New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
3 October 2018

REVISED
2 December 2018

ACCEPTED FOR PUBLICATION
20 December 2018

PUBLISHED
31 January 2019

PAPER

Ultrafast synchronization via local observation

Hai-Tao Zhang[1], Ming-Can Fan[2], Yue Wu[1], Jianxi Gao[3], H Eugene Stanley[4], Tao Zhou[5,6] and Ye Yuan[1,6]

[1] School of Automation, Huazhong University of Science and Technology, Wuhan 430074, People's Republic of China
[2] Department of Mathematics, Huizhou University, Huizhou 516007, Guangdong, People's Republic of China
[3] Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY 12180, United States of America
[4] Center for Polymer Studies and Department of Physics, Boston University, Boston, MA 02215, United States of America
[5] CompleX Lab, Web Sciences Center, University of Electronic Science and Technology of China, Chengdu 611731, People's Republic of China
[6] Authors to whom any correspondence should be addressed.

E-mail: zhutou@ustc.edu **and** yye@hust.edu.cn

Keywords: synchronization, prediction, consensus

Abstract

This paper proposes a Hankel matrix-based statistical study to calculate the final synchronization state of the entire network via local observation of just a single node for a time period significantly shorter than the synchronization process. It surfaces that synchronization can be achieved more quickly than the routine rhythm for networks no matter with linear or nonlinear dynamics. This finding refines our understanding of the abundant ultrafast synchronization phenomena observed in nature, which enables the efficient design of self-aligned robots as well.

1. Introduction

Synchronization is ubiquitous in nature [1, 2], man-made systems [3] and human behaviors [4]. Understanding synchronizing processes and regulating synchronizability have already benefited both biological and engineering systems [5], including foraging [6], predator avoidance [7], migration [8], collective control of unmanned air vehicles [9], and the self-organized formation of multi-robot systems [10]. Synchronization phenomena and such closely related concepts as collective motion and consensus have already attracted more and more attentions in many branches of science [11, 12].

Many mechanisms have been proposed to explain synchronization phenomena [12]. The best known is the neighborhood coordination mechanism [13] in which the activity of each individual is affected by their nearest neighbors. The neighbors of an individual are defined to be (i) those inside a ball-shaped vision range of a fixed radius [13, 14], (ii) those directly connected in a network [15, 16], or (iii) those, limited in number, that are closest [17]. A 'hierarchical leadership model' was proposed [18] to explain the flock of pigeons, where each pigeon follows its leader and is in turn followed by other pigeons, resulting in a hierarchical leader-follower network.

Empirical studies have found that synchronization emerges quickly in real-world ecological and biological systems [19–21]. In contrast, the synchronization produced by the neighborhood coordination mechanism is gradual. Although many methods have been proposed to speed up the synchronizing process [22–24], the neighborhood coordination mechanism cannot achieve extremely rapid synchronization or coordination as observed in real-world systems. The hierarchical leadership model has not been validated in large-scale systems [25]. Two candidate mechanisms, information propagation [21, 26] and predictive protocol [27–30], have been proposed to explain ultrafast synchronization. The former argues that direction change information can quickly propagate throughout the flock without attenuation, and the latter shows that an individual, such as a bird or a fish, is able to predict the near-future moving trajectories of neighbors, and thus is more able to anticipate collective motion. Understanding ultrafast synchronization is still an open challenge, because these two proposed mechanisms need further experimental validation. In addition, it is probable that the observed phenomena are the result of the integrated effects of multiple mechanisms.

We here implement a statistical analysis using Hankel matrix-based prediction method [30], for reaching ultrafast network synchronization, where the Hankel matrix is constructed only from the finite-time and local history of the networked system. In connected networks, we find that the record of the past states of the observed node can be used to achieve ultrafast synchronization. Monitoring additional nodes in the neighborhood of the initial node further accelerates the synchronization. We demonstrate the ultrafast synchronizing speed of this mechanism using simulations of representative network models and of a variety of real networks, no matter with linear or nonlinear dynamics.

2. Methods and results

For simplicity, we take an N-node general directed network with linear dynamics for example. When there is an edge from node j to node i, $a_{ij} = 1$ in the adjacency matrix $\mathcal{A}$. Otherwise, $a_{ij} = 0$. The state x_i of an arbitrary node i follows a discrete-time linear dynamics

$$x_i(t+1) = x_i(t) + \epsilon \sum_{j=1}^{N} a_{ij}[x_j(t) - x_i(t)], \tag{1}$$

where ϵ is the sampling period, which is small enough ($\epsilon \leqslant 1/d_{\max}$ and $d_{\max}$ is the maximal out-degree) to guarantee convergence [31]. Then the dynamics of the entire network is

$$x(t+1) = Px(t), \tag{2}$$

where $x = (x_1, x_2, \cdots, x_N)^{\mathsf{T}}$, $P = I - \epsilon(\mathcal{D} - \mathcal{A})$, I is the unit matrix, and $\mathcal{D} = \mathrm{diag}\{1^{\mathsf{T}}\mathcal{A}\}$ with $\mathbf{1}$ being an N-dimensional all-1 vector. As a simple fact, the state $x(t)$ asymptotically converges to the final value $x(\infty) = \mu x(0)\,1$ if the spectral radius of P is no greater than 1. Here μ is the left eigenvector of P corresponding to eigenvalue 1, which also satisfies the normalization condition $\mu\mathbf{1} = 1$. Specifically, for undirected networks or balanced directed networks (i.e. $\sum_j a_{ij} = \sum_j a_{ji}$ for every node i), $x(\infty) = \frac{1}{N}\sum_{i=1}^{N} x_i(0)\,1$. An essential problem naturally emerges: how to predict the future dynamics of the entire network without the global knowledge of P or $x(0)$, but just by observing the state of a single node?

To this end, we designate node i to be the one from which we gather time-sequential information about itself and its ℓ neighboring nodes $i_1, \cdots, i_\ell$ ('monitored nodes') as $y_i = (x_i, x_{i_1}, \cdots, x_{i_\ell})^{\mathsf{T}}$. We define an output matrix $C_i \in \mathbb{R}^{(\ell+1)\times N}$ in which column i in the first row and columns i_j in $(j+1)$th rows are 1 ($j = 1, 2, \cdots, l$). All other elements are 0. Thus

$$y_i(t) = (x_i(t), x_{i_1}(t), \cdots, x_{i_\ell}(t))^{\mathsf{T}} = C_i x(t). \tag{3}$$

We designate D_i to be the smallest integer that satisfies condition $C_{iqi}(P) = 0$, where $q_i(\cdot)$ is a monic polynomial with degree $D_i + 1$, and $q_i(z) = \sum_{j=0}^{D_i+1} \alpha_j^{(i)} z^j$ with $\alpha_j^{(i)}$, $j = 0, \cdots, D_i$ being free parameters and $\alpha_{D_i+1}^{(i)} = 1$. Since $y_i(t+j) = C_i x(t+j) = C_i P^j x(t)$, for any time t we have

$$\sum_{j=0}^{D_i+1} \alpha_j^{(i)} y_i(t+j) = C_i q_i(P) x(t) = 0. \tag{4}$$

Noting the Z-transform $Y_i(z) = \mathcal{Z}(y_i(t))$, from equation (4) and the time-shift property of the Z-transform we have

$$Y_i(z) = \frac{\sum_{j=1}^{D_i+1} \alpha_j^{(i)} \left(\sum_{h=0}^{j-1} y_i(h) z^{j-h} \right)}{q_i(z)} \triangleq \frac{H(z)}{q_i(z)}. \tag{5}$$

According to the definition of P in (2), the only unstable root of $q_i(z)$ is the one at 1. We then define

$$p_i(z) = \frac{q_i(z)}{z-1} = \sum_{j=0}^{D_i} \beta_j z^j, \tag{6}$$

which immediately leads to that

$$\sum_{j=0}^{D_i} \beta_j (y_i(t+j+1) - y_i(t+j)) = \sum_{j=0}^{D_i} \beta_j C_i (P-I) P^j x(t) = 0. \tag{7}$$

Using the final value theorem in (6) and some simple algebra we find the consensus value $\phi\mathbf{1}$

$$\phi\mathbf{1} = \lim_{z\to 1}(z-1) Y_i(z) = \frac{H(1)}{p_i(1)} = \frac{y_{D_i}^{\mathsf{T}} \beta}{\mathbf{1}^{\mathsf{T}} \beta}, \tag{8}$$

where $y_{D_i}^{\mathsf{T}} = [\,y_i(0) \;\; y_i(1) \;\; \ldots \;\; y_i(D_i)\,]$ and $\beta_{(D_i+1)\times 1}$ is the vector of coefficients of $p_i(z)$.

We denote the Hankel matrix [32, 30]

$$\Gamma\left\{y_i(0),\, y_i(1),\, \ldots,\, y_i\left(\left\lceil \frac{k+1}{\ell+1}\right\rceil + k - 1\right)\right\} = \begin{bmatrix} y_i(0) & y_i(1) & \ldots & y_i(k) \\ y_i(1) & y_i(2) & \ldots & y_i(k+1) \\ \vdots & \vdots & \ldots & \vdots \\ y_i\left(\left\lceil \frac{k+1}{\ell+1}\right\rceil - 1\right) & y_i\left(\left\lceil \frac{k+1}{\ell+1}\right\rceil\right) & \ldots & y_i\left(\left\lceil \frac{k+1}{\ell+1}\right\rceil + k - 1\right) \end{bmatrix}.$$

Node i then stores $y_i(t)$ $(t = 0, 1, \ldots)$ in its own memory and recursively builds up the Hankel matrix $H^k_{i,\ell}$, as

$$H^k_{i,\ell} = \Gamma\left\{y_i(1) - y_i(0),\, y_i(2) - y_i(1),\, \ldots,\, y_i\left(\left\lceil \frac{k+1}{\ell+1}\right\rceil + k\right) - y_i\left(\left\lceil \frac{k+1}{\ell+1}\right\rceil + k - 1\right)\right\}, \tag{9}$$

where $\lceil x\rceil$ is the nearest integer not less than x, and $H^k_{i,\ell}$ always has more rows than columns. Node i then calculates the rank of $H^k_{i,\ell}$ and increases the dimension k until $H^k_{i,\ell}$ loses column rank and stores the first defective Hankel matrix $H^K_{i,\ell}$. Here K is a good estimation of D_i. Node i then calculates the normalized kernel $\beta = (\beta_0,\cdots,\beta_{K-1}, 1)^{\mathsf{T}}$ of $H^K_{i,\ell}$, i.e. $H^K_{i,\ell}\beta = 0$ according to equation (7). Once β is obtained and combined with the previously memorized $[y_i(0)\ \ y_i(1)\ \ \ldots\ \ y_i(K)]$, node i can use the final value theorem in equation (8) and calculate the final global synchronized value

$$\phi\mathbf{1} = \frac{[y_i(0)\ \ y_i(1)\ \ \ldots\ \ y_i(K)]\beta^{\mathsf{T}}}{\mathbf{1}}\,\beta, \tag{10}$$

within $N_{i,\ell} = \left\lceil \frac{K+1}{\ell+1}\right\rceil + K$ iteration steps, which is defined as the minimal memory length (MML). Here, $\phi = \mu x(0)$. All nodes then propel themselves toward the calculated destination ϕ. Given the observed node i and its ℓ monitored nodes, the synchronizing time of the method can thus be quantified using $N_{i,\ell}$. In fact, this synchronizing time $N_{i,\ell}$ denotes the minimal necessary recording length of the historical trajectory of the observed node i and the ℓ monitored nodes. To quantify the synchronization speed of the routine process, we directly simulate the dynamics (1) and define the minimal convergence steps (MCS) M as when the state difference of all node pairs, e.g. $\sum_{i>j}|x_i - x_j|$, drops below a small threshold δ (here we set $\delta = 10^{-3}$). The smaller the value of M, the more rapid the synchronization. The convergence of Hankel matrix-based iterations can be refereed to [30]. Once one node has achieved the final consensus value, it will simultaneously send the final synchronization value to all the other nodes of the network.

We first consider the Erdös–Rényi (ER) [33], the Barabási–Albert (BA) [34], and the Watts–Strogatz (WS) [35] models. In an ER network, node pairs are connected with a probability ρ. Initially a BA network is a small clique of m nodes, and at each time step a single node is added with m edges connecting to existing nodes. The probability of selecting an existing node is proportional to its degree. WS network is an one-dimensional lattice in which each node connects to z neighbors, and each edge has a constant probability p of being rewired. The average degree of an BA network is approximately $2m$, and the average degree of an WS network is z. We generate $S = 100$ networks of size $N = 100$ for each network model. In each network, we independently pick up an observed node i and its ℓ neighbors for $R = 100$ times. To compare the local observation scheme and the existing synchronization method (1), we define the average minimal memory length (AMML) and the average minimal convergence steps (AMCS) for these S networks as $\overline{N_\ell} = \frac{1}{R\times S}\sum_{i=1}^{R}\sum_{j=1}^{S} N^j_{i,\ell}$ and $\overline{M} = \frac{1}{S}\sum_{j=1}^{S} M_j$, respectively. Here, $N^j_{i,\ell}$ and M_j are the MML and MCS of the jth network $(j = 1, 2, \cdots, S)$ with associate output matrix C_i $(i = 1, 2, \ldots, R)$, respectively. As shown in table 1, even when $\ell = 0$, namely we know only the record of the observed node, the synchronization speed of this method is much faster than the routine process, as indicated by how much smaller the value of $\overline{N_0}$ is than $\overline{M}$. In addition, $\overline{N_\ell}$ decreases when ℓ increases, suggesting that the synchronization can be further accelerated by including the monitored nodes.

Significantly, the state of a node is directly affected by its neighbors, the state of the neighbors are in turn affected by their neighbors, and so on. Accordingly, the average number of steps required for the influence from a randomly selected node to reach another randomly selected node is equal to the average distance $\langle d\rangle$. Thus the synchronization time is strongly dependent on $\langle d\rangle$. Figure 1 shows the relationship between the synchronization time and $\langle d\rangle$ in both our method and the routine process. The synchronization time $\overline{M}$ required by routine method is much longer than $\overline{N_\ell}$ even when $\ell = 0$, and relationships $(\overline{M},\ \langle d\rangle)$ and $(\overline{N_0},\ \langle d\rangle)$ both approximately fit a linear function, but the increasing rate of $\overline{M}$ is much larger than that of $\overline{N_0}$. We thus expect that in networks with a larger $\langle d\rangle$ the advantage enjoyed by $\overline{N_0}$ will become even more significant.

We next consider a significant extension towards nonlinear dynamics. Take the prestigious Kuramoto model [36] as example, for any node i $(i = 1, 2, \cdots, N)$

Table 1. AMML and AMCS of $S = 100$ ER, BA and WS networks with size $N = 100$. The initial state of each node is randomly selected in the range $[-2, 2]$. The rewiring probability of WS networks is set as $p = 0.1$.

	AMML ($\ell = 0$)	AMML ($\ell = 1$)	AMML ($\ell = 2$)	AMML ($\ell = 4$)	AMCS
ER($\rho = 0.1$)	52.34	44.63	35.80	33.85	281.51
ER($\rho = 0.2$)	18.92	11.14	10.21	9.29	74.61
BA($m = 3$)	77.17	67.22	62.86	63.92	507.39
BA($m = 5$)	44.49	26.18	23.73	22.79	233.22
WS($z = 6$)	78.22	69.34	62.93	46.78	582.31
WS($z = 10$)	37.83	19.86	16.54	13.98	233.20

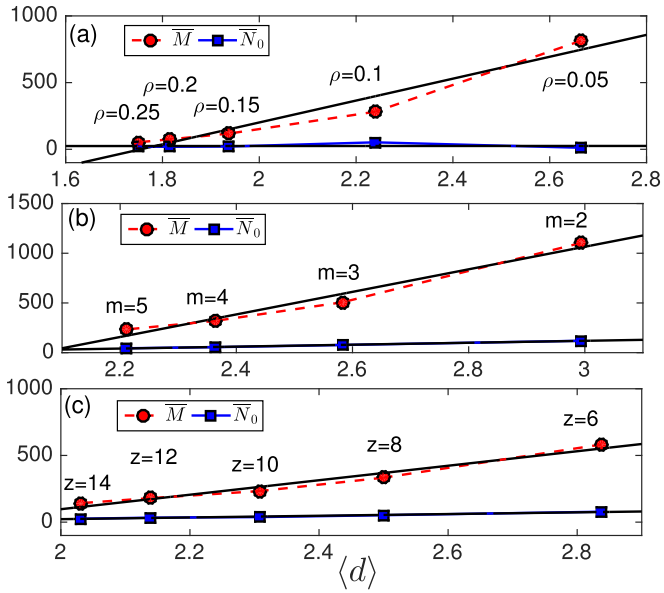

Figure 1. The synchronization time, $\overline{M}$ and $\overline{N}_0$, versus the average distance $\langle d \rangle$ for ER (a), BA (b) and WS (c) networks. The initial state of each node in these $S = 100$ ER, BA and WS networks is randomly selected in the range $[-2, 2]$ and set as fixed. The average distance of a directed network with N nodes is defined as $\langle d \rangle = \frac{1}{N(N-1)} \sum_{i \neq j} d(i, j)$, where $d(i, j)$ is the distance from node i to node j. If the network is not strongly connected, $\langle d \rangle = \infty$. The black lines represent the linear fitting by the least squares estimation. The network size is set as $N = 100$, and the rewiring probability is set as $p = 0.1$ for WS networks.

$$\theta_i(t+1) = \theta_i(t) + \epsilon \left[\omega_i + \frac{K}{\sum_j b_{ij}} \sum_j b_{ij} \sin(\theta_j(t) - \theta_i(t)) \right], \tag{11}$$

where j runs over all i's neighbors. The system is composed of N oscillators, with phases θ_i, natural frequencies ω_i, coupling K, sampling period $\epsilon = 1/N$, and edge betweenness b_{ij} [37]. It is observed that the errors of the proposed Hankel matrix-based prediction method keep less than 0.015 (see figure 2) and quickly settles down to almost zero. The AMML $\overline{N}_0$ are also approximately proportional to the average distances (see figure 3). This result is consistent to the linear case (see figure 1), suggesting the generality of the proposed prediction method.

We further test the validity of our algorithm for ten disparate real-world networks (using equation (1) as example dynamics), with the last two are directed. (i) *Karate*—a friendship network consisting of 34 members of a karate club [38]. (ii) *Power*—a US electric power grid built in the early 1960s [39]. (iii) *Dolphin*—a social contact network among 62 bottlenose dolphins in a community at Doubtful Sound, New Zealand [40]. (iv) *Lesmis*—the network of fictional characters in Victor Hugo's novel *Les Miserables*, where each edge denotes the co-appearance of the two corresponding characters [41]. (v) *Polbooks*—the network of books about recent US politics in Amazon.com, where edges represent co-purchasing relations [42]. (vi) *Football*—the network of American football games between Division IA colleges during the regular season in Fall 2000, where each node represents a team and two teams are connected if they have regular seasonal games [43]. (vii) *FFHI*—the face-to-face human interaction network in a school [44]. (viii) *Corporate*—an European corporate community in which nodes represent firms and two firms are connected if they share at least one manager or director [45]. (ix) *Bison*—the dominance relationships among American bisosn in 1972 on the National Bison Range in Moiese,

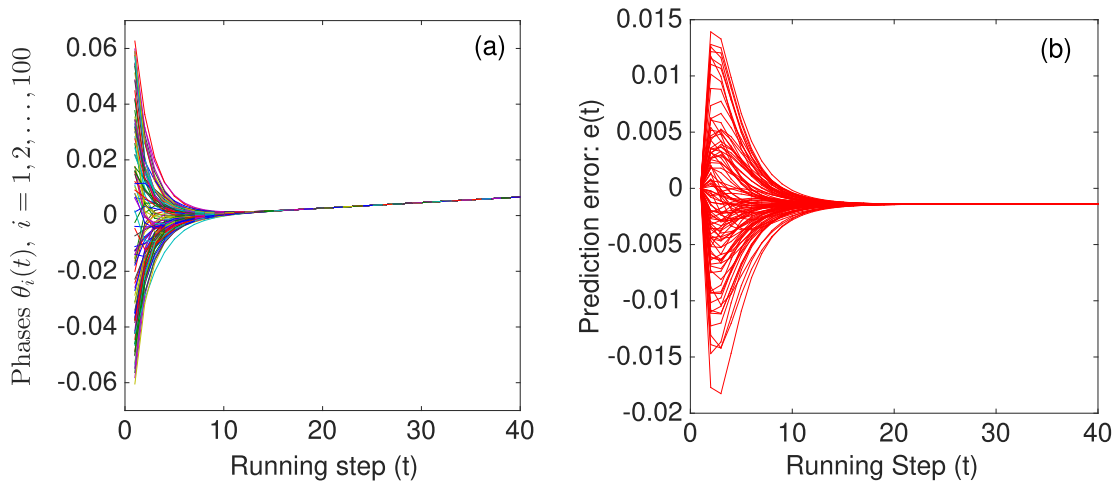

Figure 2. Temporal evolution of (a) $\theta_i(t)$ and (b) prediction errors $e(t) = \theta_i(t) - \hat{\theta}_i(t)$ for ER networks with link probability $\rho = 0.15$. The parameters are $N = 100, K = 50, \epsilon = 1/N$, $\omega_i = \frac{1}{50}$ and $\theta_i(0)$ randomly selected from $\left[-\frac{\pi}{50}, \frac{\pi}{50}\right]$.

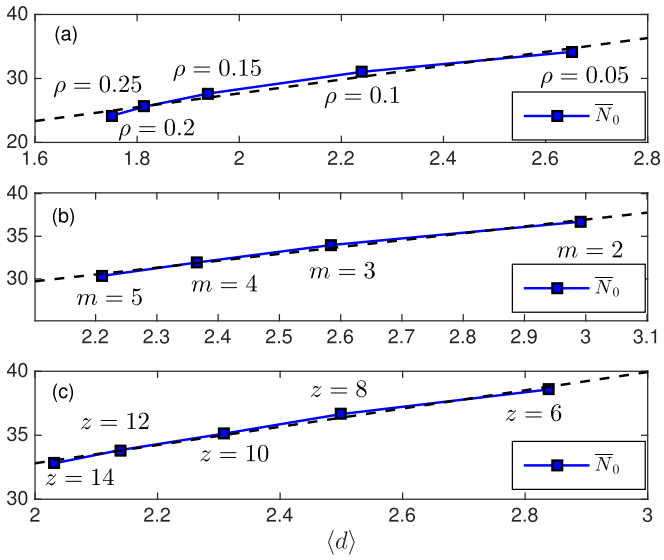

Figure 3. $\overline{N}_0$ versus $\langle d \rangle$ for (a) ER, (b) BA and (c) WS networks for dynamics in equation (11). Here, the initial state values $\theta_i(0)$, $i = 1, 2, \ldots, S$, are the same as in the simulation for figure 2, and $\overline{N}_0 = \frac{1}{R \times S}\sum_{i=1}^{R}\sum_{j=1}^{S} N_{i,0}^{j}$. The black lines represent the linear fitting by the least squares estimation. The rewiring probability for WS networks is $p = 0.1$ and other parameters are the same as figure 2.

Montana, where each node denotes a bison and each directed edge represents the dominance relationship [46]. (x) *Highschool*—the network of friendships among boys in a high school in Illinois, where a directed link from *i* to *j* means *i* identifies *j* as his friend [46]. Table 2 provides the structural statistics and synchronization time of the ten real-world networks. When we compare the last two columns, it is observed that the present method produces much faster synchronization. In most cases our method is ten times faster than the routine procedure even without monitored nodes. Figure 4 shows that despite some oscillations, the synchronization time $\overline{N}_\ell$ further decreases as ℓ increases, indicating additional benefit when monitored nodes are introduced. The oscillations in figure 4 lie in the specific topology of each real network. Similar to artificial network models, as shown in figure 5, both $\overline{N}_0$ and M increase with $\langle d \rangle$ for real-world networks. Figure 5 further shows linear fittings for visual guidance, where it is clear that the increasing rate of M is much larger than $\overline{N}_0$. Thus extensive experimental analyses of disparate real-world networks once again demonstrate the results obtained from network models, i.e. (i) our method speeds up synchronization, (ii) monitoring more neighbors further accelerates synchronization, and (iii) the synchronization time is positively correlated with $\langle d \rangle$, while the present method grows much more slowly. Note that we do NOT need a CPU for each node, but a programmable chip with limited functions like information storage, and neighboring communication. Besides, a useful monitored

Table 2. Topological features and synchronization time for the 10 real networks under consideration. N, E, $\langle k\rangle$ and $\langle d\rangle$ represent the number of nodes, the number of edges, the average degree and the average distance, respectively. The former 8 networks are undirected while the last two are directed. As clearly observed from this table, AMML $\overline{N_0}$ is much smaller than M (the MCS of the real network), indicating the advantage of the present Hankel matrix-based prediction method.

	N	E	$\langle k\rangle$	$\langle d\rangle$	$\bar{N}_0$	M
Karate	34	78	4.588	2.408	55.93	501
Power	57	78	2.737	4.954	214.04	4247
Dolphin	62	159	5.129	3.357	133.61	1595
Lesmis	77	254	6.597	2.641	103.33	2328
Polbooks	105	441	8.4	3.079	130.94	2285
Football	115	613	10.661	2.508	57.04	527
FFHI	180	2239	24.667	2.148	49.68	621
Corporate	197	801	8.132	2.106	160.36	3712
Bison	26	314	12.0769	1.571	18.92	161
Highschool	70	366	5.229	∞	157	768

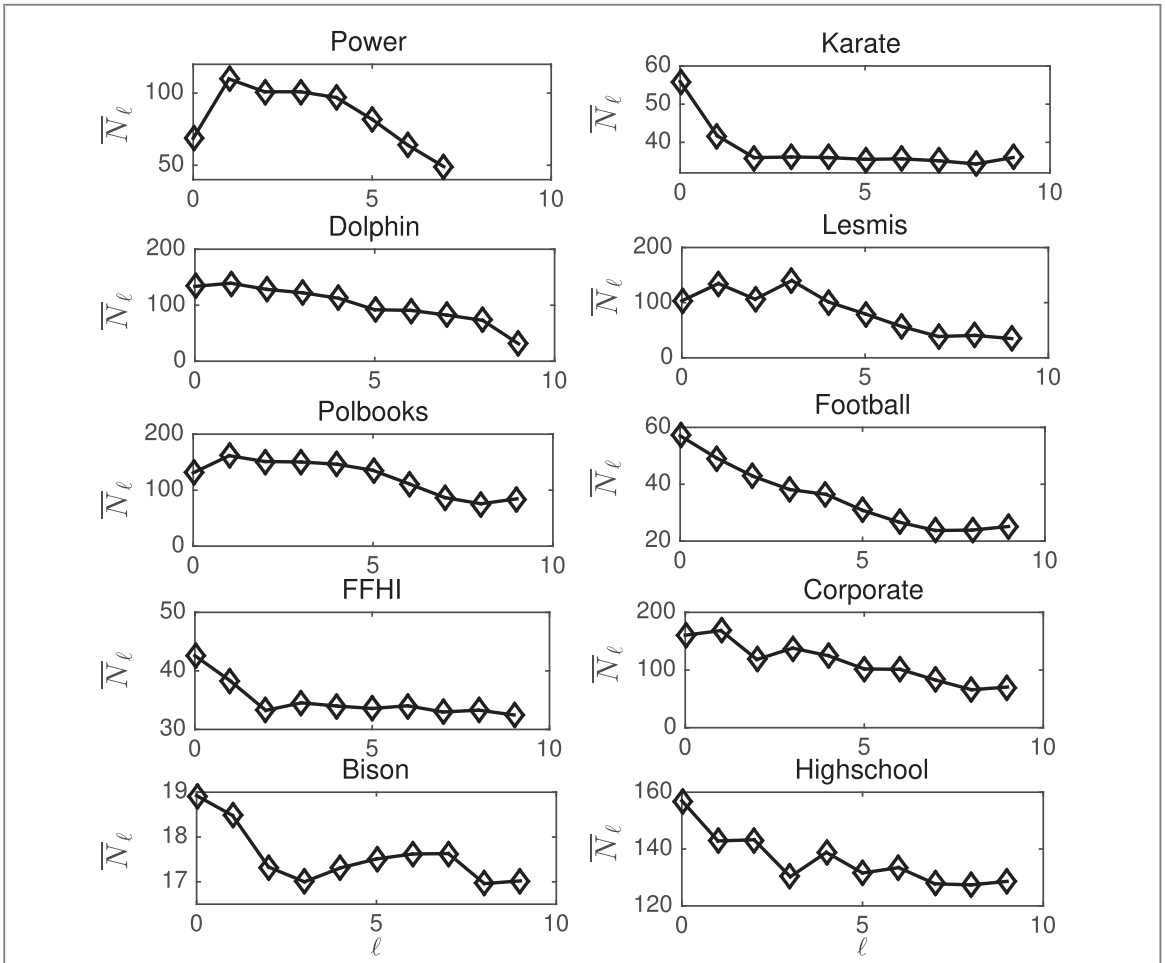

Figure 4. AMML versus the number of monitored neighbors ℓ for the 10 real networks. Here, AMML is defined as $\overline{N}_\ell = \frac{1}{R}\sum_{i=1}^{R} N_{i,\ell}$, where $N_{i,\ell}$ is the MML of the real network with associate output matrix C_i ($i = 1, 2, \ldots, R$). Given ℓ, we only select those nodes with degree no less than ℓ as observed nodes to implement the simulation and then get average MML over these nodes. Since the maximal degree of power is 7, the corresponding maximal ℓ in the first plot is 7 as well.

node selection technique is picking out the ones with short average distances to all the other nodes, like the central node of a star-shaped graph. This will reduce the broadcasting time from them to the monitored nodes.

[7] The y-axis units of figures 1 and 3 are the same, which are set as one step (or one epoch). Hence, synchronization time = epochs × sampling period, and the synchronization times of the present and previous methods are comparable by assuming identical sampling period.

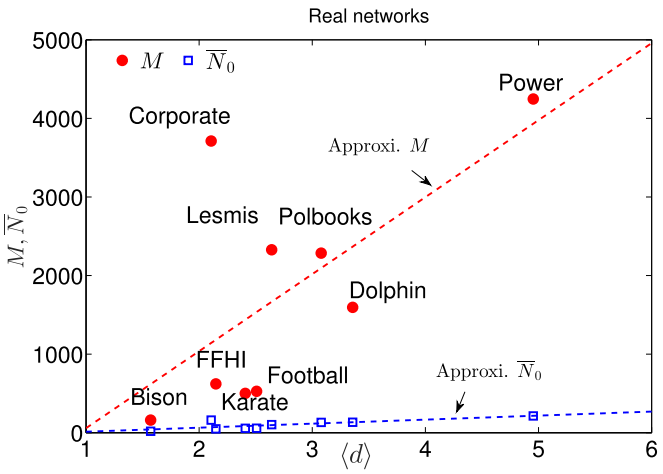

Figure 5. The synchronization time, M and $\overline{N}_0$, versus the average distance $\langle d \rangle$ for the 9 real networks except highschool whose average distance is infinite. The red circles and blue squares denote the synchronization times by the routine procedure and the present method, respectively. The red and blue lines represent the linear fittings for data points by the least squares estimation.

3. Conclusions

In summary, we have found a mechanism that leads to the ultrafast synchronization while only requires the historical dynamical trajectory of the observed node. In a networked dynamical system, the state of a node is directly affected by its neighbors, who are directly affected by their neighbors, and so on. Thus the state of a node will affect and be affected by all other nodes after a sufficiently long period of time. Our major contribution here is successfully realizing this theoretical possibility by applying Hankel matrix analysis. The different choices of the observed and monitored nodes will affect the synchronization speed. An intuitive idea is to select the node with the largest centrality as the observed node since it usually locates in the central position with shorter average distance to others [47]. The design of efficient algorithms to accurately locate the optimal observed node will be our future work.

Compared to the information propagation [21, 26] and predictive protocol [27–29], the present mechanism requires little intelligence from most individuals but a higher level of intelligence from the observed node. This includes both the memory to store the historical dynamical trajectory and the ability to analyze this trajectory. In a biological system, it is unlikely that a leader would use a Hankel matrix-based method to figure out the future travel direction to lead the flock. Instead, we believe that this mechanism will have significant applications in engineering systems. A group with one super leader is unlikely in the biological world but easy to be designed and implemented in artificial systems. A distributed sensor network in which each sensor communicates and interacts with its neighbors must be able to align and move together in such scenarios as field investigation or battleground detection. Our proposed mechanism does not require a large number of low-intelligence sensors but only one sensor with moderate memory and computational capacity. Modern information technology (in particular, the rapid development of intelligent hardware) allows us to produce a smart sensor with a sufficiently long memory and the ability to analyze the Hankel matrix. Thus this smart sensor could predict the future global state of networked dynamics and shape the consensus of the entire sensor group.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China under Grant Nos. U1713203, 61673189, 61751303, 91748112, 61433014 and 61703175. T.Z. acknowledges the Science Promotion Programme of UESTC (No. Y03111023901014006).

ORCID iDs

Hai-Tao Zhang https://orcid.org/0000-0002-8819-8829

References

[1] Buck J B 1988 *Q. Rev. Biol.* **63** 265

[2] Chen C, Liu S, Shi X Q, Chaté H and Wu Y 2017 *Nature* **542** 210
[3] Kapitaniak M, Czolczynski K, Perlikowski P, Stefanski A and Kapitaniak T 2012 *Phys. Rep.* **517** 1
[4] Néda Z, Ravasz E, Brechet Y, Vicsek T and Barabási A L 2000 *Nature* **403** 849
[5] Strogatz S H 2003 *Sync: The Emerging Science of Spontaneous Order* (New York: Hyperion)
[6] Parris J and Edelstein-Keshet L 1999 *Science* **284** 99
[7] Loannou C C, Guttal V and Couzin I D 2012 *Science* **337** 1212
[8] Weimerskirch H, Martin J, Clerquin Y, Alexandre P and Jiraskova S 2001 *Nature* **413** 697
[9] Augugliaro F *et al* 2014 *IEEE Control Syst. Mag* **34** 46
[10] Rubenstein M, Cornejo A and Nagpal R 2014 *Science* **345** 795
[11] Arenas A, Diaz-Guilera A, Kurths J, Moreno Y and Zhou C 2008 *Phys. Rep.* **469** 93
[12] Vicsek T and Zafeiris A 2012 *Phys. Rep.* **517** 71
[13] Vicsek T, Czirok A, Ben-Jacob E, Cohen I and Shochet O 1995 *Phys. Rev. Lett.* **75** 1226
[14] Couzin I D, Krause J, James R, Ruxton G D and Franks N R 2002 *J. Theor. Biol.* **218** 1
[15] Barahona M and Pecora L M 2002 *Phys. Rev. Lett.* **89** 054101
[16] Nishikawa T, Motter A E, Lai Y C and Hoppensteadt F C 2003 *Phys. Rev. Lett.* **91** 014101
[17] Herbert-Read J E, Perna A, Mann R P, Schaert T M, Sumpter D J T and Ward A J W 2011 *Proc. Natl Acad. Sci. USA* **108** 18726
[18] Nagy M, Akos Z, Biro D and Vicsek T 2010 *Nature* **464** 890
[19] Buhl J *et al* 2006 *Science* **312** 1402
[20] Couzin I D 2007 *Nature* **445** 715
[21] Attanasi A *et al* 2014 *Nat. Phys.* **10** 691
[22] Zhang J *et al* 2009 *Physica* A **399** 1237
[23] Gao J, Chen Z, Cai Y and Xu X 2010 *Phys. Rev.* E **81** 041918
[24] Gao J, Havlin S, Xu X and Stanley H E 2011 *Phys. Rev.* E **84** 046115
[25] Xu X K, Kattas G D and Small M 2012 *Phys. Rev.* E **85** 026120
[26] Cavagna A *et al* 2015 *J. Stat. Phys.* **158** 601
[27] Zhang H T, Chen M Z, Zhou T and Stan G B 2008 *EPL* **83** 40003
[28] Zhang H T, Chen M Z, Stan G B, Zhou T and Maciejowski J M 2008 *IEEE Circuits Syst. Mag.* **8** 67
[29] Zhang H T, Chen M Z and Zhou T 2009 *Phys. Rev.* E **79** 016113
[30] Yuan Y, Stan G B, Shi L, Barahona M and Goncalves J 2013 *Automatica* **49** 1227
[31] Olfati-Saber R and Murray R 2004 *IEEE Trans. Autom. Control* **49** 1520
[32] Partington J R 1988 *An introduction to Hankel operators* (Cambridge: Cambridge University Press)
[33] Erdös P and Rényi A 1960 *Publ. Math. Inst. Hung. Acad. Sci.* **5** 17
[34] Barabási A L and Albert R 1999 *Science* **286** 509
[35] Watts D J and Strogatz S H 1998 *Nature* **393** 440
[36] Acebrón J A, Bonilla L L, Vicente C J P, Ritort F and Spigler R 2005 *Rev. Mod. Phys.* **77** 137
[37] Freeman L C 1977 *Sociometry* **40** 35
[38] Zachary W W 1977 *J. Anthropol. Res.* **33** 452
[39] http://labs.ece.uw.edu/pstca/
[40] Lusseau D, Schneider K, Boisseau O J, Haase P, Slooten E and Dawson S M 2003 *Behav. Ecol. Sociobiol.* **54** 396
[41] Newman M E J and Girvan M 2004 *Phys. Rev.* E **69** 026113
[42] http://www-personal.umich.edu/~mejn/netdata/
[43] Girvan M and Newman M E J 2002 *Proc. Natl Acad. Sci. USA* **99** 7821
[44] Starnini M, Baronchelli A and Pastor-Satorras R 2013 *Phys. Rev. Lett.* **110** 168701
[45] Kogut B M 2012 *The Small Worlds Of Corporate Governance* (Cambridge, MA: MIT Press)
[46] Coleman J S 1964 *Introduction to Mathematical Sociology* (New York: The Free Press)
[47] Lü L, Chen D, Ren X L, Zhang Q M, Zhang Y C and Zhou T 2016 *Phys. Rep.* **650** 1