019).
[15] H. T. Zhang, M. Z. Q. Chen, and G. B. Stan *et al.*, IEEE Circ. Syst. Mag. **8**, 67 (2008).
[16] H. T. Zhang, M. Z. Q. Chen, T. Zhou, and G. B. Stan, Europhys. Lett. **83**, 40003 (2008).
[17] B. M. Tian, H. X. Yang, W. Li, W. X. Wang, B. H. Wang, and T. Zhou, Phys. Rev. E **79**, 052102 (2009).
[18] M. Ballerini, N. Cabibbo, and R. Candelier *et al.*, Proc. Natl. Acad. Sci. U.S.A. **105**, 1232 (2008).
[19] M. Nagy, Z. Ákos, D. Biro, and T. Vicsek, Nature **464**, 890 (2010).
[20] H. T. Zhang, Z. Chen, T. Vicsek, G. Feng, L. Sun, R. Su, and T. Zhou, Sci. Rep. **4**, 5805 (2014).
[21] D. Chen, T. Vicsek, X. Liu, T. Zhou, and H. T. Zhang, Europhys. Lett. **114**, 60008 (2016).
[22] G. D. Kattas, X. K. Xu, and M. Small, PLoS Comput. Biol. **8**, e1002449 (2012).
[23] X. K. Xu, G. D. Kattas, and M. Small, Phys. Rev. E **85**, 026120 (2012).
[24] C. W. J. Granger, Econometrica **37**, 424 (1969).
[25] D. D. Zhang, H. F. Lee, C. Wang, B. Li, Q. Pei, J. Zhang, and Y. An, Proc. Natl. Acad. Sci. U.S.A. **108**, 17296 (2011).
[26] S. Gerber and I. Horenko, Proc. Natl. Acad. Sci. U.S.A. **111**, 14651 (2014).
[27] C. Grabow, J. Macinko, D. Silver D, and M. Porfiri, Chaos **26**, 083113 (2016).
[28] A. Sheikhattar, S. Miran, and J. Liu *et al.*, Proc. Natl. Acad. Sci. U.S.A. **115**, E3869 (2018).
[29] R. F. Engle and C. W. J. Granger, Econometrica **55**, 251 (1987).
[30] T. Schreiber, Phys. Rev. Lett. **85**, 461 (2000).
[31] G. Sugihara, R. May, and H. Ye, Science **338**, 496 (2012).
[32] E. R. Deyle, R. M. May, S. B. Munch, and G. Sugihara, Proc. R. Soc. B **283**, 2015 (2016).
[33] C. Koutlis and D. Kugiumtzis, Chaos **26**, 093120 (2016).
[34] M. Zanin and D. Papo, Chaos **27**, 047403 (2017).
[35] J. Runge, Chaos **28**, 075310 (2018).
[36] Y. Chen, S. L. Bressler, K. H. Knuth, W. A. Truccolo, and M. Ding, Chaos **16**, 026113 (2006).
[37] S. K. Stavroglou, A. A. Pantelous, H. E. Stanley, and K. M. Zuev, Proc. Natl. Acad. Sci. U.S.A. **116**, 10646 (2019).
[38] E. Ferrante, A. E. Turgut, M. Dorigo, and C. Huepe, Phys. Rev. Lett. **111**, 268302 (2013).
[39] A. Attanasi *et al.*, Nat. Phys. **10**, 691 (2014).
[40] T. Mora, A. M. Walczak, and L. D. Castello *et al.*, Nat. Phys. **12**, 1153 (2016).
[41] D. Chen, B. Xu, T. Zhu, T. Zhou, and H. T. Zhang, Phys. Rev. E **96**, 022411 (2017).
[42] W. Pan, Y. Yuan, J. Gonçalves, and G. B. Stan, IEEE Trans. Automat. Contr. **61**, 182 (2015).
[43] J. Sun, D. Taylor, and E. M. Bollt, SIAM J. Appl. Dyn. Syst. **14**, 73 (2015).
[44] M. Nagy, G. Vásárhelyi, B. Pettit, I. Roberts-Mariani, T. Vicsek, and D. Biro, Proc. Natl. Acad. Sci. U.S.A. **110**, 13049 (2013).
[45] A. Kraskov, H. Stögbauer, and P. Grassberger, Phys. Rev. E **69**, 066138 (2004).
[46] A. Tsimpiris, I. Vlachos, and D. Kugiumtzis, Expert Syst. Appl. **39**, 12697 (2012).
[47] W. M. Lord, J. Sun, N. T. Ouellette, and E. M. Bollt, IEEE Trans. Mol. Biol. Multi-Scale Commun. **2**, 107 (2016).
[48] M. Yomosa, T. Mizuguchi, G. Vásárhelyi, and M. Nagy, PLoS ONE **10**, e0140558 (2015).
[49] G. R. Martin, J. Comp. Physiol. A **159**, 545 (1986).
[50] G. R. Martin and G. Katzir, Brain. Behav. Evol. **53**, 55 (1999).