

Non-linear consensus dynamics on temporal hypergraphs with random noisy higher-order interactions

YILUN SHANG[†]

*Department of Computer and Information Sciences, Northumbria University, Newcastle upon Tyne
NE1 8ST, UK*

[†]Corresponding author. Email: yilun.shang@northumbria.ac.uk

[Received on 29 November 2022; editorial decision on 28 February 2023; accepted on 2 March 2023]

Complex networks encoding the topological architecture of real-world complex systems have recently been undergoing a fundamental transition beyond pairwise interactions described by dyadic connections among nodes. Higher-order structures such as hypergraphs and simplicial complexes have been utilized to model group interactions for varied networked systems from brain, society, to biological and physical systems. In this article, we investigate the consensus dynamics over temporal hypergraphs featuring non-linear modulating functions, time-dependent topology and random perturbations. Based upon analytical tools in matrix, hypergraph, stochastic process and real analysis, we establish the sufficient conditions for all nodes in the network to reach consensus in the sense of almost sure convergence and \mathcal{L}^2 convergence. The rate of consensus and the moments of the equilibrium have been determined. Our results offer a theoretical foundation for the recent series of numerical studies and physical observations in the multi-body non-linear dynamical systems.

Keywords: consensus; higher-order network; hypergraph; nonlinear dynamics; random noise; temporal network.

1. Introduction

Network theory has gained considerable prominence over the past decades in the physics community. This is partly due to its capability to represent and solve real-world problems, offering substantial insights into topological and dynamical aspects of networked complex systems [1]. By investigating the system-level characterization of network interactions between nodes and edges, universal structural features such as heavy-tailed distributions [2] and small-world effect [3] and emergent collective behaviours including synchronization and consensus in myriad complex biological, physical and social systems have been revealed [1, 4].

The network architecture underlying a complex system is naturally represented by graphs with interactions between nodes limited to dyadic connections. This traditional representation falls short to capture group interactions involving three or more nodes simultaneously, which very recently triggers a significant transition towards the high-order networks [5–8]. In fact, higher-order interactions beyond the pairwise scheme are ubiquitous in many real-world complex systems ranging from neuronal activity patterns in brains [9, 10] to cooperation evolution in social networks [11], from multi-protein complexes prediction [12] to species coexistence in ecological systems [13, 14]. Taking this into account, group interactions lend essential descriptive power to the network theory in that it not only extracts the higher-order organization of complex networks [15] but also helps to clarify intrinsic non-linearity in system dynamics. For example, some information diffusion mechanisms such as opinion dynamics models [16] and threshold models in social contagions [17] rely on the group decision among neighbours of an

individual. These interactions, however, can be factorized as a linear combination of dyadic connections between the individual and each of its neighbours and therefore are not inherently non-linear. Moreover, it is shown in [18, 19] that group interactions in any linear time-invariant consensus dynamics can always be decomposed as an aggregation of pairwise couplings over a rescaled network and hence do not encode genuine higher-order effects.

From a mathematical perspective, higher-order networks often take the form of hypergraphs and simplicial complexes [6]. A simplicial complex is comprised of filled cliques, that is, simplexes, with different dimensions. For instance, 1-simplex is an edge and 2-simplex is a filled triangle, etc. Simplicial complexes are constructed by gluing simplexes along their nodes, edges and faces. This formulation gives simplicial complexes an edge in network geometric representation [20, 21]. The simplicial structure has shown to impact the dynamical processes such as synchronization [22, 23], percolation [24] and social evolutionary processes [25] significantly compared to classical networks with diadic ties. The hypergraph structure on the other hand is an extension to graphs, which contains hyperedges formed by sets of two or more nodes. In contrast to simplicial complexes, a higher-order interaction captured by hypergraphs does not need the existence of all lower-order interactions. Hypergraphs have also been used to model group interactions for some important network dynamics including consensus processes [18, 19], random walks [26] and epidemic spreading [27]. A higher-order interaction framework via directed hypergraphs has been developed in [16].

In parallel to the above development, extending the static network structures to temporal networks is of great interest in the network science community. In a temporal network, nodes and edges may be created and destroyed over time mimicking many real-life complex networks, which have shown fundamentally different dynamical behaviours and emergent phenomena in comparison to artificially aggregated static network substrates [28, 29]. Some initial studies have already explored the effect of the interplay between time order and higher-order interactions on network-based evolving processes such as synchronization [30], consensus [31], diffusion [32] and social motif formation [33]. For example, it is observed from numerical studies [31] that non-linear consensus processes with 3-way interactions are slowed down and that local majority in the early stage has a stronger influence on the final consensus value of the system as compared to networks with pairwise connections.

Up till now, the dynamical collective behaviours stemming from higher-order interactions in complex systems have mostly been dealt with through numerical simulations and semi-analytical methods such as mean-field analysis [5, 6]. Moreover, in the existing literature concerning higher-order interactions, the complex systems are generally assumed to be noise free. Understanding the influence of random noises, which are usually inevitable in all real dynamical processes over networked complex systems, remains challenging and is beyond any rigorous treatment in the context of higher-order networks so far.

Motivated by the above consideration, we follow the line of research in [18, 19, 31] to study nonlinear consensus processes over hypergraphs. Specifically, we investigate a discrete-time non-linear consensus process over temporal hypergraphs and the system is subject to random noises with general non-linear modulating functions. We introduce a mathematical framework to understand the topological conditions underpinning the consensus process. The analytical methods are mainly based on a combination of stochastic matrix analysis, real analysis, hypergraph, martingale convergence and the Lyapunov stability theory. We find that consensus can be achieved in the sense of both almost sure convergence and \mathcal{L}^2 (i.e. mean square) convergence when the underlying hypergraph is jointly uniformly connected. The final consensus value can also be determined in a stochastic sense. In particular, our results offer a mathematical proof for the numerical observations and experiments performed in [18, 19, 31]. It is hoped that the framework may be extended to investigate other relevant dynamical processes paving a theoretical avenue to understanding higher-order interactions.

The remainder of the article is organized as follows. In Section 2, we introduce the consensus model over temporal hypergraphs. The main results are given in Section 3, with the technical proofs and further extensions presented in Section 4. We conclude the article in Section 5.

2. Model

Let \mathbb{N} be the set of non-negative integers. At any time step $t \in \mathbb{N}$, we consider n dynamical agents encoded in the temporal hypergraph structure $\mathcal{H}(t) = (V, E(t), A(t))$, where $V = \{1, 2, \dots, n\}$ represents the node set, and $E(t)$ is the set of hyperedges containing some triples of the form $\{i, j, k\}$ with $i, j, k \in V$. The hypergraph $\mathcal{H}(t)$ is called a 3-uniform hypergraph in hypergraph theory [34] since each hyperedge is a subset of V with cardinality 3. Given $t \in \mathbb{N}$, the topology of $\mathcal{H}(t)$ is captured by the adjacency tensor $A(t) = (a_{ijk}(t)) \in \mathbb{R}^{n \times n \times n}$ with $a_{ijk}(t) = 1$ if $\{i, j, k\} \in E(t)$ and $a_{ijk}(t) = 0$ otherwise. By convention, $a_{ijk}(t) = 0$ if the edge becomes degenerate, that is, $i = j$, $i = k$ or $j = k$.

Two nodes i and j in $\mathcal{H}(t)$ are adjacent if they appear in the same hyperedge of $E(t)$. i and j are said to be connected by a path if there exists a sequence of nodes starting from i ending at j , and every two contiguous nodes in the sequence are adjacent. The hypergraph $\mathcal{H}(t)$ is called connected if any two nodes in V are connected by a path [34]. In our temporal hypergraph framework, we only assume that some piecewise unions of hypergraphs over the time axis are connected. Formally, we call $\mathcal{H}(t)$ jointly uniformly connected with bound $\ell > 0$ if there is an infinite sequence of time steps $0 = t_0 < t_1 < t_2 < \dots < t_l < \dots$ satisfying $t_l - t_{l-1} \leq \ell$ and $\cup_{s=t_l}^{t_{l+1}-1} \mathcal{H}(s)$ being connected for any $l \in \mathbb{N}$.

The set of equations governing the dynamics of n nodes is

$$x_i(t+1) = x_i(t) + \gamma u_i(t), \quad i \in V, t \in \mathbb{N}, \quad (2.1)$$

where $\gamma > 0$ is a constant gain, $x_i(t) \in \mathbb{R}$ is the state of node i and $u_i(t) \in \mathbb{R}$ is the interaction function delineating the influence of other nodes over node i at time t . We consider the following interaction function

$$\begin{aligned} u_i(t) = & \sum_{j,k=1}^n a_{ijk}(t) [(x_j(t) - x_i(t)) + (x_k(t) - x_i(t))] \\ & + \sum_{j,k=1}^n a_{ijk}(t) f(|x_j(t) - x_k(t)|) \\ & \cdot [g(x_j(t) - x_i(t)) w_{ijk}(t) + g(x_k(t) - x_i(t)) w_{ijk}(t)], \end{aligned} \quad (2.2)$$

where the two modulating functions f and $g : \mathbb{R} \rightarrow \mathbb{R}$ satisfy the following Assumption 1 and the random noise $w_{ijk}(t)$ satisfies Assumption 2 below.

ASSUMPTION 1 There exists two constants $\beta_1, \beta_2 > 0$ such that $|f(x)| \leq \beta_1$ and $|g(x)| \leq \beta_2|x|$ for any $x \in \mathbb{R}$.

ASSUMPTION 2 The random variables $\{w_{ij}(t)\}_{i,j \in V, t \in \mathbb{N}}$ are independent and identically distributed for different i, j and t . Moreover, $E(w_{ij}(t)) = 0$ and $\text{Var}(w_{ij}(t)) = 1$, where E and Var mean expectation and variance, respectively. We assume that $w_{ijk}(t)$ is an independent copy of $w_{ij}(t)$ (or equivalently, $w_{ik}(t)$).

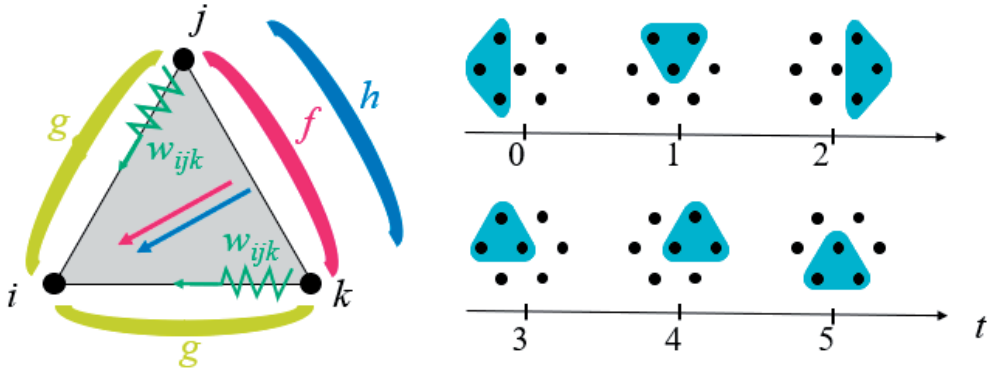


FIG. 1. Right: Schematic of a hyperedge $\{i, j, k\}$ with interaction function on node i . Left: Temporal evolution with a jointly uniformly connected topology $\mathcal{H}(t)$ with $\ell = 3$.

The perturbation $w_{ijk}(t)$ represents a random noise associated with the hyperedge $\{i, j, k\} \in E(t)$. A schematic of the interaction and evolution of $\mathcal{H}(t)$ is shown in Fig. 1. The modulating function f is bounded and captures the influence of state differences between nodes j and k on node i . To model the group reinforcement effect on consensus dynamics, f is naturally chosen as a monotonic function such as exponential functions and the Heaviside function in [18, 19, 31]. If f is decreasing, it is in line with the bounded confidence opinion model [35] and the flocking models [36] in biology, where closer states exert a stronger cohesive force among each other. Our assumption on the modulating function g is mild. A typical choice is a linear $g(x) = \beta x$ for some constant $\beta > 0$. The noises are assumed to be standardized random variables for the ease of presentation.

It is also worth noting that in the above interaction function (2.2), for simplicity, we deliberately do not include a modulating function $h(|x_j(t) - x_k(t)|)$ in the deterministic coupling part like

$$\sum_{j,k=1}^n a_{ijk}(t) h(|x_j(t) - x_k(t)|) [(x_j(t) - x_i(t)) + (x_k(t) - x_i(t))], \quad (2.3)$$

and our main focus here is on dealing with random noisy higher-order interactions. However, it can be seen below, if such a function h is bounded and positive, the results of this article still hold; c.f. Corollaries 2 and 3 in Section 4. The decreasing exponential function considered in [18], for instance, is a special candidate for h .

Since the random noises are factored in the dynamics, we are interested in the consensus behaviour of the system (2.1) with interaction function (2.2) in the following two senses.

- Almost sure consensus: There exists some random variable \hat{x} such that the probability $\mathbb{P}(\lim_{t \rightarrow \infty} x_i(t) = \hat{x}) = 1$ holds true for any $i \in V$ and initial configuration $\{x_i(0)\}_{i \in V}$.
- \mathcal{L}^2 consensus: There exists some random variable \hat{x} such that $\lim_{t \rightarrow \infty} \mathbb{E}(|x_i(t) - \hat{x}|^2) = 0$ and $\sup_{t \in \mathbb{N}} \mathbb{E}(|x_i(t)|^2) < \infty$ hold true for any $i \in V$ and initial configuration $\{x_i(0)\}_{i \in V}$.

It is well known that almost sure convergence does not imply \mathcal{L}^2 convergence and vice versa; see for example [37].

3. Results

Here, we present the main theoretical results of this article with detailed proofs deferred to Section 4. For $i \in V$, it follows from (2.1), (2.2) and Assumption 2 that

$$\begin{aligned}
x_i(t+1) &= x_i(t) + 2\gamma \sum_{j,k \in V} a_{ijk}(t)(x_j(t) - x_i(t)) \\
&\quad + 2\gamma \sum_{j,k \in V} a_{ijk}(t)f(|x_j(t) - x_k(t)|)g(x_j(t) - x_i(t))w_{ij}(t) \\
&= x_i(t) + 2\gamma \sum_{j \in V} b_{ij}(t)(x_j(t) - x_i(t)) \\
&\quad + 2\gamma \sum_{j \in V} c_{ij}(t)g(x_j(t) - x_i(t))w_{ij}(t),
\end{aligned} \tag{3.1}$$

where $b_{ij}(t) := \sum_{k \in V} a_{ijk}(t)$ and $c_{ij}(t) := \sum_{k \in V} a_{ijk}(t)f(|x_j(t) - x_k(t)|)$. The system (3.1) can be further rewritten as

$$\begin{aligned}
x_i(t+1) &= \left(1 - 2\gamma \sum_{j \in V} b_{ij}(t)\right)x_i(t) + 2\gamma \sum_{j \in V} b_{ij}(t)x_j(t) \\
&\quad + 2\gamma \cdot (c_{i1}(t), c_{i2}(t), \dots, c_{in}(t)) \\
&\quad \cdot \text{diag}(g(x_1(t) - x_i(t)), g(x_2(t) - x_i(t)), \dots, g(x_n(t) - x_i(t))) \\
&\quad \cdot (w_{i1}(t), w_{i2}(t), \dots, w_{in}(t))^\top,
\end{aligned} \tag{3.2}$$

where $\text{diag}(g_1, g_2, \dots, g_n)$ represents the diagonal matrix with diagonal components g_1, g_2, \dots, g_n , and \top means the matrix transpose. Note that the dimension of the diagonal matrix depends on the diagonal components.

Let $x(t) = (x_1(t), \dots, x_n(t))^\top \in \mathbb{R}^n$ be the state vector. Define the Laplacian matrix $L(t) := (l_{ij}(t)) \in \mathbb{R}^{n \times n}$ with $l_{ii}(t) = \sum_{j \in V} b_{ij}(t)$ for $i \in V$ and $l_{ij}(t) = -b_{ij}(t)$ for $i \neq j$. Furthermore, define the two system matrices $B(t) := I_n - 2\gamma L(t)$, $C(t) := \text{diag}(c_1(t), c_2(t), \dots, c_n(t)) \in \mathbb{R}^{n \times n^2}$, where $I_n \in \mathbb{R}^{n \times n}$ is the identity matrix and $c_i(t) := (c_{i1}(t), c_{i2}(t), \dots, c_{in}(t)) \in \mathbb{R}^{1 \times n}$ for $i \in V$. Let $w(t) := (w_1(t)^\top, w_2(t)^\top, \dots, w_n(t)^\top)^\top \in \mathbb{R}^{n^2 \times 1}$ be the random noise vector, where $w_i(t) := (w_{i1}(t), w_{i2}(t), \dots, w_{in}(t))^\top \in \mathbb{R}^n$ for $i \in V$. Define the modulating matrix $G(t) := \text{diag}(G_1(t), G_2(t), \dots, G_n(t)) \in \mathbb{R}^{n^2 \times n^2}$, where $G_i(t) := \text{diag}(g(x_1(t) - x_i(t)), g(x_2(t) - x_i(t)), \dots, g(x_n(t) - x_i(t))) \in \mathbb{R}^{n \times n}$ for $i \in V$. With these notations, the system of equations in (3.2) can be recast as

$$x(t+1) = B(t)x(t) + 2\gamma C(t)G(t)w(t) \tag{3.3}$$

for $t \in \mathbb{N}$.

Our first result concerns the system behaviour with degenerate noises.

THEOREM 1 Consider the system (3.3) with degenerate noises $w_{ij}(t) \equiv 0$ for all $i, j \in V$ and $t \in \mathbb{N}$. Assume that $\gamma < \left(2 \sup_{t \in \mathbb{N}} \max_{i \in V} \sum_{j,k \in V} a_{ijk}(t)\right)^{-1}$ and $\mathcal{H}(t)$ is jointly uniformly connected with bound ℓ and time sequence $\{t_l\}_{l \in \mathbb{N}}$. Then the following hold:

- (i) There exists a probability vector sequence $\alpha(t) = (\alpha_1(t), \alpha_2(t), \dots, \alpha_n(t))^T \in \mathbb{R}^n$ for $t \in \mathbb{N}$ such that $\sum_{i \in V} \alpha_i(t) = 1$, $\alpha_i(t) \geq \eta$ and

$$\alpha(t+1)^T B(t) = \alpha(t)^T \quad (3.4)$$

for $i \in V$ and $t \in \mathbb{N}$, where $0 < \eta < 1$.

- (ii) $\Gamma(t_{l+1}) \leq (1 - \eta'\gamma)\Gamma(t_l)$ holds for $l \in \mathbb{N}$, where $0 < \eta' < \gamma^{-1}$ and

$$\Gamma(t) := \sum_{i \in V} \alpha_i(t) [x_i(t) - \alpha(t)^T x(t)]^2. \quad (3.5)$$

Both η and η' are absolute constants that do not rely on the gain γ .

Several quick remarks are in order. Firstly, it is easy to see that the upper bound for γ in Theorem 1 always holds if we choose $\gamma < \frac{1}{(n-1)(n-2)}$. Secondly, an important implication of Theorem 1 is that the system (3.3) reaches consensus exponentially fast when noises are absent; see Corollary 1 in Section 4 below for a precise presentation. Moreover, $\{B(t)\}_{t \in \mathbb{N}}$ is a sequence of symmetric stochastic matrices. For any $s \in \mathbb{N}$, we define $F(s, s) = I_n \in \mathbb{R}^{n \times n}$ and

$$F(s+l+1, s) = B(s+l)F(s+l, s) \quad (3.6)$$

for $l \in \mathbb{N}$. Clearly, $\{F(t, s)\}_{t, s \in \mathbb{N}}$ is also a collection of stochastic matrices. Theorem 1(ii) will contribute to the convergence of the general case of system via the quadratic Lyapunov function $\Gamma(t)$ defined in (3.5). Our main random consensus result is the following.

THEOREM 2 Consider the system (3.3) under Assumptions 1 and 2. Suppose that $\gamma < \min \left\{ \left(2 \sup_{t \in \mathbb{N}} \max_{i \in V} \sum_{j, k \in V} a_{ijk}(t) \right)^{-1}, \eta' \cdot (4\xi)^{-1} \right\}$, where $\eta' > 0$ is given in Theorem 1 and $\xi > 0$ is a constant defined by (4.40). Assume that $\mathcal{H}(t)$ is jointly uniformly connected with bound ℓ and time sequence $\{t_l\}_{l \in \mathbb{N}}$. Then, the following hold:

- (i) There exists a random variable \hat{x} such that the almost sure consensus and \mathcal{L}^2 consensus are achieved.
(ii) We have

$$\mathbb{E} \hat{x} = \hat{\alpha}(0)x(0) \quad (3.7)$$

and

$$\text{Var} \hat{x} = 4\gamma^2 \sum_{s=0}^{\infty} \hat{\alpha}(s+1) (\mathbb{E} M(s)) \hat{\alpha}(s+1)^T, \quad (3.8)$$

where $\hat{\alpha}(s) = (\hat{\alpha}_1(s), \hat{\alpha}_2(s), \dots, \hat{\alpha}_n(s)) \in \mathbb{R}^{1 \times n}$ is the left eigenvector of the matrix $\lim_{t \rightarrow \infty} F(t, s)$ with respect to the eigenvalue 1 satisfying $\sum_{i \in V} \hat{\alpha}_i(s) = 1$ and $\hat{\alpha}_i(s) > 0$ for $i \in V$ and $s \in \mathbb{N}$. The matrix $M(s) \in \mathbb{R}^{n \times n}$ is defined as $M(s) := \text{diag}(\mu_1(s), \mu_2(s), \dots, \mu_n(s))$, where $\mu_i(s) = \sum_{j \in V} \left[\left(\sum_{k \in V} a_{ijk}(s) \cdot f(|x_j(s) - x_k(s)|) \right)^2 \cdot g(x_j(s) - x_i(s))^2 \right]$ for $i \in V$.

Theorem 2 shows the random consensus for the system (3.3) and identifies the final equilibrium \hat{x} . The expectation of the equilibrium is determined by the initial configuration of the agents while the variance expressed in (3.8) is a convergent series, which will be clear in the proof below. When considering the generalized system with an additional modulating function $h(\cdot)$ as in (2.3), we can readily extend Theorem 2 to Corollary 3 (see Section 4 below), which is the most flexible result in this article. Under the jointly uniformly connectedness of the hypergraph, Corollary 3 guarantees stochastic consensus when the non-linear system is governed by three types of modulating functions over the hypergraph structure: a positive bounded function h concerning the 3-way coupling strength, a bounded function f concerning the gap among agents in an edge, and a function g with bounded derivative concerning the strength of random noises. Many of the previous frameworks [18, 19, 31] can be regarded as special cases.

REMARK 1 Before proceeding to the technical proof of all these results, we point out that although we have stressed the random noise interpretation, our framework is actually capable of accommodating more general random phenomenon such as random hypergraphs [38, 39]. For instance, we consider the interaction function of (2.1) as

$$u_i(t) = \sum_{j,k=1}^n a_{ijk}(t) \theta_{ijk}(t) h(|x_j(t) - x_k(t)|) \cdot [(x_j(t) - x_i(t)) + (x_k(t) - x_i(t))], \quad (3.9)$$

where $\{\theta_{ijk}(t)\}_{(i,j,k) \in E(t)}$ are independent Bernoulli random variables satisfying $\theta_{ijk}(t) = 1$ with probability $p_{ijk}(t)$ and $\theta_{ijk}(t) = 0$ with probability $1 - p_{ijk}(t)$ for $t \in \mathbb{N}$ and $p_{ijk}(t) \in [0, 1]$. The function (3.9) can be recast as

$$\begin{aligned} u_i(t) &= \sum_{j,k=1}^n a_{ijk}(t) p_{ijk}(t) h(|x_j(t) - x_k(t)|) [(x_j(t) - x_i(t)) + (x_k(t) - x_i(t))] \\ &\quad + \sum_{j,k=1}^n a_{ijk}(t) h(|x_j(t) - x_k(t)|) \\ &\quad \cdot [(x_j(t) - x_i(t)) w_{ijk}(t) + (x_k(t) - x_i(t)) w_{ijk}(t)], \end{aligned} \quad (3.10)$$

where $w_{ijk}(t) = \theta_{ijk}(t) - p_{ijk}(t)$. Clearly, $E w_{ijk}(t) = 0$. Hence, this system fits into the current framework and we can apply Corollary 3 to conclude the random consensus result.

4. Proofs of main results and corollaries

In this section, we present the proofs for the main results of Theorems 1 and 2 in the above section together with some corollaries and remarks. The mathematical methods are based upon hypergraph theory, matrix analysis, real analysis, probability, martingale convergence and the Lyapunov stability techniques.

Proof of Theorem 1. (i) Using the assumption that the hypergraph $\mathcal{H}(t)$ is jointly uniformly connected, we have an infinite sequence of time steps $0 = t_0 < t_1 < t_2 < \dots < t_l < \dots$ satisfying $t_l - t_{l-1} \leq \ell$ and $\cup_{s=t_l}^{t_{l+1}-1} \mathcal{H}(s)$ connected for any $l \in \mathbb{N}$. Given $t \in \mathbb{N}$, let $\mathcal{G}(t) = (V, E_{\mathcal{G}}(t))$ be the graph on V with

an edge $\{i, j\} \in E_{\mathcal{G}}(t)$ present if $b_{ij}(t) > 0$. Therefore, the union $\bigcup_{s=t_l}^{t_{l+1}-1} \mathcal{G}(s)$ is connected as a graph (c.f. [34]) for all $l \in \mathbb{N}$.

Since $w(t) = 0_{n^2} \in \mathbb{R}^{n^2}$, by (3.3) we are dealing with the deterministic system

$$x(t+1) = (I_n - 2\gamma L(t))x(t) = B(t)x(t) \quad (4.1)$$

in Theorem 1. By the choice of γ , $\{B(t)\}_{t \in \mathbb{N}}$ is a sequence of stochastic matrices with all diagonal elements having a positive lower bound. This lower bound is an absolute constant irrelevant to γ . In view of the definition (3.6) and the comment above regarding connectivity, the same can be said for the matrix sequence $\{F(t_{l+1}, t_l)\}_{l \in \mathbb{N}}$.

By our choice of γ , using (3.6) and (4.1), we calculate

$$\begin{aligned} F(t_{l+1}, t_l) &= (I_n - 2\gamma L(t_{l+1} - 1)) \cdot (I_n - 2\gamma L(t_{l+1} - 2)) \cdots (I_n - 2\gamma L(t_l)) \\ &= I_n - 2\gamma \sum_{s=t_l}^{t_{l+1}-1} L(s) + O(\gamma^2) 1_n 1_n^\top, \end{aligned} \quad (4.2)$$

where $1_n \in \mathbb{R}^n$ is the vector with all elements being 1 and the Landau notation $a = O(b)$ indicates $|a/b| \leq c$ for some constant $c > 0$. For a sequence of $n \times n$ stochastic matrices $\{W(t) = (W_{ij}(t))\}_{t \in \mathbb{N}}$, it is called cut-balanced [40] if there is a constant $c > 0$ such that $\sum_{i \in S, j \notin S} W_{ij}(t) \geq c \sum_{i \notin S, j \in S} W_{ij}(t)$ for all $\emptyset \neq S \subset V$ and $t \in \mathbb{N}$. In view of the connectivity of the union graphs $\bigcup_{s=t_l}^{t_{l+1}-1} \mathcal{G}(s)$ ($l \in \mathbb{N}$) and the above calculation, it is clear that $\{F(t_{l+1}, t_l)\}_{l \in \mathbb{N}}$ is cut-balanced. The constant c involved here is irrelevant to γ .

Since $\{B(t)\}_{t \in \mathbb{N}}$ is a sequence of symmetric stochastic matrices, the same can be said for $\{F(t_{l+1}, t_l)\}_{l \in \mathbb{N}}$ by (3.6). With the above discussions, according to Lemmas 8 and 9 in [41], there exists a probability vector sequence $\alpha(t_l) = (\alpha_1(t_l), \alpha_2(t_l), \dots, \alpha_n(t_l))^\top \in \mathbb{R}^n$ for $l \in \mathbb{N}$ such that $\sum_{i \in V} \alpha_i(t_l) = 1$, $\alpha_i(t_l) \geq \eta$ and

$$\alpha(t_{l+1})^\top F(t_{l+1}, t_l) = \alpha(t_l)^\top \quad (4.3)$$

for $i \in V$ and $l \in \mathbb{N}$, where the constant $\eta \in (0, 1)$ is irrelevant to γ . To conclude the proof of (i), we need to extend the definition of the probability vector $\alpha(\cdot)$ to all-time steps. By (4.3), we have

$$\alpha(t_{l+1})^\top B(t_{l+1} - 1) B(t_{l+1} - 2) \cdots B(t_l) = \alpha(t_l)^\top. \quad (4.4)$$

Taking $l = 0$ in (4.4), we can define $\alpha(1)^\top = \alpha(t_1)^\top B(t_1 - 1) B(t_1 - 2) \cdots B(1)$. Clearly, $\alpha(1)$ is still a probability vector. As the diagonal elements of any matrix in $\{B(t)\}_{t=t_l}^{t_{l+1}-1}$ have a positive lower bound irrelevant to γ and all these matrices are symmetric and stochastic, the property $\alpha_i(1) \geq \eta$ still holds for all $i \in V$. Since $t_0 = 0$, we have $\alpha(1)^\top B(0) = \alpha(0)^\top$. Similarly, we define $\alpha(2)^\top = \alpha(t_1)^\top B(t_1 - 1) B(t_1 - 2) \cdots B(2)$. Likewise, $\alpha(2)$ is also a probability vector satisfying $\alpha_i(2) \geq \eta$ ($i \in V$) and $\alpha(2)^\top B(1) = \alpha(1)^\top$. Since $t_1 - t_0 \leq \ell$, in the same manner we can define each probability vector $\alpha(t)$ for $1 = t_0 + 1 \leq t \leq t_1 - 1$.

Taking $l = 1$ in (4.4), we can analogously define each probability vector $\alpha(t)$ for $t_1 + 1 \leq t \leq t_2 - 1$ one by one as above. Repeating this procedure by taking $l = 2, 3, 4, \dots$, the probability vector $\alpha(t)$ can be defined for all time step $t \in \mathbb{N}$ and the properties specified in (i) hold. This concludes the proof of (i).

(ii) To show the inequality for the quadratic Lyapunov function given in (3.5), we first note that

$$\begin{aligned}\Gamma(t_l) &\leq \sum_{i \in V} \alpha_i(t_l) \left[\max_{i \in V} x_i(t_l) - \min_{i \in V} x_i(t_l) \right]^2 \\ &= \left[\max_{i \in V} x_i(t_l) - \min_{i \in V} x_i(t_l) \right]^2, \quad l \in \mathbb{N},\end{aligned}\quad (4.5)$$

where we have used the fact that $\alpha(t_l)$ is a probability vector. On the other hand, by the results of (i) and (3.5), we derive

$$\begin{aligned}\Gamma(t_l) &\geq \min_{i \in V} \alpha_i(t_l) \cdot \sum_{i \in V} [x_i(t_l) - \alpha(t_l)^\top x(t_l)]^2 \\ &\geq \min_{i \in V} \alpha_i(t_l) \cdot \left\{ \left[\max_{i \in V} x_i(t_l) - \alpha(t_l)^\top x(t_l) \right]^2 + \left[\min_{i \in V} x_i(t_l) - \alpha(t_l)^\top x(t_l) \right]^2 \right\} \\ &\geq \frac{1}{2} \min_{i \in V} \alpha_i(t_l) \cdot \left[\max_{i \in V} x_i(t_l) - \min_{i \in V} x_i(t_l) \right]^2 \\ &\geq \frac{\eta}{2} \cdot \left[\max_{i \in V} x_i(t_l) - \min_{i \in V} x_i(t_l) \right]^2,\end{aligned}\quad (4.6)$$

where the third inequality is an application of the inequality of quadratic and arithmetic means. Combining (4.5) and (4.6) we know that

$$\Gamma(t_l) = \Theta \left(\left[\max_{i \in V} x_i(t_l) - \min_{i \in V} x_i(t_l) \right]^2 \right), \quad l \in \mathbb{N}, \quad (4.7)$$

where the Landau notation $a = \Theta(b)$ means $a = O(b)$ and $b = O(a)$. Here, the constant involved in $\Theta(\cdot)$ is irrelevant to γ .

For the stochastic matrix $B(t)$, suppose there is a probability vector $\bar{\alpha}(t) = (\bar{\alpha}_1(t), \dots, \bar{\alpha}_n(t))^\top \in \mathbb{R}^n$ satisfying $\bar{\alpha}(t+1)^\top B(t) = \bar{\alpha}(t)^\top$ for $t \in \mathbb{N}$. Define $\bar{\Gamma}(t) := \sum_{i \in V} \bar{\alpha}_i(t) [x_i(t) - \bar{\alpha}(t)^\top x(t)]^2 \in \mathbb{R}$ and a matrix $\bar{D}(t) = (\bar{d}_{ij}(t)) \in \mathbb{R}^{n \times n}$ as $\bar{D}(t) := B(t)^\top \text{diag}(\bar{\alpha}(t+1)) B(t)$ for $t \in \mathbb{N}$. By a direct calculation (c.f. [41, Lemma 3]), we know that

$$\bar{\Gamma}(t+1) - \bar{\Gamma}(t) = - \sum_{i < j, i, j \in V} \bar{d}_{ij}(t) (x_i(t) - x_j(t))^2 \leq 0 \quad (4.8)$$

for $t \in \mathbb{N}$. Note that $\bar{D}(t)$ is a non-negative matrix since $B(t)$ is a stochastic matrix. Based on (3.5), (4.3) and the previous analysis, the role of $B(t)$, $\bar{\alpha}(t)$, $\bar{\Gamma}(t)$ and $t \in \mathbb{N}$ can be replaced by $F(t_{l+1}, t_l)$, $\alpha(t_l)$, $\Gamma(t_l)$ and $l \in \mathbb{N}$, respectively, under the current setting. Therefore, we obtain

$$\Gamma(t_{l+1}) - \Gamma(t_l) = - \sum_{i < j, i, j \in V} d_{ij}(t_l) (x_i(t_l) - x_j(t_l))^2, \quad l \in \mathbb{N}, \quad (4.9)$$

with

$$D(t_l) := F(t_{l+1}, t_l)^\top \text{diag}(\alpha(t_{l+1})) F(t_{l+1}, t_l)$$

$$\begin{aligned}
&= \left[I_n - 2\gamma \sum_{s=t_l}^{t_{l+1}-1} L(s) + O(\gamma^2) \mathbf{1}_n \mathbf{1}_n^\top \right]^\top \cdot \text{diag}(\alpha(t_{l+1})) \\
&\quad \cdot \left[I_n - 2\gamma \sum_{s=t_l}^{t_{l+1}-1} L(s) + O(\gamma^2) \mathbf{1}_n \mathbf{1}_n^\top \right], \tag{4.10}
\end{aligned}$$

where (4.3) is used in the first equality and (4.2) is used in the second equality. The matrix $D(t_l)$ in (4.10) can be further written as $D(t_l) = D'(t_l) + O(\gamma^2) \mathbf{1}_n \mathbf{1}_n^\top$, where $D'(t_l) = (d'_{ij}(t_l)) \in \mathbb{R}^{n \times n}$ with

$$\begin{aligned}
D'(t_l) := & \text{diag}(\alpha(t_{l+1})) - 2\gamma \left[\left(\sum_{s=t_l}^{t_{l+1}-1} L(s) \right)^\top \cdot \text{diag}(\alpha(t_{l+1})) \right. \\
& \left. + \text{diag}(\alpha(t_{l+1})) \cdot \sum_{s=t_l}^{t_{l+1}-1} L(s) \right]. \tag{4.11}
\end{aligned}$$

Hence, $d_{ij}(t_l) = \Theta(\gamma)$ for all $i < j$ and $l \in \mathbb{N}$.

Given $l \in \mathbb{N}$, let $i_a := \max\{i \in V : x_i(t_l) = \max_{j \in V} x_j(t_l)\}$ and $i_b := \max\{i \in V : x_i(t_l) = \min_{j \in V} x_j(t_l)\}$. Since $\cup_{s=t_l}^{t_{l+1}-1} \mathcal{G}(s)$ is connected, there is a path $\{i_a, i_{a+1}\}, \{i_{a+1}, i_{a+2}\}, \dots, \{i_{b-1}, i_b\}$ with $\{i_r, i_{r+1}\} \in \cup_{s=t_l}^{t_{l+1}-1} \mathcal{E}_{\mathcal{G}}(s)$ for all $r = a, a+1, \dots, b-1$. Apparently, we can always choose two indices $a, b \in \mathbb{N}$ with $a < b$ satisfying the above arrangement. [In fact, if $a = b$, we obtain $\Gamma(t_{l+1}) = \Gamma(t_l)$ by (4.9). Hence, the result in (ii) holds trivially.] It follows from (4.11) that $d'_{i_r, i_{r+1}}(t_l) = d'_{i_{r+1}, i_r}(t_l) = \Theta(\gamma)$ for $a \leq r < b$. By the relationship between $D(t_l)$ and $D'(t_l)$, we further know that $0 \leq d_{i_r, i_{r+1}}(t_l) = d_{i_{r+1}, i_r}(t_l) = \Theta(\gamma)$ for $a \leq r < b$. Consequently,

$$\begin{aligned}
\sum_{i < j, i, j \in V} d_{ij}(t_l) (x_i(t_l) - x_j(t_l))^2 &\geq \sum_{r=a}^{b-1} d_{i_r, i_{r+1}}(t_l) [x_{i_r}(t_l) - x_{i_{r+1}}(t_l)]^2 \\
&\geq \frac{1}{b-a} [x_{i_a}(t_l) - x_{i_b}(t_l)]^2 \cdot \min_{a \leq r < b} d_{i_r, i_{r+1}}(t_l) \\
&= \Theta(\gamma) \cdot [x_{i_a}(t_l) - x_{i_b}(t_l)]^2, \tag{4.12}
\end{aligned}$$

where the second inequality above is due to an application of the inequality of quadratic and arithmetic means.

Combining (4.7) and (4.12), we have for $l \in \mathbb{N}$,

$$\sum_{i < j, i, j \in V} d_{ij}(t_l) (x_i(t_l) - x_j(t_l))^2 \geq \Theta(\gamma) \cdot \Gamma(t_l) \geq \eta' \cdot \gamma \cdot \Gamma(t_l) \tag{4.13}$$

for some constant η' such that $0 < \eta' < \gamma^{-1}$ and η' is irrelevant to γ . Using (4.9) and (4.13), we finally derive

$$\Gamma(t_{l+1}) - \Gamma(t_l) \leq -\eta' \gamma \Gamma(t_l), \tag{4.14}$$

which readily completes the Proof of Theorem 1. \square

An immediate corollary of Theorem 1 is that all node states $\{x_i(t)\}_{i \in V}$ of the system (3.3) achieve consensus at an exponential rate (in a deterministic way) when the noises are absent.

COROLLARY 1 Under the conditions of Theorem 1, we have for $t \in \mathbb{N}$,

$$\max_{i,j \in V} |x_i(t) - x_j(t)| \leq \frac{\sqrt{2\Gamma(0)}}{\eta} \cdot (1 - \eta'\gamma)^{\frac{t}{2}}, \quad (4.15)$$

where all notations are defined as in Theorem 1. Hence, $|x_i(t) - x_j(t)| \rightarrow 0$ exponentially fast for all $i, j \in V$.

Proof. By (4.14), we have $\Gamma(t_l) \leq (1 - \eta'\gamma)^l \Gamma(0)$ for $l \in \mathbb{N}$ since $t_0 = 0$. For any $t \in \mathbb{N}$, suppose that $t \in [t_l, t_{l+1})$ for some $l \in \mathbb{N}$ without loss of generality. We obtain

$$\Gamma(t) \leq \Gamma(t_l) \leq (1 - \eta'\gamma)^l \Gamma(0) \leq (1 - \eta'\gamma)^{\frac{t}{2}} \Gamma(0), \quad (4.16)$$

where the first inequality comes from the monotonicity in (4.8) and the last one is based on the joint uniform connectivity of $\mathcal{H}(t)$.

Since $(x_i(t) - x_j(t))^2 \leq [x_i(t) - \alpha(t)^\top x(t)]^2 + [x_j(t) - \alpha(t)^\top x(t)]^2$, multiplying $\alpha_i(t)\alpha_j(t)$ on both sides and taking sums over all $i, j \in V$ give rise to

$$\begin{aligned} \eta^2 \max_{i,j \in V} (x_i(t) - x_j(t))^2 &\leq \eta^2 \sum_{i,j \in V} (x_i(t) - x_j(t))^2 \\ &\leq \sum_{i,j \in V} \alpha_i(t)\alpha_j(t) (x_i(t) - x_j(t))^2 \\ &\leq 2 \sum_{i \in V} \alpha_i(t) [x_i(t) - \alpha(t)^\top x(t)]^2 = 2\Gamma(t), \end{aligned} \quad (4.17)$$

where we used Theorem 1(i) and the definition (3.5). The result (4.15) follows from (4.16) and (4.17). \square

If we consider the generalized system (2.3) with the additional modulating function $h(\cdot)$, we can obtain analogous results. Let $b'_{ij}(t) := \sum_{k \in V} a_{ijk}(t)h(|x_j(t) - x_k(t)|)$ for $i, j \in V$ and $L'(t) = (l'_{ij}(t)) \in \mathbb{R}^{n \times n}$ with $l'_{ii}(t) = \sum_{j \in V} b'_{ij}(t)$ for $i \in V$ and $l'_{ij}(t) = -b'_{ij}(t)$ for $i \neq j$. Let $B'(t) := I_n - 2\gamma L'(t)$. The system (2.3) can then be rewritten as

$$x(t+1) = B'(t)x(t). \quad (4.18)$$

COROLLARY 2 Consider the system (2.3) with $0 < h(\cdot) \leq \beta_3$ for some $\beta_3 > 0$. Assume that $\gamma < (2\beta_3 \sup_{t \in \mathbb{N}} \max_{i \in V} \sum_{j,k \in V} a_{ijk}(t))^{-1}$ and $\mathcal{H}(t)$ is jointly uniformly connected with bound ℓ and time sequence $\{t_l\}_{l \in \mathbb{N}}$. Then, the following hold:

- (i) There exists a probability vector sequence $\alpha(t) = (\alpha_1(t), \alpha_2(t), \dots, \alpha_n(t))^\top \in \mathbb{R}^n$ for $t \in \mathbb{N}$ such that $\sum_{i \in V} \alpha_i(t) = 1, \alpha_i(t) \geq \eta$ and

$$\alpha(t+1)^\top B'(t) = \alpha(t)^\top \quad (4.19)$$

for $i \in V$ and $t \in \mathbb{N}$, where $0 < \eta < 1$.

- (ii) $\Gamma(t_{l+1}) \leq (1 - \eta'\gamma)\Gamma(t_l)$ holds for $l \in \mathbb{N}$, where $0 < \eta' < \gamma^{-1}$ and

$$\Gamma(t) := \sum_{i \in V} \alpha_i(t) [x_i(t) - \alpha(t)^\top x(t)]^2. \quad (4.20)$$

Both η and η' are absolute constants that do not rely on the gain γ .

Proof. (i) Note that $\{B'(t)\}_{t \in \mathbb{N}}$ is a sequence of stochastic matrices although they are no longer symmetric. In analogy to (3.6), for any $s \in \mathbb{N}$, we define $F'(s, s) = I_n \in \mathbb{R}^{n \times n}$ and

$$F'(s+l+1, s) = B'(s+l)F'(s+l, s) \quad (4.21)$$

for $l \in \mathbb{N}$. Therefore, $\{F'(t, s)\}_{t, s \in \mathbb{N}}$ also forms a family of stochastic matrices. The same line of proof in Theorem 1(i) can be applied here as we did not actually rely on symmetry of these matrices.

(ii) Here, it follows again from the same line of Proof in Theorem 1(ii). We replace $D(t)$ with $D''(t) := B'(t)^\top \text{diag}(\alpha(t+1))B'(t)$ for $t \in \mathbb{N}$. The same proof can be applied. \square

Clearly, Corollary 1 also holds verbatim for the system (2.3). As a byproduct, our results present a solid theoretical support for the consensus phenomena observed in the recent works [18, 19, 31]. We will see later in Corollary 3 that even with random perturbation, the consensus results still hold. The proof of our main contribution to Theorem 2 is given below.

Proof of Theorem 2. (i) To show the almost sure and \mathcal{L}^2 convergence, we will again examine the asymptotic behaviour of the quadratic Lyapunov function $\Gamma(t)$. However, different from Theorem 1, we need to cope with the random noises.

Recall that the hypergraph $\mathcal{H}(t)$ is jointly uniformly connected. We obtain an infinite sequence of time steps $0 = t_0 < t_1 < t_2 < \dots < t_l < \dots$ satisfying $t_l - t_{l-1} \leq \ell$ and $\cup_{s=t_l}^{t_{l+1}-1} \mathcal{H}(s)$ connected for any $l \in \mathbb{N}$. It follows from (3.3) that $x(t_l+1) = (I_n - 2\gamma L(t_l))x(t_l) + 2\gamma C(t_l)G(t_l)w(t_l)$. Another application of (3.3) to the above expression gives $x(t_l+2) = (I_n - 2\gamma L(t_l+1))(I_n - 2\gamma L(t_l))x(t_l) + 2\gamma(I_n - 2\gamma L(t_l+1))C(t_l)G(t_l)w(t_l) + 2\gamma C(t_l+1)G(t_l+1)w(t_l+1)$. Since $t_{l+1} - t_l \leq \ell$, a repetition of the above procedure leads to

$$\begin{aligned} x(t_{l+1}) = & (I_n - 2\gamma L(t_{l+1} - 1)) \cdots (I_n - 2\gamma L(t_l))x(t_l) \\ & + 2\gamma(I_n - 2\gamma L(t_{l+1} - 1)) \cdots (I_n - 2\gamma L(t_l + 1))C(t_l)G(t_l)w(t_l) + \cdots \end{aligned}$$

$$\begin{aligned}
& + 2\gamma(I_n - 2\gamma L(t_{l+1} - 1))(I_n - 2\gamma L(t_{l+1} - 2)) \\
& \quad \cdot C(t_{l+1} - 3)G(t_{l+1} - 3)w(t_{l+1} - 3) \\
& + 2\gamma(I_n - 2\gamma L(t_{l+1} - 1))C(t_{l+1} - 2)G(t_{l+1} - 2)w(t_{l+1} - 2) \\
& + 2\gamma C(t_{l+1} - 1)G(t_{l+1} - 1)w(t_{l+1} - 1) \\
& = F(t_{l+1}, t_l)x(t_l) + 2\gamma F(t_{l+1}, t_l + 1)C(t_l)G(t_l)w(t_l) + \dots \\
& \quad + 2\gamma F(t_{l+1}, t_{l+1} - 1)C(t_{l+1} - 2)G(t_{l+1} - 2)w(t_{l+1} - 2) \\
& \quad + 2\gamma F(t_{l+1}, t_{l+1})C(t_{l+1} - 1)G(t_{l+1} - 1)w(t_{l+1} - 1) \\
& = F(t_{l+1}, t_l)x(t_l) + y(t_l), \quad l \in \mathbb{N},
\end{aligned} \tag{4.22}$$

where we applied the recursive expression (3.6) and $y(t_l) := 2\gamma \sum_{s=t_l}^{t_{l+1}-1} F(t_{l+1}, s+1)C(s)G(s)w(s) \in \mathbb{R}^n$. By Theorem 1(i), there is a probability vector sequence $\alpha(t) = (\alpha_1(t), \alpha_2(t), \dots, \alpha_n(t))^\top \in \mathbb{R}^n$ satisfying $\sum_{i \in V} \alpha_i(t) = 1$, $\alpha_i(t) \geq \eta$ and (3.4) for $i \in V$ and $t \in \mathbb{N}$, where $0 < \eta < 1$ is a constant. Define $\Gamma(t)$ as in (3.5), and we obtain the increment of Γ along the time sequence $\{t_l\}_{l \in \mathbb{N}}$ as

$$\begin{aligned}
\Gamma(t_{l+1}) - \Gamma(t_l) &= \sum_{i \in V} \alpha_i(t_{l+1}) [x_i(t_{l+1}) - \alpha^\top(t_{l+1})x(t_{l+1})]^2 \\
&\quad - \sum_{i \in V} \alpha_i(t_l) [x_i(t_l) - \alpha^\top(t_l)x(t_l)]^2 \\
&= \sum_{i \in V} \alpha_i(t_{l+1}) \left[x_i(t_{l+1})^2 - 2x_i(t_{l+1})\alpha^\top(t_{l+1})x(t_{l+1}) \right. \\
&\quad \quad \left. + (\alpha^\top(t_{l+1})x(t_{l+1}))^2 \right] \\
&\quad - \sum_{i \in V} \alpha_i(t_l) \left[x_i(t_l)^2 - 2x_i(t_l)\alpha^\top(t_l)x(t_l) + (\alpha^\top(t_l)x(t_l))^2 \right] \\
&= x(t_{l+1})^\top \text{diag}(\alpha(t_{l+1}))x(t_{l+1}) - (\alpha(t_{l+1})^\top x(t_{l+1}))^2 \\
&\quad - x(t_l)^\top \text{diag}(\alpha(t_l))x(t_l) + (\alpha(t_l)^\top x(t_l))^2.
\end{aligned} \tag{4.23}$$

In view of (4.22) and (4.3), the expression (4.23) can be further written as

$$\begin{aligned}
\Gamma(t_{l+1}) - \Gamma(t_l) &= [x(t_l)^\top F(t_{l+1}, t_l)^\top + y(t_l)^\top] \text{diag}(\alpha(t_{l+1})) \\
&\quad \cdot [F(t_{l+1}, t_l)x(t_l) + y(t_l)] \\
&\quad - [\alpha(t_{l+1})^\top F(t_{l+1}, t_l)x(t_l) + \alpha(t_{l+1})^\top y(t_l)]^2 \\
&\quad - x(t_l)^\top \text{diag}(\alpha(t_l))x(t_l) + (\alpha(t_l)^\top x(t_l))^2 \\
&= x(t_l)^\top F(t_{l+1}, t_l)^\top \text{diag}(\alpha(t_{l+1}))F(t_{l+1}, t_l)x(t_l) \\
&\quad + x(t_l)^\top F(t_{l+1}, t_l)^\top \text{diag}(\alpha(t_{l+1}))y(t_l) \\
&\quad + y(t_l)^\top \text{diag}(\alpha(t_{l+1}))F(t_{l+1}, t_l)x(t_l)
\end{aligned}$$

$$\begin{aligned}
& + y(t_l)^\top \text{diag}(\alpha(t_{l+1}))y(t_l) \\
& - 2\alpha(t_l)^\top x(t_l)\alpha(t_{l+1})^\top y(t_l) - (\alpha(t_{l+1})^\top y(t_l))^2 \\
& - x(t_l)^\top \text{diag}(\alpha(t_l))x(t_l) \\
& = x(t_l)^\top \Xi_1(t_l)x(t_l) - 2x(t_l)^\top \Xi_2(t_l)y(t_l) \\
& + y(t_l)^\top \Xi_3(t_l)y(t_l), \tag{4.24}
\end{aligned}$$

where $\Xi_1(t_l) := F(t_{l+1}, t_l)^\top \text{diag}(\alpha(t_{l+1}))F(t_{l+1}, t_l) - \text{diag}(\alpha(t_l)) \in \mathbb{R}^{n \times n}$, $\Xi_2(t_l) := F(t_{l+1}, t_l)^\top \text{diag}(\alpha(t_{l+1})) - \alpha(t_l)\alpha(t_{l+1})^\top \in \mathbb{R}^{n \times n}$, and $\Xi_3(t_l) := \text{diag}(\alpha(t_{l+1})) - \alpha(t_{l+1})\alpha(t_{l+1})^\top \in \mathbb{R}^{n \times n}$. In the following, we will estimate the three (random) terms on the righthand side of (4.24).

Since $F(t_{l+1}, t_l)$ is a stochastic matrix, by using (4.3) we obtain

$$\begin{aligned}
\Xi_1(t_l)1_n &= F(t_{l+1}, t_l)^\top \text{diag}(\alpha(t_{l+1}))1_n - \text{diag}(\alpha(t_l))1_n \\
&= F(t_{l+1}, t_l)^\top \alpha(t_{l+1}) - \alpha(t_l) = 0_n, \tag{4.25}
\end{aligned}$$

where $0_n \in \mathbb{R}^n$ is a zero vector. Recalling the definition of $D(t_l) \in \mathbb{R}^{n \times n}$ in (4.10), we derive $\Xi_1(t_l) = D(t_l) - \text{diag}(\alpha(t_l))$ with the sum of each row being 0 as implied by (4.25). Hence,

$$\begin{aligned}
& x(t_l)^\top \Xi_1(t_l)x(t_l) \\
& = x(t_l)^\top D(t_l)x(t_l) - x(t_l)^\top \text{diag}(\alpha(t_l))x(t_l) \\
& = x(t_l)^\top \begin{pmatrix} \alpha_1(t_l) - \sum_{i \in V \setminus \{1\}} d_{1i}(t_l) & d_{12}(t_l) & \cdots & d_{1n}(t_l) \\ d_{21}(t_l) & \alpha_2(t_l) - \sum_{i \in V \setminus \{2\}} d_{2i}(t_l) & \cdots & d_{2n}(t_l) \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1}(t_l) & d_{n2}(t_l) & \cdots & \alpha_n(t_l) - \sum_{i \in V \setminus \{n\}} d_{ni}(t_l) \end{pmatrix} \\
& \quad \cdot x(t_l) - x(t_l)^\top \text{diag}(\alpha(t_l))x(t_l) \\
& \leq - \sum_{i < j, i, j \in V} d_{ij}(t_l) (x_i(t_l) - x_j(t_l))^2 \tag{4.26}
\end{aligned}$$

since $D(t_l)$ is a nonnegative matrix. Combining (4.14), (4.24) and (4.26) we have

$$\Gamma(t_{l+1}) \leq (1 - \eta'\gamma)\Gamma(t_l) - 2x(t_l)^\top \Xi_2(t_l)y(t_l) + y(t_l)^\top \Xi_3(t_l)y(t_l). \tag{4.27}$$

Let \mathcal{F}_t be the σ -algebra generated by the list $\{x(0), w(0), w(1), \dots, w(t)\}$. Then, $\{\mathcal{F}_t\}_{t \in \mathbb{N}}$ is a filtration and $x(t_l)$ is \mathcal{F}_{t_l-1} -measurable. Assumption 2 leads to $\mathbb{E}(x(t_l)^\top \Xi_2(t_l)y(t_l)) = 0$ and $\mathbb{E}(w(s_1)^\top w(s_2)) = 0$ for $s_1, s_2 \in \mathbb{N}$ with $s_1 \neq s_2$. Therefore,

$$\begin{aligned}
\mathbb{E}(y(t_l)^\top \Xi_3(t_l)y(t_l)) &= \mathbb{E} \left\{ \left[2\gamma \sum_{s=t_l}^{t_{l+1}-1} F(t_{l+1}, s+1)C(s)G(s)w(s) \right]^\top \Xi_3(t_l) \right. \\
& \quad \left. \cdot \left[2\gamma \sum_{s=t_l}^{t_{l+1}-1} F(t_{l+1}, s+1)C(s)G(s)w(s) \right] \right\}
\end{aligned}$$

$$\begin{aligned}
&= 4\gamma^2 \mathbb{E} \left\{ \sum_{s=t_l}^{t_{l+1}-1} w(s)^\top G(s)^\top C(s)^\top F(t_{l+1}, s+1)^\top \Xi_3(t_l) \right. \\
&\quad \left. \cdot F(t_{l+1}, s+1) C(s) G(s) w(s) \right\}. \tag{4.28}
\end{aligned}$$

With the above comments, taking expectation over (4.27) yields

$$\begin{aligned}
\mathbb{E}(\Gamma(t_{l+1})) &\leq (1 - \eta'\gamma) \mathbb{E} \Gamma(t_l) \\
&\quad + 4\gamma^2 \sum_{s=t_l}^{t_{l+1}-1} \mathbb{E} \left(w(s)^\top G(s)^\top \Xi(t_l, s) G(s) w(s) \right), \tag{4.29}
\end{aligned}$$

where $\Xi(t_l, s) := C(s)^\top F(t_{l+1}, s+1)^\top \Xi_3(t_l) F(t_{l+1}, s+1) C(s) \in \mathbb{R}^{n^2 \times n^2}$. Note that both $\Xi_3(t_l)$ and $\Xi(t_l, s)$ are symmetric matrices for $l \in \mathbb{N}$. In the following, we aim to bind the second term on the right-hand side of (4.29) in terms of the expected Lyapunov $\mathbb{E} \Gamma(t_l)$, similar to the first term on the right-hand side of (4.29).

To this end, let ξ_1 be the maximal spectral norm of the matrices $\Xi(t_l, s)$ across $l \in \mathbb{N}$ and $s \in \{t_l, t_l + 1, \dots, t_{l+1} - 1\}$. Namely,

$$\xi_1 = \sup_{l \in \mathbb{N}} \max_{t_l \leq s \leq t_{l+1}-1} \sigma(\Xi(t_l, s)) > 0, \tag{4.30}$$

where $\sigma(\cdot)$ means the largest singular value of a matrix. Using the condition $t_{l+1} - t_l \leq \ell$, Assumption 1, as well as the construction and symmetry of $\Xi(t_l, s)$, we know that $\xi_1 < \infty$ by applying the Gershgorin circle theorem. Clearly, the constant ξ_1 is irrelevant to γ . It then follows from Assumptions 1, 2 and the Rayleigh–Ritz theorem that

$$\begin{aligned}
\mathbb{E} \left(w(s)^\top G(s)^\top \Xi(t_l, s) G(s) w(s) \right) &\leq \xi_1 \sum_{ij \in V} \mathbb{E} [g(x_j(s) - x_i(s))^2] \mathbb{E} [w_{ij}(s)^2] \\
&\leq \xi_1 \beta_2^2 \sum_{ij \in V} \mathbb{E} [(x_j(s) - x_i(s))^2]. \tag{4.31}
\end{aligned}$$

Define the matrices $P_1 = [1, 0_{n-1}^\top] \in \mathbb{R}^{1 \times n}$, $P_2 = [1_{n-1}, -I_{n-1}] \in \mathbb{R}^{(n-1) \times n}$ and $P = \begin{bmatrix} P_1 \\ P_2 \end{bmatrix}$. For $t \in \mathbb{N}$, we define the vector $z(t) \in \mathbb{R}^{n-1}$ through the following transformation $[x_1(t), z(t)^\top]^\top = Px(t) = [x_1(t), x_1(t) - x_2(t), x_1(t) - x_3(t), \dots, x_1(t) - x_n(t)]^\top \in \mathbb{R}^n$. Multiplying P on both side of (3.3) gives $Px(t+1) = PB(t)x(t) + 2\gamma PC(t)G(t)w(t)$, which further implies that

$$\begin{aligned}
x_1(t+1) &= x_1(t) - 2\gamma(l_{11}(t)x_1(t) + l_{12}(t)x_2(t) + \dots + l_{1n}(t)x_n(t)) \\
&\quad + 2\gamma P_1 C(t) G(t) w(t) \\
&= x_1(t) - 2\gamma L^{P_1}(t) z(t) + 2\gamma P_1 C(t) G(t) w(t), \tag{4.32}
\end{aligned}$$

where $L^{P_1}(t) := [-l_{12}(t), -l_{13}(t), \dots, -l_{1n}(t)] \in \mathbb{R}^{1 \times (n-1)}$, and that

$$\begin{aligned}
& z(t+1) \\
&= z(t) - 2\gamma \begin{pmatrix} l_{11}(t) - l_{21}(t) & l_{12}(t) - l_{22}(t) & \cdots & l_{1n}(t) - l_{2n}(t) \\ l_{11}(t) - l_{31}(t) & l_{12}(t) - l_{32}(t) & \cdots & l_{1n}(t) - l_{3n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ l_{11}(t) - l_{n1}(t) & l_{12}(t) - l_{n2}(t) & \cdots & l_{1n}(t) - l_{nn}(t) \end{pmatrix}_{(n-1) \times n} \\
&\quad \cdot \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix} + 2\gamma P_2 C(t) G(t) w(t) \\
&= z(t) - 2\gamma \begin{pmatrix} l_{22}(t) - l_{12}(t) & l_{23}(t) - l_{13}(t) & \cdots & l_{2n}(t) - l_{1n}(t) \\ l_{32}(t) - l_{12}(t) & l_{33}(t) - l_{13}(t) & \cdots & l_{3n}(t) - l_{1n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ l_{n2}(t) - l_{12}(t) & l_{n3}(t) - l_{13}(t) & \cdots & l_{nn}(t) - l_{1n}(t) \end{pmatrix}_{(n-1) \times (n-1)} \\
&\quad \cdot \begin{pmatrix} x_1(t) - x_2(t) \\ x_1(t) - x_3(t) \\ \vdots \\ x_1(t) - x_n(t) \end{pmatrix} + 2\gamma P_2 C(t) G(t) w(t) \\
&:= z(t) - 2\gamma L^{P_2}(t) z(t) + 2\gamma P_2 C(t) G(t) w(t), \tag{4.33}
\end{aligned}$$

where $L^{P_2}(t) \in \mathbb{R}^{(n-1) \times (n-1)}$ is defined as above.

Let $B^{P_2}(t) := I_{n-1} - 2\gamma L^{P_2}(t) \in \mathbb{R}^{(n-1) \times (n-1)}$. Define two constants

$$\xi_2 = \sup_{t \in \mathbb{N}} \sigma(B^{P_2}(t)^\top B^{P_2}(t)) > 0 \text{ and } \xi_3 = \sup_{t \in \mathbb{N}} \sigma(C(t)^\top P_2^\top P_2 C(t)) > 0. \tag{4.34}$$

Similarly as in (4.30), we have $\xi_2 < \infty$ and $\xi_3 < \infty$ by using the construction and symmetry of the two involved matrices and the Gershgorin circle theorem. It is easy to see that $\xi_2 = O(\gamma)$ and ξ_3 is irrelevant to γ . In view of Assumptions 1 and 2, we obtain for $t \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{E}(z(t+1)^\top z(t+1)) &= \mathbb{E}(z(t)^\top B^{P_2}(t)^\top B^{P_2}(t) z(t)) \\
&\quad + 4\gamma^2 \mathbb{E}(w(t)^\top G(t)^\top C(t)^\top P_2^\top P_2 C(t) G(t) w(t)) \\
&\leq \xi_2 \mathbb{E}(z(t)^\top z(t)) \\
&\quad + 4\gamma^2 \xi_3 \sum_{ij \in V} \mathbb{E}[g(x_j(t) - x_i(t))^2] \mathbb{E}[w_{ij}(t)^2] \\
&\leq \xi_2 \mathbb{E}(z(t)^\top z(t)) \\
&\quad + 4\gamma^2 \xi_3 \beta_2^2 \sum_{ij \in V} \mathbb{E}[(x_j(t) - x_i(t))^2], \tag{4.35}
\end{aligned}$$

where we have applied (4.33), (4.34) and the Rayleigh–Ritz theorem. Noting that

$$\begin{aligned} \sum_{i,j \in V} (x_j(t) - x_i(t))^2 &= \sum_{i,j \in V} (x_j(t) - x_1(t) + x_1(t) - x_i(t))^2 \\ &\leq 2 \left(\sum_{i,j \in V} (x_j(t) - x_1(t))^2 + \sum_{i,j \in V} (x_1(t) - x_i(t))^2 \right) \\ &= 4 \sum_{i,j \in V} z(t)^\top z(t), \end{aligned} \quad (4.36)$$

we further write (4.35) as

$$\begin{aligned} \mathbb{E} (z(t+1)^\top z(t+1)) &\leq \xi_2 \mathbb{E} (z(t)^\top z(t)) + 16\gamma^2 \xi_3 \beta_2^2 \sum_{i,j \in V} \mathbb{E} (z(t)^\top z(t)) \\ &\leq (\xi_2 + 2^4 \gamma^2 \xi_3 \beta_2^2 n^2) \mathbb{E} (z(t)^\top z(t)). \end{aligned} \quad (4.37)$$

By a similar argument in (4.6), we derive for any $i, j \in V$ and $t \in \mathbb{N}$,

$$\begin{aligned} \mathbb{E} (z(t)^\top z(t)) &\leq (n-1) \mathbb{E} \left[\left(\max_{k \in V} x_k(t) - \min_{k \in V} x_k(t) \right)^2 \right] \\ &\leq \frac{2(n-1)E\Gamma(t)}{\eta}. \end{aligned} \quad (4.38)$$

With (4.31), (4.36), (4.37) and (4.38) at hand, the second term on the right-hand side of (4.29) can finally be estimated as

$$\begin{aligned} &\sum_{s=t_l}^{t_{l+1}-1} \mathbb{E} \left(w(s)^\top G(s)^\top \Xi(t_l, s) G(s) w(s) \right) \\ &\leq \sum_{s=t_l}^{t_{l+1}-1} \xi_1 \beta_2^2 \sum_{i,j \in V} \mathbb{E} [(x_j(s) - x_i(s))^2] \\ &\leq 4\xi_1 \beta_2^2 \sum_{s=t_l}^{t_{l+1}-1} \sum_{i,j \in V} \mathbb{E} (z(s)^\top z(s)) \\ &\leq 4n^2 \xi_1 \beta_2^2 \sum_{s=t_l}^{t_{l+1}-1} \mathbb{E} (z(s)^\top z(s)) \\ &\leq 4n^2 \xi_1 \beta_2^2 \sum_{s=0}^{\ell-1} (\xi_2 + 2^4 \gamma^2 \xi_3 \beta_2^2 n^2)^s \mathbb{E} (z(t_l)^\top z(t_l)) \end{aligned}$$

$$\begin{aligned} &\leq 2^3(n-1)n^2\xi_1\beta_2^2\eta^{-1}\sum_{s=0}^{\ell-1}(\xi_2+2^4\gamma^2\xi_3\beta_2^2n^2)^s\mathbb{E}\Gamma(t_l) \\ &:=\xi\mathbb{E}\Gamma(t_l), \end{aligned} \quad (4.39)$$

where we used the condition $t_{l+1}-t_l\leq\ell$ for $l\in\mathbb{N}$ and the constant

$$\xi=\frac{2^3(n-1)n^2\xi_1\beta_2^2}{\eta}\cdot\sum_{s=0}^{\ell-1}(\xi_2+2^4\gamma^2\xi_3\beta_2^2n^2)^s>0. \quad (4.40)$$

Clearly, we have $\xi=\Theta(1)+O(\gamma)<\infty$, where the constant $\Theta(1)$ is irrelevant to γ . Feeding (4.39) into (4.29) yields the following estimate

$$\mathbb{E}\Gamma(t_{l+1})\leq(1-\eta'\gamma+4\xi\gamma^2)\mathbb{E}\Gamma(t_l) \quad (4.41)$$

for $l\in\mathbb{N}$. By Theorem 1(ii), $1-\eta'\gamma>0$ and hence the coefficient in (4.41) is $1-\eta'\gamma+4\xi\gamma^2>0$. Moreover, we have

$$0<1-\eta'\gamma+4\xi\gamma^2<1 \quad (4.42)$$

since $4\xi\gamma<\eta'$ by assumption. It is easy to see that the inequality (4.42) or equivalently $4\xi\gamma<\eta'$ is feasible when $\gamma>0$ is sufficiently small. Since $t_0=0$, it follows from (4.41) that

$$\mathbb{E}\Gamma(t_l)\leq(1-\eta'\gamma+4\xi\gamma^2)^l\Gamma(0). \quad (4.43)$$

Namely, the expectation $\mathbb{E}\Gamma(t_l)$ converges exponentially fast as $l\rightarrow\infty$.

To show the consensus of system (3.3) in the sense of the definitions given at the end of Section 2, there are still several steps to go. For example, we need to kind of get rid of the subscript l and the expectation operator \mathbb{E} in (4.43).

By repeatedly applying the update rule of (3.3) in a similar way as (4.22), we obtain

$$x(t)=F(t,0)x(0)+y'(t), \quad t\in\mathbb{N}, \quad (4.44)$$

where $y'(t):=2\gamma\sum_{s=0}^{t-1}F(t,s+1)C(s)G(s)w(s)\in\mathbb{R}^n$. For any $t\in\mathbb{N}$, let $\bar{l}:=\max\{l\in\mathbb{N}:t_l<t\}$. Thanks to Assumption 2, we have

$$\begin{aligned} &\mathbb{E}(y'(t)^\top y'(t)) \\ &=4\gamma^2\sum_{s=0}^{t-1}\mathbb{E}(w(s)^\top G(s)^\top C(s)^\top F(t,s+1)^\top F(t,s+1)C(s)G(s)w(s)) \\ &=4\gamma^2\left(\sum_{s=0}^{t_1-1}+\sum_{s=t_1}^{t_2-1}+\cdots+\sum_{s=t_{\bar{l}-1}}^{t_{\bar{l}}-1}+\sum_{s=t_{\bar{l}}}^{t-1}\right)\mathbb{E}(w(s)^\top G(s)^\top C(s)^\top F(t,s+1)^\top \end{aligned}$$

$$\begin{aligned} & \cdot F(t, s + 1)C(s)G(s)w(s) \\ & \leq \xi' E \Gamma(0) + \xi' E \Gamma(t_1) + \cdots + \xi' E \Gamma(t_{l-1}) + \xi' E \Gamma(t_l), \end{aligned} \quad (4.45)$$

where we have applied a similar argument as (4.39) for each term in the last step, where $\xi' > 0$ is a constant. In view of (4.42), (4.43) and $E \Gamma(0) = \Gamma(0)$, the upper bound in (4.45) is a sum of geometric sequence and hence we obtain

$$E(y'(t)^\top y'(t)) \leq \frac{\xi' \Gamma(0)}{\eta' \gamma - 4\xi \gamma^2}. \quad (4.46)$$

It can be seen from (3.5) and (4.8) that $\Gamma(t)$ is a (random) non-increasing function with respect to time and it is non-negative. Therefore, there exists a random variable $v \in \mathbb{R}$ such that $\lim_{t \rightarrow \infty} \Gamma(t) = v$ almost surely and $v \geq 0$. Clearly, $E v = 0$. [In fact, if $E v > 0$, we take $t_l \leq t$ and $E \Gamma(t) \leq E \Gamma(t_l) \leq (1 - \eta' \gamma + 4\xi \gamma^2)^l \Gamma(0) \rightarrow 0$ as $t \rightarrow \infty$ and $l \rightarrow \infty$ by (4.43), which contracts with $E \Gamma(t) \rightarrow E v > 0$.] Hence, $v = 0$ and accordingly $\lim_{t \rightarrow \infty} \Gamma(t) = 0$ almost surely. It follows from (3.5) that

$$\lim_{t \rightarrow \infty} x_i(t) - x_j(t) = 0 \quad (4.47)$$

almost surely for all $i, j \in V$. Taking expectation on both sides of (3.5) leads to $E \Gamma(t) = \sum_{i \in V} \alpha_i(t) E[(x_i(t) - \alpha(t)^\top x(t))^2] \rightarrow 0$ as $t \rightarrow \infty$. Therefore, for any $i \in V$, $E[(x_i(t) - \alpha(t)^\top x(t))^2] \rightarrow 0$ as $t \rightarrow \infty$. Thus, we derive the \mathcal{L}^2 convergence

$$\begin{aligned} & \lim_{t \rightarrow \infty} E[|x_i(t) - x_j(t)|^2] \\ & \leq \lim_{t \rightarrow \infty} E[|x_i(t) - \alpha(t)^\top x(t)|^2] + \lim_{t \rightarrow \infty} E[|x_j(t) - \alpha(t)^\top x(t)|^2] = 0 \end{aligned} \quad (4.48)$$

for all $i, j \in V$.

For the first part (i) of Theorem 2, what remains to show is the existence of a limit random variable. The trick up our sleeves is the following martingale convergence theorem; c.f. [37, Theorem 4.4.6]: Suppose that $\{X(t)\}_{t \in \mathbb{N}}$ is a martingale with $\sup_{t \in \mathbb{N}} E(X(t)^2) < \infty$. Then, there is a random variable X such that $X(t) \rightarrow X$ almost surely and in \mathcal{L}^2 .

Consider the system (3.3) without noises, we have $x(t) = F(t, s)x(s)$ for any $t \geq s$, $t, s \in \mathbb{N}$, where $\{F(t, s)\}_{t, s \in \mathbb{N}}$ is a collection of stochastic matrices. By Corollary 1, $\lim_{t \rightarrow \infty} x_i(t) - x_j(t) = 0$ for all $i, j \in V$. Therefore, for any $s \in \mathbb{N}$ the limit matrix $\lim_{t \rightarrow \infty} F(t, s)$ exists, and it takes the form of $1_n \cdot (\hat{\alpha}_1(s), \hat{\alpha}_2(s), \dots, \hat{\alpha}_n(s)) \in \mathbb{R}^{n \times n}$, where $\hat{\alpha}(s) := (\hat{\alpha}_1(s), \hat{\alpha}_2(s), \dots, \hat{\alpha}_n(s))$ is a row probability vector. Furthermore, $\hat{\alpha}(s) \lim_{t \rightarrow \infty} F(t, s) = \hat{\alpha}(s)$, namely, the row vector $\hat{\alpha}(s)$ is also the left eigenvector of the limit matrix with respect to eigenvalue 1. By the Perron–Frobenius theorem, $\hat{\alpha}(s)$ is a positive vector.

It follows from (3.3) and (4.44) that $x(t)$ and $y'(t)$ are \mathcal{F}_{t-1} -measurable. Accordingly,

$$\begin{aligned} E(y'(t) | \mathcal{F}_{t-2}) &= 2\gamma E\left(\sum_{s=0}^{t-1} F(t, s+1)C(s)G(s)w(s)\right) \\ &= 2\gamma \sum_{s=0}^{t-2} F(t, s+1)C(s)G(s) E(w(s) | \mathcal{F}_{t-2}) \end{aligned}$$

$$\begin{aligned}
& + 2\gamma F(t, t)C(t-1)G(t-1)E(w(t-1)|\mathcal{F}_{t-2}) \\
& = y'(t-1)
\end{aligned} \tag{4.49}$$

since $E(w(s)|\mathcal{F}_{t-2}) = w(s)$ and $E(w(t-1)|\mathcal{F}_{t-2}) = Ew(t-1) = 0_{n^2}$. Let $\|\cdot\|_1$ be the 1-norm of a vector. By (4.46) and the Cauchy–Schwartz inequality, we have

$$\begin{aligned}
E\|y'(t)\|_1 & \leq \sqrt{n} E\left[(y'(t)^\top y'(t))^{\frac{1}{2}}\right] \\
& \leq \sqrt{n} E\left[(y'(t)^\top y'(t))^{\frac{1}{2}}\right] \\
& \leq \frac{\sqrt{n\xi\Gamma(0)}}{\sqrt{\eta'\gamma - 4\xi\gamma^2}},
\end{aligned} \tag{4.50}$$

where the second inequality is due to Hölder's inequality in probability. Therefore, $\{y'(t)\}_{t \in \mathbb{N}}$ is a martingale (relevant to $\{\mathcal{F}_{t-1}\}_{t \geq 1}$). It follows from the aforementioned martingale convergence theorem that there exists a random vector $\hat{y} \in \mathbb{R}^n$ such that $y'(t) \rightarrow \hat{y}$ almost surely and in \mathcal{L}^2 as $t \rightarrow \infty$. Capitalizing (4.44), (4.47), (4.48) and $\lim_{t \rightarrow \infty} F(t, s) = 1_n \hat{\alpha}(s)$, we obtain that there exists a random variable \hat{x} such that $x_i(t) \rightarrow \hat{x}$ almost surely and in \mathcal{L}^2 as $t \rightarrow \infty$ for any $i \in V$. This finally concludes the proof of (i).

(ii) Define $M(s) := \text{diag}(\mu_1(s), \mu_2(s), \dots, \mu_n(s)) \in \mathbb{R}^{n \times n}$, where $\mu_i(s) = \sum_{j \in V} \left[\left(\sum_{k \in V} a_{ijk}(s) f(|x_j(s) - x_k(s)|) \right)^2 \cdot g(x_j(s) - x_i(s))^2 \right]$ for $i \in V$. The random matrix sequence $\{M(s)\}_{s \in \mathbb{N}}$ will play a central role in the derivation of properties of random variable \hat{x} .

Since $E\hat{y} = 0_n$ using Assumption 2 and the fact $y'(t) \rightarrow \hat{y}$ almost surely, we obtain by employing (4.44) that

$$\begin{aligned}
& \text{Cov}(\hat{y}) \\
& = E(\hat{y}\hat{y}^\top) \\
& = 4\gamma^2 \lim_{t \rightarrow \infty} \sum_{s=0}^{t-1} F(t, s+1) E\left(C(s)G(s)w(s)w(s)^\top G(s)^\top C(s)^\top\right) F(t, s+1)^\top \\
& = 4\gamma^2 \lim_{t \rightarrow \infty} \sum_{s=0}^{t-1} F(t, s+1) E(M(s)) F(t, s+1)^\top,
\end{aligned} \tag{4.51}$$

where $\text{Cov}(\cdot)$ represents the covariance matrix of a random vector. The largest singular value of $E(M(s))$ can be estimated as follows

$$\begin{aligned}
\sigma(E(M(s))) & = \max_{i \in V} |E\mu_i(s)| \\
& \leq n\beta_1^2 \beta_2^2 \max_{i \in V} \sum_{j \in V} E[(x_j(s) - x_i(s))^2] \\
& \leq \frac{2n^2 \beta_1^2 \beta_2^2}{\eta} E\Gamma(s)
\end{aligned} \tag{4.52}$$

by Assumption 1 and (4.38). Applying (4.43) and the uniform bound ℓ for the hypergraph connectivity, we write (4.52) further as

$$\sigma(\mathbb{E}(M(s))) \leq \frac{2n^2\beta_1^2\beta_2^2}{\eta}(1 - \eta'\gamma + 4\xi\gamma^2)^{\frac{1}{2}}\Gamma(0). \quad (4.53)$$

In view of (4.53) and the fact $\sigma(F(t, s+1)) = \sigma(F(t, s+1)^\top) < \infty$, we derive that the series $\lim_{t \rightarrow \infty} \sum_{s=0}^{t-1} \sigma(F(t, s+1) \mathbb{E}(M(s)) F(t, s+1)^\top) \leq \lim_{t \rightarrow \infty} \sum_{s=0}^{t-1} \sigma(F(t, s+1)) \sigma(\mathbb{E}(M(s))) \sigma(F(t, s+1)^\top) < \infty$. As the series is convergent, for any $\varepsilon > 0$, there is $s_1 \in \mathbb{N}$ such that for any $t > s_1$ we have

$$\sum_{s=s_1}^{t-1} \sigma(F(t, s+1) \mathbb{E}(M(s)) F(t, s+1)^\top) \leq \frac{\varepsilon}{3}. \quad (4.54)$$

Likewise, since $\sigma(1_n \hat{\alpha}(s')) = \sigma(\hat{\alpha}(s')^\top 1_n^\top) < \infty$ for any $s' \in \mathbb{N}$, we can choose the above s_1 sufficiently large such that we also have

$$\sum_{s=s_1}^{\infty} \sigma(1_n \hat{\alpha}(s) \mathbb{E}(M(s)) \hat{\alpha}(s)^\top 1_n^\top) \leq \frac{\varepsilon}{3}. \quad (4.55)$$

for any $s' \in \mathbb{N}$. Recall $\lim_{t \rightarrow \infty} F(t, s) = 1_n \hat{\alpha}(s)$ for any $s \in \mathbb{N}$ and therefore, there must exist $s_2 \geq s_1$ such that $\sigma(F(t, s) - 1_n \hat{\alpha}(s)) = \sigma(F(t, s)^\top - \hat{\alpha}(s)^\top 1_n^\top) \leq \varepsilon$ for any $t \geq s_2$. We can further choose $s_3 \geq s_2$ such that for any $t \geq s_3$,

$$\begin{aligned} & \sigma \left(\sum_{s=0}^{s_1-1} F(t, s+1) \mathbb{E}(M(s)) F(t, s+1)^\top \right. \\ & \quad \left. - \sum_{s=0}^{s_1-1} 1_n \hat{\alpha}(s+1) \mathbb{E}(M(s)) \hat{\alpha}(s+1)^\top 1_n^\top \right) \\ & \leq \sum_{s=0}^{s_1-1} \sigma(F(t, s+1) - 1_n \hat{\alpha}(s+1)) \\ & \quad \cdot \sigma(\mathbb{E}(M(s)) F(t, s+1)^\top - \mathbb{E}(M(s)) \hat{\alpha}(s+1)^\top 1_n^\top) \\ & \leq \varepsilon \cdot \sum_{s=0}^{s_1-1} \sigma(\mathbb{E}(M(s))) \cdot \sigma(F(t, s+1)^\top - \hat{\alpha}(s+1)^\top 1_n^\top) \\ & \leq \varepsilon^2 \sum_{s=0}^{s_1-1} \sigma(\mathbb{E}(M(s))) \leq \frac{2\varepsilon^2 n^2 \beta_1^2 \beta_2^2}{\eta(\eta'\gamma - 4\xi\gamma^2)} \Gamma(0), \end{aligned} \quad (4.56)$$

where we have employed properties of singular values and applied (4.53) in the last step. If we take

$$0 < \varepsilon \leq \frac{\eta(\eta'\gamma - 4\xi\gamma^2)}{6n^2\beta_1^2\beta_2^2\Gamma(0)}, \quad (4.57)$$

then the bound in (4.56) is no more than $\varepsilon/3$. Hence, combining (4.54), (4.55) and (4.56) we arrive at

$$\begin{aligned}
& \sigma \left\{ \sum_{s=0}^{t-1} F(t, s+1) E(M(s)) F(t, s+1)^\top \right. \\
& \quad \left. - \sum_{s=0}^{\infty} 1_n \hat{\alpha}(s+1) E(M(s)) \hat{\alpha}(s+1)^\top 1_n^\top \right\} \\
& \leq \sigma \left\{ \left(\sum_{s=0}^{s_1-1} + \sum_{s=s_1}^{t-1} \right) F(t, s+1) E(M(s)) F(t, s+1)^\top \right. \\
& \quad \left. - \left(\sum_{s=0}^{s_1-1} + \sum_{s=s_1}^{\infty} \right) 1_n \hat{\alpha}(s+1) E(M(s)) \hat{\alpha}(s+1)^\top 1_n^\top \right\} \\
& \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon
\end{aligned} \tag{4.58}$$

for any $t \geq s_3$. Therefore,

$$\begin{aligned}
& \lim_{t \rightarrow \infty} \sum_{s=0}^{t-1} F(t, s+1) E(M(s)) F(t, s+1)^\top \\
& = \left(\sum_{s=0}^{\infty} \hat{\alpha}(s+1) E(M(s)) \hat{\alpha}(s+1)^\top \right) 1_n 1_n^\top.
\end{aligned} \tag{4.59}$$

Invoking (4.44), we obtain

$$E \hat{x} 1_n = \lim_{t \rightarrow \infty} E x(t) = \lim_{t \rightarrow \infty} F(t, 0) x(0) + \lim_{t \rightarrow \infty} E y'(t) = 1_n \hat{\alpha}(0) x(0), \tag{4.60}$$

where we have applied the definition of $y'(t)$ and Assumption 2. This yields (3.7). On the other hand,

$$\begin{aligned}
& E(\hat{x}^2) 1_n 1_n^\top \leftarrow E(x(t)x(t)^\top) \\
& = E \left[(F(t, 0)x(0) + y'(t)) \cdot (x(0)^\top F(t, 0)^\top + y'(t)^\top) \right] \\
& = E \left(F(t, 0)x(0)x(0)^\top F(t, 0)^\top \right) + E \left(F(t, 0)x(0)y'(t)^\top \right) \\
& \quad + E \left(y'(t)x(0)^\top F(t, 0)^\top \right) + E \left(y'(t)y'(t)^\top \right).
\end{aligned} \tag{4.61}$$

The four terms on the right-hand side of (4.61) can be further calculated as follows. The first term tends to $1_n \hat{\alpha}(0) x(0) x(0)^\top \hat{\alpha}(0) 1_n^\top = (E \hat{x}^2) 1_n 1_n^\top$ as $t \rightarrow \infty$. The second and third terms are zero by Assumption 2. Hence, as $t \rightarrow \infty$,

$$E \left(y'(t)y'(t)^\top \right) \rightarrow (E \hat{x}^2) - (E \hat{x})^2 1_n 1_n^\top = (\text{Var } \hat{x}) 1_n 1_n^\top. \tag{4.62}$$

Employing (4.51), (4.59) and (4.62), we derive

$$(\text{Var } \hat{x}) 1_n 1_n^\top = \lim_{t \rightarrow \infty} E \left(y'(t)y'(t)^\top \right) = E \left(\hat{y} \hat{y}^\top \right)$$

$$=4\gamma^2 \left(\sum_{s=0}^{\infty} \hat{\alpha}(s+1) E(M(s)) \hat{\alpha}(s+1)^\top \right) \mathbf{1}_n \mathbf{1}_n^\top, \quad (4.63)$$

which gives rise to (3.8). The Proof of Theorem 2 is complete. \square

Instead of the system (3.3), if we consider the generalized system with the extra modulating function $h(\cdot)$, namely,

$$x(t+1) = B'(t)x(t) + 2\gamma C(t)G(t)w(t), \quad (4.64)$$

where $B'(t)$ is defined as in (4.18). We have the following result.

COROLLARY 3 Consider the system (4.64) under Assumptions 1, 2 and $0 < h(\cdot) \leq \beta_3$ for some $\beta_3 > 0$. Suppose that $\gamma < \min \left\{ (2\beta_3 \sup_{t \in \mathbb{N}} \max_{i \in V} \sum_{j,k \in V} a_{ijk}(t))^{-1}, \eta' \cdot (4\xi')^{-1} \right\}$, where $\eta' > 0$ is given in Corollary 2 and $\xi' > 0$ is a constant defined by (4.70). Assume that $\mathcal{H}(t)$ is jointly uniformly connected with bound ℓ and time sequence $\{t_l\}_{l \in \mathbb{N}}$. Then, the following hold:

- (i) There exists a random variable \hat{x} such that the almost sure consensus and \mathcal{L}^2 consensus are achieved.
- (ii) We have

$$E\hat{x} = \hat{\alpha}(0)x(0) \quad (4.65)$$

and

$$\text{Var} \hat{x} = 4\gamma^2 \sum_{s=0}^{\infty} \hat{\alpha}(s+1) (E M(s)) \hat{\alpha}(s+1)^\top, \quad (4.66)$$

where $\hat{\alpha}(s) = (\hat{\alpha}_1(s), \hat{\alpha}_2(s), \dots, \hat{\alpha}_n(s)) \in \mathbb{R}^{1 \times n}$ is the left eigenvector of the matrix $\lim_{t \rightarrow \infty} F(t, s)$ with respect to the eigenvalue 1 satisfying $\sum_{i \in V} \hat{\alpha}_i(s) = 1$ and $\hat{\alpha}_i(s) > 0$ for $i \in V$ and $s \in \mathbb{N}$. The matrix $M(s) \in \mathbb{R}^{n \times n}$ is defined as $M(s) := \text{diag}(\mu_1(s), \mu_2(s), \dots, \mu_n(s))$, where $\mu_i(s) = \sum_{j \in V} \left[\left(\sum_{k \in V} a_{ijk}(s) f(|x_j(s) - x_k(s)|) \right)^2 \cdot g(x_j(s) - x_i(s))^2 \right]$ for $i \in V$.

Proof. (i) The proof largely follows the same line in Theorem 2(i) by taking advantage of Corollary 2 instead of Theorem 1. The only difference lies in dealing with the transformation matrix $P = \begin{bmatrix} P_1 \\ P_2 \end{bmatrix}$, where $P_1 = [1, 0_{n-1}^\top] \in \mathbb{R}^{1 \times n}$ and $P_2 = [1_{n-1}, -I_{n-1}] \in \mathbb{R}^{(n-1) \times n}$. As $[x_1(t), z(t)^\top]^\top = Px(t) = [x_1(t), x_1(t) - x_2(t), x_1(t) - x_3(t), \dots, x_1(t) - x_n(t)]^\top \in \mathbb{R}^n$, multiplying P on both side of (4.64) gives $Px(t+1) = PB'(t)x(t) + 2\gamma PC(t)G(t)w(t)$. This means

$$\begin{aligned} x_1(t+1) &= x_1(t) - 2\gamma(l'_{11}(t)x_1(t) + l'_{12}(t)x_2(t) + \dots + l'_{1n}(t)x_n(t)) \\ &\quad + 2\gamma P_1 C(t)G(t)w(t) \\ &= x_1(t) - 2\gamma L^{P_1}(t)z(t) + 2\gamma P_1 C(t)G(t)w(t), \end{aligned} \quad (4.67)$$

where $L^{P_1}(t) := [-l'_{12}(t), -l'_{13}(t), \dots, -l'_{1n}(t)] \in \mathbb{R}^{1 \times (n-1)}$, and that

$$\begin{aligned}
& z(t+1) \\
& = z(t) - 2\gamma \begin{pmatrix} l'_{11}(t) - l'_{21}(t) & l'_{12}(t) - l'_{22}(t) & \cdots & l'_{1n}(t) - l'_{2n}(t) \\ l'_{11}(t) - l'_{31}(t) & l'_{12}(t) - l'_{32}(t) & \cdots & l'_{1n}(t) - l'_{3n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ l'_{11}(t) - l'_{n1}(t) & l'_{12}(t) - l'_{n2}(t) & \cdots & l'_{1n}(t) - l'_{nn}(t) \end{pmatrix}_{(n-1) \times n} \\
& \quad \cdot \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix} + 2\gamma P_2 C(t) G(t) w(t) \\
& = z(t) - 2\gamma \begin{pmatrix} l'_{22}(t) - l'_{12}(t) & l'_{23}(t) - l'_{13}(t) & \cdots & l'_{2n}(t) - l'_{1n}(t) \\ l'_{32}(t) - l'_{12}(t) & l'_{33}(t) - l'_{13}(t) & \cdots & l'_{3n}(t) - l'_{1n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ l'_{n2}(t) - l'_{12}(t) & l'_{n3}(t) - l'_{13}(t) & \cdots & l'_{nn}(t) - l'_{1n}(t) \end{pmatrix}_{(n-1) \times (n-1)} \\
& \quad \cdot \begin{pmatrix} x_1(t) - x_2(t) \\ x_1(t) - x_3(t) \\ \vdots \\ x_1(t) - x_n(t) \end{pmatrix} + 2\gamma P_2 C(t) G(t) w(t) \\
& := z(t) - 2\gamma L^{P_2}(t) z(t) + 2\gamma P_2 C(t) G(t) w(t), \tag{4.68}
\end{aligned}$$

where $L^{P_2}(t) \in \mathbb{R}^{(n-1) \times (n-1)}$ is defined as above.

Let $B^{P_2}(t) := I_{n-1} - 2\gamma L^{P_2}(t) \in \mathbb{R}^{(n-1) \times (n-1)}$. Instead of using ξ_2 in (4.34), we define

$$\xi'_2 = \sup_{t \in \mathbb{N}} \sigma(B^{P_2}(t)^\top B^{P_2}(t)) > 0. \tag{4.69}$$

The rest of the proof can be followed analogously. In particular, the constant ξ in (4.40) is replaced with ξ' defined as

$$\xi' = \frac{2^3(n-1)n^2\xi_1\beta_2^2}{\eta} \cdot \sum_{s=0}^{\ell-1} (\xi'_2 + 2^4\gamma^2\xi_3\beta_2^2n^2)^s > 0. \tag{4.70}$$

(ii) The second part of the proof follows verbatim as in Theorem 2(ii). \square

5. Concluding remarks

Higher-order interactions and models have emerged as a new frontier in modelling complex networked systems, as they represent more realistic backbones encoded in group interactions in real-world systems [42]. This article attempts an initial step in establishing analytical theory of consensus dynamics over networks with higher-order interactions accommodating non-linear modulating functions, temporal networks and random perturbation. We show that almost sure consensus and \mathcal{L}^2 consensus can be achieved

when the network topology represented by a temporal hypergraph is jointly uniformly connected. To model group interactions, three types of modulating functions have been introduced, which governs the 3-way coupling strength, in-group discrepancy and the noise intensity. Random hypergraphs can also be viewed as a special case in our framework. Building upon the mathematical tools of real analysis, matrix analysis, hypergraph, martingale and Lyapunov stability theory, we are able to characterize the statistics of the random consensus process and its equilibrium.

In the noise-free case, we note that the convergence is shown to be exponentially fast. It would be interesting to further investigate the consensus rate under random noises. From a mathematical point of view, an analysis of tail probability estimates and large deviation bounds regarding the equilibrium state would also be desirable. Since we have only considered 3-uniform hypergraphs, understanding how the topology of more general uniform hypergraphs and non-uniform hypergraphs affects the consensus dynamics is also a worthwhile future research direction.

Acknowledgements

The author is very thankful to the two anonymous reviewers for their careful reading and constructive comments.

Data availability

Data sharing is not applicable to this article, as no datasets were generated or analysed during the current study.

REFERENCES

1. NEWMAN, M. (2018) *Networks*, 2nd ed. Oxford: Oxford University Press.
2. BARABÁSI, A.-L. & ALBERT, R. (1999) Emergence of scaling in random networks. *Science*, **286**, 509–512.
3. WATTS, D. J. & STROGATZ, S. H. (1998) Collective dynamics of ‘small-world’ networks. *Nature*, **393**, 440–442.
4. BOCCALETTI, S., LATORA, V., MORENO, Y., CHAVEZ, M. & HWANG, D.-U. (2006) Complex networks: structure and dynamics. *Phys. Rep.*, **424**, 175–308.
5. BATTISTON, F., AMICO, E., BARRAT, A., BIANCONI, G., DE ARRUDA, G. F., FRANCESCHIELLO, B., IACOPINI, I., KÉFI, S., LATORA, V., MORENO, Y., MURRAY, M. M., PEIXOTO, T. P., VACCARINO, F. & PETRI, G. (2021) The physics of higher-order interactions in complex systems. *Nat. Phys.*, **17**, 1093–1098.
6. BATTISTON, F., CENCETTI, G., IACOPINI, I., LATORA, V., LUCAS, M., PATANIA, A., YOUNG, J.-G. & PETRI, G. (2020) Networks beyond pairwise interactions: structure and dynamics. *Phys. Rep.*, **874**, 1–92.
7. SHANG, Y. (2022) A system model of three-body interactions in complex networks: consensus and conservation. *Proc. R. Soc. A.*, **478**, 20210564.
8. TORRES, L., BLEVINS, A. S., BASSETT, D. & ELIASSI-RAD, T. (2021) The why, how and when of representation for complex systems. *SIAM Rev.*, **63**, 435–485.
9. ANDJELKOVIĆ, M., TADIĆ, B. & MELNIK, R. (2020) The topology of higher-order complexes associated with brain hubs in human connectomes. *Sci. Rep.*, **10**, 17320.
10. SIZEMORE, A. E., GIUSTI, C., KAHN, A., VETTEL, J. M., BETZEL, R. F. & BASSETT, D. S. (2018) Cliques and cavities in the human connectome. *J. Comput. Neurosci.*, **44**, 115–145.
11. ALVAREZ-RODRIGUEZ, U., BATTISTON, F., DE ARRUDA, G. F., MORENO, Y., PERC, M. & LATORA, V. (2021) Evolutionary dynamics of higher-order interactions in social networks. *Nat. Hum. Behav.*, **5**, 586–595.
12. KLIMM, F., DEANE, C. M. & REINERT, G. (2021) Hypergraphs for predicting essential genes using multiprotein complex data. *J. Complex Netw.*, **9**, cnaa028.
13. ANGULO, M. T., KELLEY, A., MONTEJANO, L., SONG, C. & SAAVEDRA, S. (2021) Coexistence holes characterize the assembly and disassembly of multispecies systems. *Nat. Ecol. Evol.*, **5**, 1091–1101.

14. LEVINE, J. M., BASCOMPTE, J., ADLER, P. B. & ALLESINA, S. (2017) Beyond pairwise mechanisms of species coexistence in complex communities. *Nature*, **546**, 56–64.
15. BENSON, A. R., GLEICH, D. F. & LESKOVEC, J. (2016) Higher-order organization of complex networks. *Science*, **353**, 163–166.
16. SHANG, Y. (2021) Consensus formation in networks with neighbor-dependent synergy and observer effect. *Commun. Nonlinear Sci. Numer. Simul.*, **95**, 105632.
17. VASCONCELOS, V. V., LEVIN, S. A. & PINHEIRO, F. L. (2019) Consensus and polarization in competing complex contagion processes. *J. R. Soc. Interface*, **16**, 20190196.
18. NEUHÄUSER, L., MELLOR, A. & LAMBIOTTE, R. (2020) Multibody interactions and nonlinear consensus dynamics on networked systems. *Phys. Rev. E*, **101**, 032310.
19. SAHASRABUDDHE, R., NEUHÄUSER, L. & LAMBIOTTE, R. (2021) Modelling non-linear consensus dynamics on hypergraphs. *J. Phys. Complex.*, **2**, 025006.
20. DEVRIENDT, K. & MIEGHEM, P. V. (2019) The simplex geometry of graphs. *J. Complex Netw.*, **7**, 469–490.
21. SHANG, Y. (2022) Sombor index and degree-related properties of simplicial networks. *Appl. Math. Comput.*, **419**, 126881.
22. GAMBUTTA, L. V., PATTI, F. D., GALLO, L., LEPRI, S., ROMANCE, M., CRIADO, R., FRASCA, M., LATORA, V. & BOCCALETTI, S. (2021) Stability of synchronization in simplicial complexes. *Nat. Commun.*, **12**, 1255.
23. GHORBANCHIAN, R., RESTREPO, J. G., TORRES, J. J. & BIANCONI, G. (2021) Higher-order simplicial synchronization of coupled topological signals. *Commun. Phys.*, **4**, 120.
24. SUN, H., ZIFF, R. M. & BIANCONI, G. (2020) Renormalization group theory of percolation on pseudofractal simplicial and cell complexes. *Phys. Rev. E*, **102**, 012308.
25. BURGIO, G., ARENAS, A., GÓMEZ, S., MATAMALAS, J. T. (2021) Network clique cover approximation to analyze complex contagions through group interactions. *Commun. Phys.*, **4**, 111.
26. CARLETTI, T., BATTISTON, F., CENCETTI, G. & FANELLI, D. (2020) Random walks on hypergraphs. *Phys. Rev. E*, **101**, 022308.
27. ST-ONGE, G., SUN, H., ALLARD, A., HÉBERT-DUFRESNE, L. & BIANCONI, G. (2021) Universal nonlinear infection kernel from heterogeneous exposure on higher-order networks. *Phys. Rev. Lett.*, **127**, 158301.
28. HOLME, P. & SARAMÁKI, J. (2012) Temporal networks. *Phys. Rep.*, **519**, 97–125.
29. LI, A., CORNELIUS, S. P., LIU, Y.-Y., WANG, L. & BARABÁSI, A.-L. (2017) The fundamental advantages of temporal networks. *Science*, **358**, 1042–1046.
30. ZHANG, Y., LATORA, V. & MOTTER, A. E. (2021) Unified treatment of synchronization patterns in generalized networks with higher-order, multilayer, and temporal interactions. *Commun. Phys.*, **4**, 195.
31. NEUHÄUSER, L., LAMBIOTTE, R. & SCHAUB, M. T. (2021) Consensus dynamics on temporal hypergraphs. *Phys. Rev. E*, **104**, 064305.
32. CHOWDHARY, S., KUMAR, A., CENCETTI, G., IACIPINI, I. & BATTISTON, F. (2021) Simplicial contagion in temporal higher-order networks. *J. Phys. Complex.*, **2**, 035019.
33. CENCETTI, G., BATTISTON, F., LEPRI, B. & KARSAI, M. (2021) Temporal properties of higher-order interactions in social networks. *Sci. Rep.*, **11**, 7028.
34. BRETTO, A. (2013) *Hypergraph Theory: An Introduction*. Switzerland: Springer.
35. SHANG, Y. (2013) Deffuant model with general opinion distributions: first impression and critical confidence bound. *Complexity*, **19**, 38–49.
36. CATTIAUX, P., DELEBECQUE, F. & PEDECHES, L. (2018) Stochastic Cucker–Smale models: old and new. *Ann. Appl. Probab.*, **28**, 3239–3286.
37. DURRETT, R. (2019) *Probability: Theory and Examples*, 5th edn. New York: Cambridge University Press.
38. COOLEY, O., GARBE, F., HNG, E. K., KANG, M., SANHUEZA-MATAMALA, N. & ZALLA, J. (2021) Longest paths in random hypergraphs. *SIAM J. Discrete Math.*, **35**, 2430–2458.
39. LIU, Y. P. & ZHAO, Y. (2021) On the upper tail problem for random hypergraphs. *Random Struct. Alg.*, **58**, 179–220.

40. HENDRICKX, J. M. & TSITSIKLIS, J. N. (2013) Convergence of type-symmetric and cut-balanced consensus seeking systems. *IEEE Trans. Autom. Contr.*, **58**, 214–218.
41. TOURI, B. & NEDIĆ, A. (2014) Product of random stochastic matrices. *IEEE Trans. Autom. Contr.*, **59**, 437–448.
42. MAJHI, S., PERC, M. & GHOSH, D. (2022) Dynamics on higher-order networks: a review. *J. R. Soc. Interface*, **19**, 20220043.