

Data Rate for Distributed Consensus of Multi-agent Systems with High Order Oscillator Dynamics

Zhirong Qiu¹, Lihua Xie^{†,1}, and Yiguang Hong²

Abstract

Distributed consensus with data rate constraint is an important research topic of multi-agent systems. Some results have been obtained for consensus of multi-agent systems with integrator dynamics, but it remains challenging for general high-order systems, especially in the presence of unmeasurable states. In this paper, we study the quantized consensus problem for a special kind of high-order systems and investigate the corresponding data rate required for achieving consensus. The state matrix of each agent is a $2m$ -th order real Jordan block admitting m identical pairs of conjugate poles on the unit circle; each agent has a single input, and only the first state variable can be measured. The case of harmonic oscillators corresponding to $m = 1$ is first investigated under a directed communication topology which contains a spanning tree, while the general case of $m \geq 2$ is considered for a connected and undirected network. In both cases it is concluded that the sufficient number of communication bits to guarantee the consensus at an exponential convergence rate is an integer between m and $2m$, depending on the location of the poles.

I. INTRODUCTION

Distributed consensus is a basic problem in distributed control of multi-agent systems, which aims to reach an interested common value of the states for a team of agents or subsystems by exchanging information with their neighbors. A variety of consensus protocols have been proposed for different kinds of applications; see the survey papers [2], [3], [4] and the reference therein. Nonetheless, to apply the consensus protocol in a digital network with limited bandwidth, it is necessary to introduce quantization and devise the corresponding encoding-decoding scheme. With static uniform quantization, quantized consensus was first studied in [5] to achieve the approximate average consensus for integer-valued agents by applying gossip algorithms. For a large class of averaging algorithms of real-valued agents, [6] established the bounds of the steady-state error and the convergence times, as well as their dependence on the number of quantization levels. Logarithmic quantizers with infinite quantization levels were adopted in [7] to guarantee the asymptotic average consensus. To achieve the asymptotic average consensus with finite quantization levels, a static finite-level uniform quantizer with a dynamic encoding scheme was proposed in [8], and used to show that an exponentially fast consensus can be ensured by finite-level quantizers for multi-agent systems with general linear dynamics, whether the state is fully measurable [9], or the state is only partially measurable and yet detectable [10]. However, the lower bound of sufficient data rate for the consensus obtained in these works are overly conservative, and it is more appealing to achieve the consensus with fewer bits of information exchange from the perspective of reducing communication load.

Some works have been devoted to exploring the sufficient data rate to guarantee the consensus of multi-agent systems with integrator dynamics, and single-integrator systems receive the most attention. With a presumed bound of the initial state of each agent, Li et. al. [8] showed that the average consensus can be achieved by 1 bit of information exchange for a fixed and undirected network, which was further extended to the case when the network is balanced and contains a spanning tree [11]. In an undirected network where the duration of link failure is bounded, 5-level quantizers suffices for the consensus [8], which also holds when the network is periodically strongly connected [12]. With a novel update protocol carefully screening the quantized message, the presumed bound of initial values was shown to be unnecessary in [13] and it was concluded that ternary messages are sufficient for the average consensus under a periodically connected network. Then, for double-integrator systems with only position being measurable, [14] concluded that 2 bits of communications suffice for the consensus. By employing a totally different technique based on matrix perturbation, n bits were found to be sufficient to achieve the consensus of multi-agent systems with n -th integrator agent dynamics in [15]. Still, it is unclear about the sufficient data rate to guarantee the consensus for general high-order systems, especially when the state variables are only partially measured.

In this paper, we explore the data rate problem in achieving quantized consensus of another kind of discrete-time high-order critical systems as a complement of integrator systems. The dynamics of each agent is described by a $2m$ -th order real Jordan block admitting m identical pairs of conjugate poles on the unit circle with single input, and only the first state variable can be measured. We design the encoding-decoding scheme on the basis of the constructability of the state variables of each

*This work is partially supported by National Research Foundation of Singapore under grant NRF-CRP8-2011-03 and NNSF of China under Grant 61174071. The work in this paper has been partially presented in [1].

[†]Author for correspondence.

¹Z. Qiu and L. Xie are with School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798. Email: qiuazhr@gmail.com; elhxie@ntu.edu.sg

²Y. Hong is with the Key Laboratory of Systems and Control, Institute of Systems Science, Chinese Academy of Sciences, Beijing 100190, China. Email: yghong@iss.ac.cn

individual system: at each time instant the quantizer will produce a signal to make an estimate for the current measurable state, which is combined with the previous $2m - 1$ estimates of the measurable state to obtain the estimate of the current full state. The same quantized signal will also be sent to neighbor agents to generate the identical estimate of state. The control input is constructed in terms of the estimate of its own state, as well as those of its neighbor agents. For harmonic oscillators ($m = 1$), it is shown that 2 bits of communications suffice to guarantee the exponentially fast consensus for a directed network containing a spanning tree. For higher-order case of $m \geq 2$, the exponentially fast consensus can be achieved with at most $2m$ bits under an undirected network, provided that the undirected communication topology is connected. The exact number of bits for achieving consensus in both the cases is an integer between m and $2m$, depending on the frequency of oscillators or the location of poles on the unit circle.

Although the analysis of consensus and data rate in this paper employs similar perturbation techniques as in [15], the problem posed here is much different, and it is much more challenging to obtain an explicit data rate required for consensus in the oscillator case (corresponding to complex eigenvalues). In contrast to [15] where the special structure of integrator dynamics enables a direct connection between the encoder's past outputs and those at the present moment which leads to a convenient iteration in the encoding scheme, a similar iteration is no longer available for the estimation of state variables in the case of oscillator dynamics. As such, a new observer-based encoding scheme is devised. However, such an encoding scheme leads to the involvement of control inputs into the estimation error, which makes the consensus analysis challenging. Furthermore, the expression of data rate for the oscillator case requires calculating a linear combination of some rows of a matrix which is a multiplication of the $(2m - 1)$ -th power of the system matrix and the inverse of the observability matrix, and is hard to obtain by a direct computation. To overcome this difficulty, we transform the linear combination into a set of linear equations and employ techniques of combinatorics. It is shown that a data rate between m and $2m$, depending on the frequencies of the oscillations, suffices to achieve the consensus. It is worthy noting that the result not only provides a sufficient data rate for consensus of the systems under consideration but also reveals an interesting connection between the data rate and the system dynamics. We believe it will shed some further light on the data rate problem for multi-agent systems of general dynamics.

The rest of the paper is organized as follows. Some preliminaries about graph theory and the problem formulation are presented in Section II. Then the data rate problem for distributed consensus of the coupled harmonic oscillators is conducted in Section III, which is followed by the general case of $m \geq 2$ in Section IV. For illustration, a numeric example is given in Section V. Some concluding remarks are drawn in Section VI. The proofs of the main lemmas can be found in the Appendix.

Some notations listed below will be used throughout this paper. For a matrix U , $U(i, j)$ and $U(i, \cdot)$ respectively denote its (i, j) -th entry and i -th row; U^T is its transpose, and $\|U\|$ is its infinity-norm. \mathbb{N}^+ is the set of positive integers, and $\lceil a \rceil$, $\lfloor a \rfloor$ respectively denote the smallest integer not less than a , and the largest integer not greater than a . $C(n, k)$ is the number of k -combinations from a given set of n elements. $\mathbf{1}_N$ is the N dimensional vector with every component being 1, and I_m is the identity matrix of order m . $j = \sqrt{-1}$ is the unit imaginary number. $J_{\lambda, n}$ denotes the n dimensional Jordan block with eigenvalue λ . $A \otimes B$ denotes the Kronecker product between matrices A and B . $\langle \cdot, \cdot \rangle$ denotes the standard inner product in Euclidean spaces.

II. PROBLEM FORMULATION

Consider a multi-agent system in the following form:

$$\begin{cases} x^i(t+1) = Ax^i(t) + bu_i(t), \\ y_i(t) = x_{i1}(t), \end{cases} \quad (1)$$

where $x^i(t) = [x_{i1}(t), x_{i2}(t), \dots, x_{i,2m}(t)]^T \in \mathbb{R}^{2m}$, $y_i(t), u_i(t) \in \mathbb{R}$, $i = 1, \dots, N$ represent the state, output and input of

agent i , respectively. Moreover, $A = \begin{bmatrix} Q & I_2 & & \\ & Q & \ddots & \\ & & \ddots & I_2 \\ & & & Q \end{bmatrix} \in \mathbb{R}^{2m \times 2m}$ is a real Jordan form consisting of m pairs of conjugate

eigenvalues $\cos \theta + j \sin \theta$ with $\sin \theta \neq 0$ and $Q = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$; $b = [0, \dots, 0, 1]^T \in \mathbb{R}^{2m}$.

Suppose that the total number of agents is N . Assumed to be error-free, the digital communication channels between agents are modeled as edges of a directed or undirected graph. A graph \mathcal{G} consists of a node set $\mathcal{V} = \{1, \dots, N\}$ and an edge set $\mathcal{E} = \{(i, v) : i, v \in \mathcal{V}\}$ where self-loop (i, i) is excluded. An edge (i, v) of a directed graph implies that node v can receive information from node i , but not necessarily vice versa. In contrast, for an undirected graph, $(i, v) \in \mathcal{E}$ means mutual communications between i and v . For node i , $\mathcal{N}_i^+ = \{v : (v, i) \in \mathcal{E}\}$ and $\mathcal{N}_i^- = \{v : (i, v) \in \mathcal{E}\}$ respectively denote its in-neighbors and out-neighbors, which coincide if \mathcal{G} is undirected, and will be denoted as \mathcal{N}_i . A directed path $(i_1, i_2), (i_2, i_3), \dots$ is formed by a sequence of edges. For a directed graph \mathcal{G} , if there exists a directed path connecting all the nodes, then \mathcal{G} is said to contain a spanning tree, which is equivalent to the case of being connected when \mathcal{G} is undirected.

Usually, a nonnegative matrix $G = [g_{iv}] \in \mathbb{R}^{N \times N}$ is assigned to the weighted graph \mathcal{G} , where $g_{iv} > 0$ if and only if $(v, i) \in \mathcal{E}$, and $g_{iv} = g_{vi}$ is further required for an undirected graph. The connectivity of \mathcal{G} can be examined from an algebraic

point of view, by introducing the Laplacian matrix $L = D^G - G$, where $D^G = \text{diag}(d_1^G, \dots, d_N^G)$ and $d_i^G = \sum_{v=1}^N g_{iv}$. By $L\mathbf{1}_N = 0$, L has at least one zero eigenvalue, with the other non-zero eigenvalues on the right half plane. L has only one zero eigenvalue if and only if \mathcal{G} contains a spanning tree [16]. We can always find a nonsingular matrix $U_L = [\phi_1 \ \phi_2 \ \dots \ \phi_N]$ with $\phi_1 = \mathbf{1}_N/\sqrt{N}$ and $\|\phi_i\| = 1$, such that $U_L^{-1}LU_L = \text{diag}\{J_{0,N_1}, J_{\lambda_2,N_2}, \dots, J_{\lambda_l,N_l}\} \triangleq L_J$, where $0 \leq \text{Re}\lambda_2 \leq \dots \leq \text{Re}\lambda_l$ with λ_i being an eigenvalue of L . In particular, we denote $\psi_i = U_L^{-1}(i, \cdot)^T$. Moreover, $L_J = \text{diag}\{0, \lambda_2, \dots, \lambda_N\}$ with $0 \leq \lambda_2 \leq \dots \leq \lambda_N$ and $\psi_i = \phi_i^T$ if \mathcal{G} is undirected.

We adopt the following finite-level uniform quantizer $q_t(\cdot)$ in the encoding scheme, where $M(t) \in \mathbb{N}^+$:

$$q_t(y) = \begin{cases} 0, & -\frac{1}{2} < y < \frac{1}{2}; \\ j, & \frac{2j-1}{2} \leq y < \frac{2j+1}{2}, \ j = 1, \dots, M(t) - 1; \\ M(t), & y \geq \frac{2M(t)-1}{2}; \\ -q_t(-y), & y \leq -\frac{1}{2}. \end{cases} \quad (2)$$

Remark 2.1: Clearly, the total number of quantization levels of $q_t(\cdot)$ is $2M(t) + 1$. Demanding that agent i does not send out any signal when the output is zero, it is enough to use $\lceil \log_2(2M(t)) \rceil$ bits to represent all the signals.

The problem of distributed quantized consensus is solved if we can design a distributed control protocol based on the outputs of the encoding-decoding scheme, making the states of different agents reach the agreement asymptotically:

$$\lim_{t \rightarrow \infty} [x^i(t) - x^j(t)] = 0, \quad i, j = 1, 2, \dots, N. \quad (3)$$

III. HARMONIC OSCILLATOR CASE

In this section, we will start with the harmonic oscillator case as an example to investigate how many bits of information exchange are enough to achieve consensus exponentially fast with quantized neighbor-based control. We separate it from higher-order cases due to its speciality and simplicity: the solution of this basic case not only provides a result under a directed communication topology, but also serves to facilitate the understanding of higher-order cases. Some relevant remarks will be included in the next section, as a comparison between second-order and higher-order cases, or a summary of general cases.

Note that now the system matrix $A = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$.

A. Encoding-decoding scheme and distributed control law

An encoding-decoding scheme has a paramount importance in the quantized consensus, which should not only provide estimates for all the states from the partially measurable states, but also help reduce the data rate. Accordingly, the encoder should serve as an observer based on iterations. To be specific, inspired by the constructability in the sense that the present state of the system can be recovered from the present and past outputs and inputs, namely

$$\begin{bmatrix} x_{i1}(t) \\ x_{i2}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\csc \theta & \cot \theta \end{bmatrix} \begin{bmatrix} x_{i1}(t-1) \\ x_{i1}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ u_i(t-1) \end{bmatrix}, \quad (4)$$

we propose the following encoder φ_i for agent i :

$$\begin{cases} s_i(1) = q_t(\frac{y_i(1)}{p(0)}), \ \hat{x}_{i1}(1) = p(0)s_i(1); \\ s_i(2) = q_t(\frac{y_i(2)}{p(1)}), \ \hat{x}_{i1}(2) = p(1)s_i(2), \\ \hat{x}_{i2}(2) = \cot \theta \hat{x}_{i1}(2) - \csc \theta \hat{x}_{i1}(1); \\ s_i(t) = q_t(\frac{y_i(t) - [\cos \theta \hat{x}_{i1}(t-1) + \sin \theta \hat{x}_{i2}(t-1)]}{p(t-1)}), \\ \hat{x}_{i1}(t) = \cos \theta \hat{x}_{i1}(t-1) + \sin \theta \hat{x}_{i2}(t-1) + p(t-1)s_i(t), \\ \hat{x}_{i2}(t) = \cot \theta \hat{x}_{i1}(t) - \csc \theta \hat{x}_{i1}(t-1), \ t \geq 2, \end{cases} \quad (5)$$

where $p(t) = p_0 \gamma^t$, $0 < \gamma < 1$ is a decaying scaling function.

After $s_i(t)$ is received by one of the i -th agent's out-neighbors, say $v \in \mathcal{N}_i^-$, a decoder φ_{iv} will be activated:

$$\begin{cases} \hat{x}_{iv1}(1) = p(0)s_i(1); \\ \hat{x}_{iv1}(2) = p(1)s_i(2), \ \hat{x}_{iv2}(2) = \cot \theta \hat{x}_{iv1}(2) - \csc \theta \hat{x}_{iv1}(1); \\ \hat{x}_{iv1}(t) = \cos \theta \hat{x}_{iv1}(t-1) + \sin \theta \hat{x}_{iv2}(t-1) + p(t-1)s_i(t), \\ \hat{x}_{iv2}(t) = \cot \theta \hat{x}_{iv1}(t) - \csc \theta \hat{x}_{iv1}(t-1), \ t \geq 2. \end{cases} \quad (6)$$

Remark 3.1: As in [15], a scaled "prediction error" is quantized to generate the signal $s_i(t)$, in an effort to reduce the number of quantization levels. $s_i(t)$ is then used to construct the estimate $\hat{x}_{i1}(t)$ of the first component $x_{i1}(t)$, which is combined with $\hat{x}_{i1}(t-1)$ to obtain the estimate $\hat{x}_{i2}(t)$ for $x_{i2}(t)$. Denote $\Delta_i(t) = s_i(t) - d_i(t)$ as the quantization error, where

$$d_i(t) = \begin{cases} \frac{y_i(t)}{p(t-1)}, \ t = 1, 2; \\ \frac{y_i(t) - [\cos \theta \hat{x}_{i1}(t-1) + \sin \theta \hat{x}_{i2}(t-1)]}{p(t-1)}, \ t > 2, \end{cases} \quad (7)$$

and $e_{ij}(t) = \hat{x}_{ij}(t) - x_{ij}(t)$ as the estimation for $x_{ij}(t)$, $j = 1, 2$. Then comparing (4), (5) and (6) we have

$$\begin{cases} e_{i1}(t) = \hat{x}_{i1}(t) - x_{i1}(t) = p(t-1)\Delta_i(t), & t \geq 1; \\ e_{i2}(t) = \hat{x}_{i2}(t) - x_{i2}(t) = \cot \theta e_{i1}(t) - \csc \theta e_{i1}(t-1) - u_i(t-1), & t \geq 2. \end{cases} \quad (8)$$

Evidently the estimation error is related with control inputs in addition to quantization errors, which may impair the consensus. But as shown in the consensus analysis below, the influence of the control inputs can be ignored by making the control gains arbitrarily small.

Based on the outputs of the encoding-decoding scheme, the distributed control law of agent i is given by

$$u_i(t) = \begin{cases} 0, & t = 0, 1; \\ \sum_{j=1}^2 k_j \sum_{v \in \mathcal{N}_i^+} g_{iv} [\hat{x}_{vij}(t) - \hat{x}_{ij}(t)], & t \geq 2. \end{cases} \quad (9)$$

B. Consensus Analysis and Data Rate

Some notations are defined as follows:

$$\begin{aligned} u(t) &= [u_1(t), \dots, u_N(t)]^T, \\ \Delta(t) &= [\Delta_1(t), \dots, \Delta_N(t)]^T, \\ d(t) &= [d_1(t), \dots, d_N(t)]^T, \\ x_j(t) &= [x_{1j}(t), \dots, x_{Nj}(t)]^T, \\ \delta_j(t) &= (I_N - \phi_1 \psi_1^T) x_j(t) = [\delta_{1j}(t), \dots, \delta_{Nj}(t)], \\ e_j(t) &= [e_{1j}(t), \dots, e_{Nj}(t)]^T. \end{aligned} \quad (10)$$

We adopt the following two assumptions in the subsequent analysis.

Assumption 3.1: The communication graph \mathcal{G} contains a spanning tree.

Assumption 3.2: There exist known positive constants C^* and C_δ^* such that $\max_{j=1,2} \|x_j(0)\| \leq C^*$ and $\max_{j=1,2} \|\delta_j(0)\| \leq C_\delta^*$.

Remark 3.2: Assumption 3.1 is a standard assumption, under which we have $0 < \text{Re}\lambda_2 \leq \dots \leq \text{Re}\lambda_l$, with 0 as the simple eigenvalue. Assumption 3.2 enables us to make the quantizer $q_t(\cdot)$ unsaturated at initial steps.

The following lemma is critical in the consensus analysis.

Lemma 3.1: Denote $K = \begin{bmatrix} 0 & 0 \\ k_1 & k_2 \end{bmatrix}$ and $A_i = A - \lambda_i K$ with $\text{Re}\lambda_i > 0$. Let $k_j = c_j \varepsilon$, $j = 1, 2$ and $\varepsilon > 0$. Then the following results hold with sufficiently small ε :

1). The spectral radius ρ_i of A_i is less than 1 if $c_2 \cos \theta - c_1 \sin \theta > 0$ and $c_1 \cos \theta + c_2 \sin \theta = 0$. Moreover, $\rho_i = 1 - \frac{1}{2}(\text{Re}\lambda_i)(c_2 \cos \theta - c_1 \sin \theta)\varepsilon + o(\varepsilon)$.

2). Take c_1, c_2 as in 1). For any vector $\xi \in \mathbb{R}^2$, the entries of $A_i^s \xi$, which are denoted as ξ_{s1} and ξ_{s2} , satisfy that $|\xi_{sj}| \leq 5/2\rho_i^s$ for $j = 1, 2$.

Proof: 1). Noticing that $A = P \begin{bmatrix} e^{j\theta} & \\ & e^{-j\theta} \end{bmatrix} P^{-1}$ with $P = \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix}$ and $P^{-1} = \frac{1}{2} \begin{bmatrix} 1 & -j \\ 1 & j \end{bmatrix}$, we have

$$\begin{aligned} \mu I - A &= P \begin{bmatrix} \mu - e^{j\theta} & \\ & \mu - e^{-j\theta} \end{bmatrix} P^{-1} \\ &= \frac{1}{2} \begin{bmatrix} (\mu - e^{j\theta}) + (\mu - e^{-j\theta}) & -j(\mu - e^{j\theta}) + j(\mu - e^{-j\theta}) \\ j(\mu - e^{j\theta}) - j(\mu - e^{-j\theta}) & (\mu - e^{j\theta}) + (\mu - e^{-j\theta}) \end{bmatrix}. \end{aligned} \quad (11)$$

Consequently the characteristic polynomial of A_i can be obtained as

$$\chi_i(\mu) = (\mu - e^{j\theta})(\mu - e^{-j\theta}) + \frac{\lambda_i}{2}\varepsilon[(c_2 + c_1 j)(\mu - e^{j\theta}) + (c_2 - c_1 j)(\mu - e^{-j\theta})]. \quad (12)$$

By perturbation theory [17] it is readily seen that the two perturbed roots of (12) are given by

$$\mu_{i1} = e^{j\theta} + \mu_{i11}\varepsilon + o(\varepsilon), \quad \mu_{i2} = e^{-j\theta} + \mu_{i21}\varepsilon + o(\varepsilon). \quad (13)$$

Substituting $\mu = \mu_{i1}$ into $\chi_i(\mu) = 0$ and comparing the coefficient of ε yield

$$\mu_{i11}(2j \sin \theta) + \frac{1}{2}\lambda_i(c_2 - c_1 j)(2j \sin \theta) = 0$$

and $\mu_{i11} = -\frac{1}{2}\lambda_i(c_2 - c_1 j)$ follows immediately. Direct computation shows that $|\mu_{i1}|^2 = 1 + 2\text{Re}(\mu_{i11}e^{-j\theta})\varepsilon + o(\varepsilon)$, where

$$\text{Re}(\mu_{i11}e^{-j\theta}) = -\frac{1}{2}[a_i(c_2 \cos \theta - c_1 \sin \theta) + b_i(c_1 \cos \theta + c_2 \sin \theta)]$$

if we let $\lambda_i = a_i + b_i j$. Clearly $|\mu_{i1}| = 1 - \frac{1}{2}(\text{Re}\lambda_i)(c_2 \cos \theta - c_1 \sin \theta)\varepsilon + o(\varepsilon)$ when $c_1 \cos \theta + c_2 \sin \theta = 0$. Similarly we can show $\mu_{i21} = -\frac{1}{2}\lambda_i(c_2 + c_1 j)$ and $|\mu_{i2}| = 1 - \frac{1}{2}\text{Re}\lambda_i(c_2 \cos \theta - c_1 \sin \theta)\varepsilon + o(\varepsilon)$, which implies the conclusion.

2). Here we need to compute the Jordan decomposition of A_i . The eigenvector corresponding to the eigenvalue μ_{i1} is given by $w_{i1} = w_{i10} + w_{i11}\varepsilon + o(\varepsilon)$. Substituting it into the equation $A_i w_{i1} = \mu_{i1} w_{i1}$ and comparing the coefficients of constant term, we have $A w_{i10} = e^{j\theta} w_{i10}$. With the normalization condition $v^T w_{i1} = 1$ where $v^T = \frac{1}{2}(1 - j)$, $w_{i10} = (1 \ j)^T$. Similarly, the eigenvector corresponding to the eigenvalue μ_{i2} is given by $w_{i2} = w_{i20} + w_{i21}\varepsilon + o(\varepsilon)$ with $w_{i20} = (1 \ -j)^T$. Letting $R_i = (w_{i1} \ w_{i2}) = \begin{bmatrix} 1 + O(\varepsilon) & 1 + O(\varepsilon) \\ j + O(\varepsilon) & -j + O(\varepsilon) \end{bmatrix}$, it is clear that $R_i^{-1} = \frac{1}{\det R_i} \begin{bmatrix} -j + O(\varepsilon) & -1 + O(\varepsilon) \\ -j + O(\varepsilon) & 1 + O(\varepsilon) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 + O(\varepsilon) & -j + O(\varepsilon) \\ 1 + O(\varepsilon) & j + O(\varepsilon) \end{bmatrix}$. The result follows directly by noticing that $A_i = R_i \begin{bmatrix} \mu_{i1} & \\ & \mu_{i2} \end{bmatrix} R_i^{-1}$. ■

Remark 3.3: Denote $\rho = \max_{i=2,\dots,l} \rho_i$ and let h be a constant in $(0, \text{Re}\lambda_2]$. Taking $c_1 = -\sin\theta/h$ and $c_2 = \cos\theta/h$, we have $\rho \leq 1 - \varepsilon + o(\varepsilon) < 1 - \varepsilon/2$ with sufficiently small ε .

We also need to define some constants as follows:

$$\begin{aligned} C_0 &= \frac{1}{2}|c_1| + \frac{3}{2}|c_2 \csc\theta| = \frac{1}{2h}(|\sin\theta| + 3|\cot\theta|), \\ \Lambda &= \max_{i=2,\dots,l} |\lambda_i|, \\ C(1) &= \|U_L^{-1}\| + 2C_0\Lambda\|U_L\|, \\ C(k) &= \|U_L^{-1}\| + 2C_0(\Lambda + 1)\|U_L\| + 10(|c_1| + |c_2|)C(k-1), \quad k \geq 2, \\ \bar{C} &= 5(|c_1| + |c_2|)C(N_{\max}) + C_0\|U_L\|, \end{aligned} \tag{14}$$

where $N_{\max} = \max_{i=2,\dots,l} N_i$.

Lemma 3.2: Let $\gamma = 1 - \varepsilon/4$. Then we can choose sufficiently small ε to satisfy the following inequalities:

$$(\Lambda + 1)\bar{C}\varepsilon \leq \frac{1}{2}\gamma|\csc\theta|\|U_L\|; \tag{15a}$$

$$\frac{1}{\gamma}(2|\cos\theta| + \frac{1}{\gamma}) \leq 2|\cos\theta| + 1 + \frac{1}{2}; \tag{15b}$$

$$(N - 1)\bar{C}(\Lambda + 1)\varepsilon \leq \frac{1}{4}|\csc\theta|\gamma^3. \tag{15c}$$

Theorem 3.1: Take c_j 's as in Remark 3.3 and let $\gamma = 1 - \varepsilon/4$. Select sufficiently small ε to satisfy Lemmas 3.1 and 3.2 with $\rho < 1 - \varepsilon/2$. Then under Assumptions 3.1 and 3.2, consensus can be achieved at a convergence rate of $O(\gamma^t)$ provided that $g_0 \geq \max\{\frac{4}{3\gamma}C^*, C_\delta^*\}$ and $M(t)$ satisfies

$$\begin{cases} M(t) \geq 1, & t = 1, 2; \\ M(t) \geq |\cos\theta| + 1/2, & t = 2m + 1, \dots \end{cases} \tag{16}$$

Therefore, the number of bits used to achieve the consensus is $\lceil \log_2 2 \lceil |\cos\theta| + 1/2 \rceil \rceil$.

Proof: 1) Preparation. The closed-loop system of disagreement vectors can be established as

$$\begin{bmatrix} \delta_1(t+1) \\ \delta_2(t+1) \end{bmatrix} = \begin{bmatrix} \cos\theta I_N & \sin\theta I_N \\ -\sin\theta I_N & \cos\theta I_N \end{bmatrix} \begin{bmatrix} \delta_1(t) \\ \delta_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ u(t) \end{bmatrix}$$

with

$$u(t) = \begin{cases} 0, & t = 0, 1; \\ -\sum_{j=1}^2 k_j L(\delta_j(t) + e_j(t)), & t \geq 2, \end{cases} \tag{17}$$

by noticing (9) and $L = L(I_N - \phi_1 \psi_1^T)$. Letting $\tilde{\delta}_j(t) = U_L^{-1} \delta_j(t) = [\tilde{\delta}_{1j}(t), \dots, \tilde{\delta}_{N,j}(t)]^T$, we obtain

$$\begin{bmatrix} \tilde{\delta}_1(t+1) \\ \tilde{\delta}_2(t+1) \end{bmatrix} = \begin{bmatrix} \cos\theta I_N & \sin\theta I_N \\ -\sin\theta I_N - k_1 L_J & \cos\theta I_N - k_2 L_J \end{bmatrix} \begin{bmatrix} \tilde{\delta}_1(t) \\ \tilde{\delta}_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \eta(t) \end{bmatrix},$$

where $\eta(t) = -L_J U_L^{-1}(k_1 e_1(t) + k_2 e_2(t))$. Denote $\tilde{\delta}^i(t) = [\tilde{\delta}_{i1}(t), \tilde{\delta}_{i2}(t)]^T$ for $i = 1, \dots, N$. Clearly $\tilde{\delta}^1(t) \equiv 0$ due to that $\tilde{\delta}_{1j}(t) = \psi_1^T(I_N - \phi_1 \psi_1^T)x_j(t) = 0$ for $j = 1, 2$. Without loss of generality we assume that $N_2 = 2$ (the Jordan block with respect to λ_2 is two-dimensional) and consequently $\tilde{\delta}^2(t)$ and $\tilde{\delta}^3(t)$ are coupled in the following way:

$$\begin{aligned} \tilde{\delta}^i(t+1) &= A\tilde{\delta}^i(t), \quad t = 0, 1, \quad i = 2, 3; \\ \tilde{\delta}^2(t+1) &= A_2\tilde{\delta}^2(t) - K\tilde{\delta}^3(t) - \eta_2(t), \\ \tilde{\delta}^3(t+1) &= A_2\tilde{\delta}^3(t) - \eta_3(t), \quad t \geq 2, \end{aligned} \tag{18}$$

where A_2 and K have been defined in Lemma 3.1 and $\eta_2(t) = [0, (\lambda_2 \psi_2^T + \psi_3^T)(k_1 e_1(t) + k_2 e_2(t))]^T$, $\eta_3(t) = [0, \lambda_2 \psi_3^T(k_1 e_1(t) + k_2 e_2(t))]^T$.

2) Estimation error and exponential convergence. Remember that $e_2(t) = \cot\theta e_1(t) - \csc\theta e_1(t-1) - u(t-1)$ is dependent on the control input by (8), we have to first make an estimate for $u(t)$ before establishing the consensus result. Below we shall show $|\psi_i^T u(t)| \leq \varepsilon(\Lambda + 1)\bar{C}\rho_0\gamma^{t-2}$, $t \geq 2$ for $i \geq 2$ by induction.

With the choice of p_0 and γ it is easy to see $|s_i(t)| \leq 3/2$ when $t \leq 2$ by noticing $|y_i(t)| \leq 2C^*$, hence we obtain $\max_{t=1,2} \|\Delta(t)\| \leq 1/2$ if $M(1), M(2) \geq 1$. For $i = 2$, we have $\psi_2^T L = \lambda_2 \psi_2^T + \psi_3^T$ and as a result

$$\begin{aligned} |\psi_2^T u(2)| &= |\psi_2^T L \sum_{j=1}^2 k_j (\delta_j(2) + e_j(2))| \\ &\leq \varepsilon(\Lambda + 1) \|U_L\| [|c_1| (2\|\delta(0)\| + p(1)\|\Delta(2)\|) \\ &\quad + |c_2| (2\|\delta(0)\| + |\cot \theta| p(1)\|\Delta(2)\| + |\csc \theta| p(0)\|\Delta(1)\|)] \\ &\leq \varepsilon(\Lambda + 1) \|U_L\| (|c_1| + |c_2|) (2\|\delta(0)\| + |\csc \theta| p_0) \\ &\leq \varepsilon(\Lambda + 1) \|U_L\| (|c_1| + |c_2|) (2 + |\csc \theta|) p_0 \\ &\leq \varepsilon(\Lambda + 1) \bar{C} p_0, \end{aligned}$$

which also holds for $|\psi_i^T u(2)|$ for $i > 2$.

Now assume that

$$\begin{aligned} |\psi_i^T u(t)| &\leq \varepsilon(\Lambda + 1) \bar{C} p_0 \gamma^{t-2}, \quad t \geq 2; \\ \|\Delta(\tau)\| &\leq 1/2 (\Rightarrow |e_{i1}(\tau)| \leq \frac{1}{2} p_0 \gamma^{\tau-1}), \quad 1 \leq \tau \leq t. \end{aligned} \quad (19)$$

Then by combining (19) and (15a) it follows that

$$\begin{aligned} \|\eta_3(\tau)\| &= |\lambda_2| |\psi_3^T (k_1 e_1(\tau) + k_2 e_2(\tau))| \\ &\leq \varepsilon |\lambda_2| p_0 \gamma^{\tau-2} \|U_L\| (\frac{1}{2} |c_1| + |c_2| |\csc \theta|) + \varepsilon |\lambda_2| |c_2| |\psi_3^T u(\tau-1)| \\ &\leq \varepsilon |\lambda_2| \|U_L\| C_0 p_0 \gamma^{\tau-2}. \end{aligned} \quad (20)$$

Recalling (18) we get that for $t \geq 2$

$$\tilde{\delta}^3(t+1) = A_2^{t-1} \tilde{\delta}^3(2) - \sum_{\tau=1}^{t-1} A_2^{t-1-\tau} \eta_3(\tau+1),$$

which produces the following estimate by Lemma 3.1 and (20)

$$\begin{aligned} \|\tilde{\delta}^3(t+1)\| &\leq \frac{5}{2} (\rho_2^{t-1} \|\tilde{\delta}^3(2)\| + 4 |\lambda_2| C_0 \|U_L\| p_0 \gamma^{t-1}) \\ &\leq 5 \gamma^{t-1} (\|U_L^{-1}\| \|\delta(0)\| + 2 |\lambda_2| C_0 \|U_L\| p_0) \\ &\leq 5C(1) p_0 \gamma^{t-1}. \end{aligned} \quad (21)$$

Similarly, an estimate for $\|\tilde{\delta}_2(t+1)\|$ can be found as $\|\tilde{\delta}_2(t+1)\| \leq 5C(2) p_0 \gamma^{t-1}$, if we notice that $\|K \tilde{\delta}^3(\tau)\| \leq 5\varepsilon(|c_1| + |c_2|) C(1) p_0 \gamma^{\tau-2}$ and $\|\eta_2(\tau)\| \leq \varepsilon(|\lambda_2| + 1) \|U_L\| C_0 p_0 \gamma^{\tau-2}$ for $2 \leq \tau \leq t$. For any $i \geq 2$, by proceeding along the same line as in the above it is concluded that

$$\|\tilde{\delta}^i(t+1)\| \leq 5C(N_{\max}) p_0 \gamma^{t-1}. \quad (22)$$

3) Data rate. Now we are able to discuss the estimation for $|\psi_i^T u(t+1)|$, which is bounded by the sum of $|\sum_{j=1}^2 k_j \psi_i^T L \delta_j(t+1)|$ and $|\sum_{j=1}^2 k_j \psi_i^T L e_j(t+1)|$. For the first term, by (22) it is readily seen that

$$\begin{aligned} |\sum_{j=1}^2 k_j \psi_i^T L \delta_j(t+1)| &\leq |\sum_{j=1}^2 k_j \psi_i^T U_L L_J \tilde{\delta}_j(t+1)| \\ &\leq 5\varepsilon(|c_1| + |c_2|) C(N_{\max}) p_0 \gamma^{t-1}; \end{aligned} \quad (23)$$

while the second term is essentially related with $e_j(t+1)$, or more exactly $\Delta(t+1)$. By (7) and (8) we have

$$\begin{aligned} d(t+1) &= \frac{1}{p(t)} (-\cos \theta e_1(t) - \sin \theta e_2(t)) \\ &= \frac{-2 \cos \theta}{\gamma} \Delta(t) + \frac{1}{\gamma^2} \Delta(t-1) + \frac{1}{p(t)} \sin \theta u(t-1), \end{aligned} \quad (24)$$

which is obviously dependent on the previous quantization errors $\Delta(t)$ and $\Delta(t-1)$, as well as the previous control input $u(t-1)$. Hence with the induction assumption (19) the quantizer can be made unsaturated with sufficiently many bits at time $t+1$, and $\|\Delta(t+1)\| \leq 1/2$ follows directly. Consequently

$$|\sum_{j=1}^2 k_j \psi_i^T L e_j(t+1)| \leq \varepsilon |\lambda_2| \|U_L\| C_0 p_0 \gamma^{t-1} \quad (25)$$

as in (20). The induction is then established by combining (23). Moreover, by (22) the consensus can be achieved at a convergence rate of $O(\gamma^t)$.

Below we are to calculate the number of required quantization levels at each time step. The situation when $t \leq 2$ has been discussed. When $t > 2$, from (24) we can see that

$$\begin{aligned} \|d(t)\| &\leq \frac{1}{2\gamma} (2|\cos \theta| + \frac{1}{\gamma}) + \frac{|\sin \theta|}{p(t-1)} \|u(t-2)\| \\ &\leq \frac{1}{2} (2|\cos \theta| + 1) + \frac{1}{4} + \frac{|\sin \theta|}{\gamma^3} (N-1)(\Lambda+1) \bar{C} \\ &\leq \frac{1}{2} (2|\cos \theta| + 1) + \frac{1}{2} \end{aligned}$$

by noticing (15b), (15c) and $u(t) = \sum_{i=2}^N \phi_i \psi_i^T u(t)$ ($\psi_1^T u(t) = 0$). In summary, the proof is completed. \blacksquare

Remark 3.4: For the coupling system shown in (18), we divide it into two subsystems with disturbance. Each subsystem can be stabilized as long as the disturbance decays exponentially at a speed slower than ρ_2 , i.e. $\|\eta_3(t)\| \sim O(\gamma^t)$ and $\|K\tilde{\delta}^3(t) + \eta_2(t)\| \sim O(\gamma^t)$, with $\rho_2 < \gamma < 1$. The interference of $u(t)$ in the estimation error $e(t)$ can be ignored, as long as $\|u(t)\| \sim O(\varepsilon^\alpha)p_0\gamma^t$ with $\alpha > 0$, yielding that $\|\eta_3(t)\| \sim O(\varepsilon)p_0\gamma^t$, and then $\|\tilde{\delta}^3(t)\| \sim O(\gamma^t)$ follows. As a result, $\|K\tilde{\delta}^3(t)\| \sim O(\varepsilon)\gamma^t$ and $\|\tilde{\delta}^3(t)\| \sim O(\gamma^t)$ follows by combining $\|\eta_3(t)\| \sim O(\varepsilon)p_0\gamma^t$. Such a reasoning still applies when (18) involves more than two subsystems. Finally we show that $\|u(t)\| \sim O(\varepsilon)p_0\gamma^t$, and by (24) we conclude that the control input does not consume extra bits in exchanging the information when the control gains are sufficiently small.

IV. HIGHER-ORDER CASES

In this section, we will conduct the same task as in the last section for general higher-order cases. The analysis actually proceeds along a similar line, but the assignment of control gains to achieve consensus is much more challenging, and we have to resort to combinatorial identities for an explicit data rate. As before, we first provide an encoding-decoding scheme for all the agents and devise a control protocol in terms of the outputs of the scheme. Then we present some lemmas, which will play a crucial role in the convergence analysis and the derivation of the data rate in the final part.

A. Encoding-decoding scheme and distributed control law

As pointed out in the last section, the construction of the encoding scheme should follow two principles: firstly, the encoder is able to estimate other state variables given that only the first component is measurable; secondly, the estimation should be based on iterations in an effort to reduce quantization levels. Such an idea can be stated more clearly as follows. At each time step, the scaled difference between the output $y_i(t)$ and its estimate is quantized to obtain a signal $s_i(t)$. Based on $s_i(t)$ we construct an estimate $\hat{x}_{i1}(t)$ of the first component $x_{i1}(t)$, and combine previous estimates $\hat{x}_{i1}(t-1)$ through $\hat{x}_{i1}(t-2m+1)$ to obtain estimates of the other components $x_{i2}(t)$ through $x_{i,2m}(t)$.

$$\text{To be detailed, denote the observability matrix } \mathcal{O} = \begin{bmatrix} I_{2m}(1, \cdot) \\ A(1, \cdot) \\ \vdots \\ A^{2m-1}(1, \cdot) \end{bmatrix},$$

$$\begin{aligned} \bar{x}_i(t) &= [x_{i1}(t-2m+1), x_{i1}(t-2m+2), \dots, x_{i1}(t)]^T, \\ b_n(\theta) &= [0, \dots, 0, A(1, 2m), \dots, A^{n-1}(1, 2m)]^T \in \mathbb{R}^{2m}. \end{aligned}$$

We have

$$\bar{x}_i(t) = \mathcal{O}x^i(t-2m+1) + \sum_{n=1}^{2m-1} b_n(\theta)u_i(t-n) \quad (26)$$

if we notice by (1) that

$$\begin{aligned} x^i(t-2m+k) &= A^{k-1}x^i(t-2m+1) + \sum_{n=0}^{k-2} A^{k-2-n}bu_i(t-2m+1+n), \\ k &= 1, \dots, 2m. \end{aligned} \quad (27)$$

As a result,

$$x^i(t-2m+1) = \mathcal{O}^{-1}[\bar{x}_i(t) - \sum_{n=1}^{2m-1} b_n(\theta)u_i(t-n)] \quad (28)$$

and

$$\begin{aligned} x^i(t) &= A^{2m-1}x^i(t-2m+1) + \sum_{j=0}^{2m-2} A^{2m-2-j}bu_i(t-2m+1+j) \\ &= S\bar{x}_i(t) + \sum_{j=1}^{2m-1} \tilde{b}_j(\theta)u_i(t-j), \end{aligned} \quad (29)$$

where $S = A^{2m-1}\mathcal{O}^{-1}$ (the existence of \mathcal{O}^{-1} can be easily verified by PBH test [18] if $\sin \theta \neq 0$) and $\tilde{b}_n(\theta) = -Sb_n(\theta) + A^{n-1}(\cdot, 2m)$. Inspired by (29), the encoding scheme for agent i is implemented below: for $t \leq 2m$,

$$\left\{ \begin{array}{l} s_i(t) = q_t\left(\frac{y_i(t)}{p(t-1)}\right), \quad \hat{x}_{i1}(t) = p(t-1)s_i(t); \\ \begin{bmatrix} \hat{x}_{i2}(2m) \\ \vdots \\ \hat{x}_{i,2m}(2m) \end{bmatrix} = S_m \begin{bmatrix} \hat{x}_{i1}(1) \\ \vdots \\ \hat{x}_{i1}(2m) \end{bmatrix}; \end{array} \right. \quad (30)$$

for $t > 2m$,

$$\begin{cases} s_i(t) = q_t \left(\frac{y_i(t) - [\cos \theta \hat{x}_{i1}(t-1) + \sin \theta \hat{x}_{i2}(t-1) + \hat{x}_{i3}(t-1)]}{p^{(t-1)}} \right), \\ \hat{x}_{i1}(t) = \cos \theta \hat{x}_{i1}(t-1) + \sin \theta \hat{x}_{i2}(t-1) + \hat{x}_{i3}(t-1) + p(t-1)s_i(t), \\ \begin{bmatrix} \hat{x}_{i2}(t) \\ \vdots \\ \hat{x}_{i,2m}(t) \end{bmatrix} = S_m \begin{bmatrix} \hat{x}_{i1}(t-2m+1) \\ \vdots \\ \hat{x}_{i1}(t) \end{bmatrix}, \end{cases} \quad (31)$$

where $S_m = S(2 : 2m, \cdot)$ is a submatrix of S obtained by deleting the first row, and $p(t) = p_0 \gamma^t$, $0 < \gamma < 1$ is a decaying scaling function.

After $s_i(t)$ is generated, transmitted and received by one of agent i 's out-neighbors, say $v \in \mathcal{N}_i^-$, a decoder will be activated: for $t \leq 2m$,

$$\begin{cases} \hat{x}_{iv1}(t) = p(t-1)s_i(t); \\ \begin{bmatrix} \hat{x}_{iv2}(2m) \\ \vdots \\ \hat{x}_{iv,2m}(2m) \end{bmatrix} = S_m \begin{bmatrix} \hat{x}_{iv1}(1) \\ \vdots \\ \hat{x}_{iv1}(2m) \end{bmatrix}; \end{cases} \quad (32)$$

for $t > 2m$,

$$\begin{cases} \hat{x}_{iv1}(t) = \cos \theta \hat{x}_{iv1}(t-1) + \sin \theta \hat{x}_{iv2}(t-1) + \hat{x}_{iv3}(t-1) + p(t-1)s_i(t); \\ \begin{bmatrix} \hat{x}_{iv2}(t) \\ \vdots \\ \hat{x}_{iv,2m}(t) \end{bmatrix} = S_m \begin{bmatrix} \hat{x}_{iv1}(t-2m+1) \\ \vdots \\ \hat{x}_{iv1}(t) \end{bmatrix}. \end{cases} \quad (33)$$

Remark 4.1: Comparing (30) with (32), (31) with (33), it is clear that $\hat{x}_{ivj}(t) \equiv \hat{x}_{ij}(t)$, $j = 1, \dots, 2m$, for $v \in \mathcal{N}_i^-$, $i = 1, \dots, N$. Denote $e_{ij}(t) = \hat{x}_{ij}(t) - x_{ij}(t)$ as the estimation error, $\Delta_i(t) = s_i(t) - d_i(t)$ as the quantization error, where

$$d_i(t) = \begin{cases} \frac{y_i(t)}{p^{(t-1)}}, & t = 1, 2, \dots, 2m; \\ \frac{y_i(t) - [\cos \theta \hat{x}_{i1}(t-1) + \sin \theta \hat{x}_{i2}(t-1) + \hat{x}_{i3}(t-1)]}{p^{(t-1)}}, & t > 2m. \end{cases} \quad (34)$$

Comparing (29) with (31), the estimation errors are given by the following:

$$\begin{cases} e_{i1}(t) = p(t-1)\Delta_i(t), & t \geq 1; \\ e_{ij}(t) = \sum_{n=1}^{2m} S(j, n)e_{i1}(t-2m+n) - \sum_{n=1}^{2m-1} \tilde{b}_{nj}(\theta)u_i(t-n), \\ t \geq 2m, j = 2, \dots, 2m, \end{cases} \quad (35)$$

where \tilde{b}_{nj} is the j -th entry of \tilde{b}_n .

Remark 4.2: The encoding schemes (5) and (31) proposed in our work is different from those in [10] or [15]. Actually, to address the general dynamics with unmeasurable states, [10] designed the encoding scheme respectively for the output and control input, and used Luenberger observer to estimate the unmeasurable states. If we compare with [15], we can also see a big difference: the special structure of n -th order integrator dynamics enables it to easily “recover” the control input at n steps earlier, based on which an estimate of the unmeasurable components can be made with time delay, and the encoding scheme can be designed accordingly. However, in our case it is unlikely to achieve the same task and we resort to the constructability of the system, namely we estimate the unmeasurable states directly from $\hat{x}_{i1}(t)$ through $\hat{x}_{i1}(t-2m+1)$. Although such a method introduces the control input into the estimation errors, it is able to make an estimation without time delay, and hence avoids the stabilization of a time-delayed closed-loop system in the consensus analysis.

For agent i , the outputs of encoder are $\hat{x}_{i1}(t), \dots, \hat{x}_{i,2m}(t)$, while the outputs of decoders are $\hat{x}_{vi1}(t), \dots, \hat{x}_{vi,2m}(t)$ for $v \in \mathcal{N}_i^+$. Based on these outputs, the distributed control law of agent i is proposed as

$$u_i(t) = \begin{cases} 0, & t = 0, 1, \dots, 2m-1; \\ \sum_{j=1}^{2m} k_j \sum_{v \in \mathcal{N}_i^+} g_{iv} [\hat{x}_{vij}(t) - \hat{x}_{ij}(t)], & t \geq 2m. \end{cases} \quad (36)$$

B. Lemmas

The following two lemmas are respectively needed in analyzing consensus and data rate. The first one is to stabilize the closed-loop system of disagreements, and the second one is used for estimating the magnitude of $u_i(t)$ and $d_i(t)$.

Lemma 4.1: Denote $A_i = A - \lambda_i K$ with $\lambda_i > 0$, where $K \in \mathbb{R}^{2m \times 2m}$ and its nonzero entries are only at the last row $[k_1, k_2, \dots, k_{2m-1}, k_{2m}]$. Take

$$\begin{aligned} k_{2j-1} &= \begin{cases} c_{2j-1}\varepsilon^{m-j}, & j = 1, \dots, m-1; \\ c_{2j-1}\varepsilon, & j = m, \end{cases} \\ k_{2j} &= \begin{cases} c_{2j}\varepsilon^{m-j}, & j = 1, \dots, m-1; \\ c_{2j}\varepsilon, & j = m. \end{cases} \end{aligned} \quad (37)$$

Then we can find constants c_{2j-1} and c_{2j} ($j = 1, \dots, m$) such that, when ε is sufficiently small, the spectral radius ρ_i of A_i is less than 1 with distinct eigenvalues. Moreover, denote

$$R_m = \frac{1}{2} + \frac{1}{2}(c_{2m-1} \sin \theta - c_{2m} \cos \theta), \quad H = Re\left[\frac{c_{2m-5} - c_{2m-4}}{c_{2m-3} - c_{2m-2}} e^{-j\theta}\right]. \quad (38)$$

The requirements about c_{2j-1} 's and c_{2j} 's corresponding to different m 's are listed below.

- 1). $m = 2$: let $c_1 = -\sin 2\theta$ and $c_2 = \cos 2\theta$. If $R_2 < 0$, then $\rho_i = 1 + \frac{1}{2}\lambda_i R_2 \varepsilon + o(\varepsilon)$;
- 2). $m \geq 3$: let $c_{2m-3} = -\sin 2\theta$ and $c_{2m-2} = \cos 2\theta$. If $\lambda_i R_m + H < 0$ and $\text{Re}(\vartheta_{n1} e^{-j\theta}) < 0$ with $\vartheta_{n1}, n = 3, \dots, m$ denoting the $m-2$ distinct roots of the equation

$$\vartheta_1^{m-2}(c_{2m-2} - c_{2m-3}j) + \dots + \vartheta_1(c_4 - c_3j) + (c_2 - c_1j) = 0, \quad (39)$$

then $\rho_i = 1 + \frac{1}{2} \max_{n=3, \dots, m} \{\lambda_i R_m + H, 2\text{Re}(\vartheta_{n1} e^{-j\theta})\} \varepsilon + o(\varepsilon)$.

Lemma 4.2: Assume that Lemma 4.1 holds. When ε is sufficiently small, for any vector $\xi \in \mathbb{R}^{2m}$, the entries of $A_i^s \xi$, which are denoted as $\xi_{s,2j-1}$ and $\xi_{s,2j}$, $j = 1, \dots, m$, satisfy that

$$|\xi_{s,2j-1}|, |\xi_{s,2j}| \leq \begin{cases} \|\xi\| M_{ij} \rho_i^s \varepsilon^{j-(m-1)}, & j = 1, 2, \dots, m-2; \\ \|\xi\| M_{ij} \rho_i^s \varepsilon^{(j-m)/2}, & j = m-1, m, \end{cases} \quad (40)$$

where

$$M_{ij} = \begin{cases} \frac{5}{2\lambda_i} \left(\sum_{n=3}^m \frac{|\vartheta_{n1}|^{j-1}}{\prod_{3 \leq k \leq m, k \neq n} |\vartheta_{k1} - \vartheta_{n1}|} \right), & j = 1, \dots, m-3; \\ \frac{5}{2\lambda_i} \left(\sum_{n=3}^m \frac{|\vartheta_{n1}|^{m-3}}{\prod_{3 \leq k \leq m, k \neq n} |\vartheta_{k1} - \vartheta_{n1}|} + 1 \right), & j = m-2; \\ \frac{3}{\sqrt{2\lambda_i}}, & j = m-1; \\ 5/2, & j = m. \end{cases}$$

Remark 4.3: The proofs of the above lemmas can be found in the Appendix. As in [15], the basic idea is to combine the bifurcation analysis of the roots of characteristic polynomials and the Jordan basis of a perturbed matrix [19]. However, the situation here is much different. On one hand, the complex conjugate eigenvalues of the original matrix A complicates the analysis of the perturbed eigenvalues, as seen from the proof of Lemma 4.1. On the other hand, unlike [15] where the unperturbed matrix admits multiple eigenvalues of 0 and 1, the unperturbed matrix here admits eigenvalues of m identical pairs of complex conjugate numbers, which allows a less cumbersome calculation of the perturbed Jordan basis, as in the proof of Lemma 4.2.

Remark 4.4: Assume $0 < \lambda_2 \leq \dots \leq \lambda_N$ and let $\rho = \max_{i=2, \dots, N} \rho_i$, $h \in (0, \lambda_2]$. Given $c_{2m-3} = -\sin 2\theta$ and $c_{2m-2} = \cos 2\theta$, the other constants c_{2j-1} and c_{2j} can be selected as follows such that $\rho \leq 1 - \varepsilon + o(\varepsilon) < 1 - \varepsilon/2$ holds with sufficiently small ε :

- 1). $m = 2$: select $c_3 = -(4/h + 1) \sin \theta$, $c_4 = (4/h + 1) \cos \theta$ such that $R_2 = -2/h < 0$;
- 2). $m \geq 3$: first select $c_{2m-4}, c_{2m-5}, \dots, c_1$ such that the solutions of (39) are given by $\vartheta_{n1} = -(n-2)e^{j\theta}$, $n = 3, \dots, m$ and H is determined by (38). In fact, direct computation shows that $c_{2m-4} - c_{2m-5}j = \frac{1}{2}(m-1)(m-2)e^{3j\theta}$ and consequently $H = (m-1)(m-2)/2 > 0$. Now let $c_{2m-1} = -[(2H+4)/h+1] \sin \theta$, $c_{2m} = [(2H+4)/h+1] \cos \theta$ such that

$$\lambda_i R_m + H = H(1 - \lambda_i/h) - 2\lambda_i/h < -2\lambda_i/h \leq -2.$$

With such a selection, $M_{ij} = \frac{5}{2\lambda_i} \left(\sum_{n=1}^{m-2} \frac{n^{j-1}}{\prod_{1 \leq k \leq m-2, k \neq n} |k-n|} \right)$, $j = 1, \dots, m-3$ and $M_{i,m-2} = \frac{5}{2\lambda_i} \left(\sum_{n=1}^{m-2} \frac{n^{m-3}}{\prod_{1 \leq k \leq m-2, k \neq n} |k-n|} + 1 \right) + 1$.

To explicitly express the data rate, another lemma is required.

Lemma 4.3: Denote

$$l(\theta) = [l_0(\theta), l_1(\theta), \dots, l_{2m-2}(\theta), l_{2m-1}(\theta)] = \cos \theta S(1, \cdot) + \sin \theta S(2, \cdot) + S(3, \cdot).$$

Then

$$l_k(\theta) = (-1)^{k-1} \sum_{h=0}^{\lfloor k/2 \rfloor} C(m, k-2h) C(m - (k-2h), h) (2 \cos \theta)^{k-2h}, \quad k = 0, 1, \dots, 2m-1.$$

Moreover, $\sum_{k=0}^{2m-1} |l_k(\theta)| = [2(1 + |\cos \theta|)]^m - 1$.

The proof of Lemma 4.3 can be found in the Appendix. The idea is simple and direct: by viewing $l(\theta)$ as the solution to an equation, we are left to the verification when $l(\theta)$ takes the form in the lemma. Still, the computation is quite complicated and requires special techniques from combinatorics [20], [21].

C. Convergence analysis and data rate

The notations in (10) will still be used, except that ψ_1 is replaced by ϕ_1 . The following assumptions are adopted in the subsequent analysis.

Assumption 4.1: The communication graph \mathcal{G} is undirected and connected.

Assumption 4.2: There exist known positive constants C^* and C_δ^* such that $\max_{j=1,\dots,2m} \|x_j(0)\| \leq C^*$ and $\max_{j=1,\dots,2m} \|\delta_j(0)\| \leq C_\delta^*$.

Remark 4.5: Assumption 4.1 is a standard assumption, under which the eigenvalues of L can be rearranged as $0 = \lambda_1 < \lambda_2 \leq \dots \leq \lambda_N$. The reason that we only consider the undirected graph will be clarified in Remark 4.7. Assumption 4.2 serves the same purpose as Assumption 3.2.

We also need the following constants:

$$\begin{aligned} b^* &= \max_{l,j} \{|\tilde{b}_{lj}|\}, \quad c^* = \max_{1 \leq j \leq 2m} \{c_j\}, \\ \Lambda_i &= \max\{\lambda_i^{1/2}, \lambda_i^{3/2}\}, \quad \bar{C} = \frac{9}{\sqrt{2}}[\|U_L^{-1}\| + 5c^*mN(\|S\| + 2)]. \end{aligned} \quad (41)$$

Lemma 4.4: Let $\gamma = 1 - \varepsilon/4$. Then we can choose sufficiently small ε to satisfy the following inequalities:

$$2c^* \sum_{j \neq m-1} M_{ij} \varepsilon^{1/2} \leq M_{i,m-1}, \quad i = 2, \dots, N; \quad (42a)$$

$$\sum_{k=0}^{2m-1} |l_k(\theta)|/\gamma^{2m} \leq [2(1 + |\cos \theta|)]^m - 1/2; \quad (42b)$$

$$(2m-1)b^*(N-1)\Lambda\bar{C}\varepsilon^{1/2} \leq \frac{1}{8}\gamma^{4m-1}, \quad (42c)$$

where $\Lambda = \max_i \Lambda_i$.

Theorem 4.1: Take k_j 's as in (37), c_j 's as in Remark 4.4 and $\gamma = 1 - \varepsilon/4$. Select sufficiently small ε to satisfy Lemma 4.4 and $\rho_i < 1 - \varepsilon/2$ for $i = 2, \dots, N$. Then under Assumptions 4.1 and 4.2, consensus can be achieved at a convergence rate of $O(\gamma^t)$ provided that $M(t)$ satisfies

$$\begin{cases} M(t) \geq 1, \quad t = 1, \dots, 2m; \\ M(t) \geq 2^{m-1}(1 + |\cos \theta|)^m - \frac{1}{2}, \quad t = 2m+1, \dots, \end{cases} \quad (43)$$

and $p_0 \geq (\sqrt{2} + 1)^{2m} \max\{C^*, C_\delta^*\}$.

Therefore, we can use $\lceil \log_2 2[2^{m-1}(1 + |\cos \theta|)^m - \frac{1}{2}] \rceil$ bits of information exchange to achieve the consensus.

Proof: 1) Preparation. By (36) we have

$$u(t) = \begin{cases} 0, & t = 0, 1, \dots, 2m-1; \\ -\sum_{j=1}^{2m} k_j L(\delta_j(t) + e_j(t)), & t \geq 2m. \end{cases} \quad (44)$$

Direct computation shows

$$\begin{bmatrix} \delta_1(t+1) \\ \vdots \\ \delta_{2m}(t+1) \end{bmatrix} = (A \otimes I_N) \begin{bmatrix} \delta_1(t) \\ \vdots \\ \delta_{2m}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ u(t) \end{bmatrix}$$

Let $\tilde{\delta}_j(t) = U_L^{-1} \delta_j(t) = [\tilde{\delta}_{1j}(t), \dots, \tilde{\delta}_{N,j}(t)]^T$ and $\tilde{\delta}^i(t) = [\tilde{\delta}_{i1}(t), \dots, \tilde{\delta}_{i,2m}(t)]^T$. Then we obtain $\tilde{\delta}^1(t) \equiv 0$, and for $i = 2, \dots, N$

$$\tilde{\delta}^i(t+1) = \begin{cases} A\tilde{\delta}^i(t), & t = 0, 1, \dots, 2m-1; \\ A_i\tilde{\delta}^i(t) - \epsilon_i(t), & t \geq 2m, \end{cases} \quad (45)$$

where $\epsilon_i(t) = [0, \dots, 0, \sum_{j=1}^{2m} k_j \lambda_i \phi_i^T e_j(t)]^T \in \mathbb{R}^{2m}$.

2) Estimation error and exponential convergence. To analyze the influence of $u(t)$ on the error term $e_j(t)$, we will show $|\phi_i^T u(t)| \leq \Lambda_i \bar{C} p_0 \gamma^{t-2m} \varepsilon^{1/2}$, $t \geq 2m$ by induction.

With the choice of p_0 and γ it's easy to see $|s_i(t)| \leq 3/2$ when $t \leq 2m$ by noticing $|y_i(t)| \leq (\sqrt{2} + 1)^{2m} C^*$, hence we obtain $\max_{1 \leq t \leq 2m} \|\Delta(t)\| \leq 1/2$ provided $M(t) \geq 1$, $t = 1, \dots, 2m$. Moreover, $\|\delta_j(2m)\| \leq (\sqrt{2} + 1)^{2m} C_\delta^*$. Recalling (36) and $e_j(2m) \leq p_0 \|S\| \max_{1 \leq t \leq 2m} \|\Delta(t)\|$ we have

$$\begin{aligned} |\phi_i^T u(2m)| &= |\phi_i^T \sum_{j=1}^{2m} k_j L(\delta_j(2m) + e_j(2m))| \\ &\leq 2\lambda_i m N c^* \varepsilon ((\sqrt{2} + 1)^{2m} C_\delta^* + p_0 \|S\|/2) \\ &\leq 2\lambda_i m N c^* \varepsilon (1 + \|S\|/2) p_0 \\ &\leq \Lambda_i \bar{C} p_0 \varepsilon^{1/2} \end{aligned}$$

by noticing $p_0 \geq (\sqrt{2} + 1)^{2m} C_\delta^*$ and $\varepsilon < 1$.

Assume that

$$\begin{aligned}
|\phi_i^T u(\tau)| &\leq \Lambda_i \bar{C} p_0 \gamma^{\tau-2m} \varepsilon^{1/2}, \quad 2m \leq \tau \leq t; \\
\|\Delta(\tau)\| &\leq \frac{1}{2} (\Rightarrow |e_{i1}(\tau)| \leq \frac{1}{2} p_0 \gamma^{\tau-1}), \quad 1 \leq \tau \leq t.
\end{aligned} \tag{46}$$

$$e_j(\tau) = \begin{bmatrix} e_{11}(\tau-2m+1) & \dots & e_{11}(\tau) \\ \vdots & \ddots & \vdots \\ e_{N1}(\tau-2m+1) & \dots & e_{N1}(\tau) \end{bmatrix} \begin{bmatrix} S(j,1) \\ \vdots \\ S(j,2m) \end{bmatrix} - \sum_{n=1}^{2m-1} \tilde{b}_{nj}(\theta) u(\tau-n),$$

we have

$$|\phi_i^T e_j(\tau)| \leq N p_0 \max_{\tau-2m+1 \leq s \leq \tau} \|\Delta(s)\| \gamma^{\tau-2m} \|S\| + N(2m-1) b^* \lambda_i \varepsilon^{1/2} \bar{C} p_0 \gamma^{\tau-4m+1}$$

and

$$\begin{aligned}
\|\epsilon_i(\tau)\| &\leq \lambda_i c^* \varepsilon 2mN [p_0 \max_{\tau-2m+1 \leq s \leq \tau} \|\Delta(s)\| \gamma^{\tau-2m} \|S\| + (2m-1) b^* \lambda_i \varepsilon^{1/2} \bar{C} p_0 \gamma^{\tau-4m+1}] \\
&\leq 2\lambda_i c^* mN (\|S\|/2 + 1) p_0 \gamma^{\tau-2m} \varepsilon
\end{aligned} \tag{47}$$

by (42c), if $\|\Delta(s)\| \leq 1/2$ for $\tau-2m+1 \leq s \leq \tau$. Recalling (45) we obtain

$$\tilde{\delta}^i(t+1) = A_i^{t+1-2m} \tilde{\delta}^i(2m) - \sum_{\tau=0}^{t-2m} A_i^{t-2m-\tau} \epsilon_i(2m+\tau).$$

By applying Lemma 4.2 and taking into account (47), it yields that

$$\begin{aligned}
&|\tilde{\delta}_{i,2j-1}(t+1)|, |\tilde{\delta}_{i,2j}(t+1)| \\
&\leq \begin{cases} M_{ij} \varepsilon^{j-(m-1)} \gamma^{t+1-2m} [\|\tilde{\delta}^i(2m)\| + 4\lambda_i c^* mN (\|S\| + 2) p_0], \\ j = 1, \dots, m-2; \\ M_{ij} \varepsilon^{(j-m)/2} \gamma^{t+1-2m} [\|\tilde{\delta}^i(2m)\| + 4\lambda_i c^* mN (\|S\| + 2) p_0], \\ j = m-1, m, \end{cases}
\end{aligned} \tag{48}$$

due to $\varepsilon/(\gamma - \rho_i) < 4$.

With (48) it is ready to estimate $\phi_i^T u(t+1)$, which is a sum of $\sum_{j=1}^{2m} k_j \phi_i^T L \delta_j(t+1)$ and $\sum_{j=1}^{2m} k_j \phi_i^T e_j(t+1)$. For the first part, by (48) and $\|\tilde{\delta}^i(2m)\| \leq \|U_L^{-1}\| \|\delta_i(2m)\|$ we have

$$\begin{aligned}
&|\sum_{j=1}^{2m} k_j \phi_i^T L \delta_j(t+1)| \\
&= \lambda_i |\sum_{j=1}^{2m} k_j \tilde{\delta}_{ij}(t+1)| \\
&\leq \lambda_i p_0 \gamma^{t-2m+1} (2M_{i,m-1} \varepsilon^{1/2} + 2c^* \varepsilon \sum_{j \neq m-1} M_{ij}) \cdot \\
&\quad [\|\tilde{\delta}^i(2m)\| + 4\lambda_i c^* mN (\|S\| + 2) p_0] \\
&\leq 3\lambda_i M_{i,m-1} [\|U_L^{-1}\| + 4\lambda_i c^* mN (\|S\| + 2)] p_0 \gamma^{t-2m+1} \varepsilon^{1/2}
\end{aligned} \tag{49}$$

if we note that $|c_{2m-2}|, |c_{2m-3}| \leq 1$. For the second part, as in the second order case, it is closely related with $\|\Delta(t+1)\|$ and similarly it can be inferred from (34) that $d(t+1)$ is only dependent on the past quantization errors $\Delta(\tau)$, $t-2m+1 \leq \tau \leq t$ and the past control inputs $u(\tau)$, $t-2m+1 \leq \tau \leq t-1$. Hence with the induction assumption (46) the quantizer can be made unsaturated at time $t+1$ with finite bits, namely $\|\Delta(t+1)\| \leq 1/2$. In consequence we get an estimation similar to (47) that

$$|\sum_{j=1}^{2m} k_j \phi_i^T L e_j(t+1)| \leq 2\lambda_i c^* mN (\|S\|/2 + 1) p_0 \gamma^{t-2m+1} \varepsilon. \tag{50}$$

Combining (49) and (50), it is clear that $|\phi_i^T u(t+1)| \leq \Lambda_i \bar{C} p_0 \gamma^{t-2m+1} \varepsilon^{1/2}$, which establishes the induction. Furthermore, by (48) clearly the consensus can be achieved at a convergence rate of $O(\gamma^t)$.

3) Data rate. Below we are to discuss the number of quantization levels at each time step. The situation when $t \leq 2m$ has been discussed. When $t > 2m$, we have

$$\begin{aligned}
&\|d(t)\| \\
&\leq \left\| \frac{\cos \theta}{\gamma} \Delta(t-1) + \sum_{j=1}^{2m} \frac{\sin \theta S(2,j) + S(3,j)}{\gamma^{2m-j+1}} \Delta(t-1-2m+j) \right\| + 2b^* \sum_{j=1}^{2m-1} \frac{\|u(t-1-j)\|}{p(t-1)} \\
&= \left\| \sum_{j=1}^{2m} \frac{1}{\gamma^{2m-j+1}} [\cos \theta S(1,j) + \sin \theta S(2,j) + S(3,j)] \Delta(t-1-2m+j) \right\| \\
&\quad + 2b^* \sum_{j=1}^{2m-1} \frac{\|u(t-1-j)\|}{p(t-1)} \\
&\leq \frac{1}{2\gamma^{2m}} \sum_{k=0}^{2m-1} |l_k(\theta)| + 2b^* (2m-1)(N-1) \Lambda \bar{C} \gamma^{1-4m} \varepsilon^{1/2}
\end{aligned}$$

by noticing $S(1, \cdot) = [0, \dots, 0, 1]$ and $u(t) = \sum_{i=1}^N \phi_i \phi_i^T u(t)$, $\Lambda_1 = 0$. By taking into account (42b) and (42c) it can be seen that $\|d(t)\|$ is bounded by $2^{m-1}(1 + |\cos \theta|)^m$ and the proof is completed by remembering (43). ■

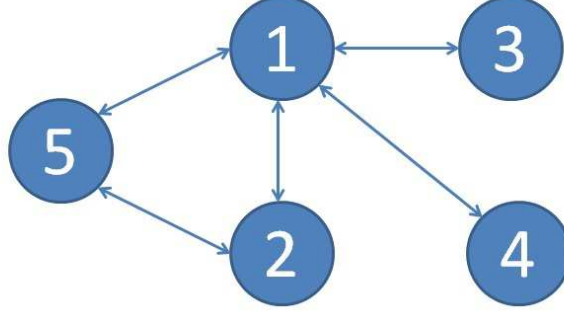


Fig. 1. Communication topology

Remark 4.6: Noticing that $\prod_{1 \leq k \leq m-2, k \neq n} |k-n| = (n-1)!(m-2-n)!$ attains the minimum at $n = \lfloor \frac{m-2}{2} \rfloor$ and multiplying by a positive λ_i on both sides does not change the direction of an inequality, (42a) can be substituted by the following stronger one, which is easier to check:

$$5c^* \left(\sum_{j=1}^{m-2} \sum_{n=1}^{m-2} \frac{n^{j-1}}{(\lfloor \frac{m}{2} \rfloor - 2)!(m-1-\lfloor \frac{m}{2} \rfloor)!} + 1 + \lambda_N \right) \varepsilon^{1/2} \leq 3\sqrt{\frac{\lambda_2}{2}}. \quad (51)$$

Remark 4.7: From the proof it is readily seen that we can still use the same number of bits to achieve the quantized consensus once the Laplacian of the directed topology satisfies that $0 < \lambda_2 \leq \dots \leq \lambda_N$. However, unlike the case of the 2nd-order oscillator, it does not hold for the general topology, when the Laplacian contains complex eigenvalues, or real Jordan blocks of multiple dimensions. For one reason, note that Lemma 4.1 does not hold for a complex λ_i . For another one, note the disparity in the order of ε between the disturbance term and the weighted sum of disagreement entries, i.e. $\|\epsilon_i(t)\| \sim O(\varepsilon)p_0\gamma^t$ and $\|K\tilde{\delta}^i(t)\| \sim O(\varepsilon^{1/2})p_0\gamma^t$. Therefore, if we assume $m = 2$ and the Jordan block corresponding to $\lambda_2 > 0$ is two-dimensional as in (18), then it follows from $\|K\tilde{\delta}^3(t)\| \sim O(\varepsilon^{1/2})p_0\gamma^t$ that $\|\tilde{\delta}^2(t)\| \sim O(\varepsilon^{-1})p_0\gamma^t$ and $\|u(t)\| \sim O(1)p_0\gamma^t$, suggesting that the input term can no longer be neglected in the estimation errors, nor in the quantization input $d(t)$. Such a situation is also encountered in [15].

Remark 4.8: At the first glance it may seem doubtful that the data rate is dependent on $|\cos \theta|$; but a little further inspection is enough to clarify. Similar to the situation of the n -th order integrator system investigated in [15], the control input does not consume any bit in exchanging the estimates of the states when ε is sufficiently small. In other words, we only need to focus on how many bits it needs to estimate the output of an individual open loop system. Take the second-order case as an example. Noticing that $y_i(t) = \cos \theta x_{i1}(t-1) + \sin \theta x_{i2}(t-1) = 2 \cos \theta x_{i1}(t-1) - x_{i1}(t-2)$, we can estimate $y_i(t)$ based on $\hat{x}_{i1}(t-1)$ and $\hat{x}_{i1}(t-2)$ with an error bound no larger than $\frac{1}{2}(2|\cos \theta| + 1) + \frac{1}{2}$. Generally speaking, when $|\cos \theta| \approx 0$ or equivalently $|\sin \theta| \approx 1$, $x_{i,2j-1}(t)$ and $x_{i,2j}(t)$ are tightly coupled, and it needs only m bits of information exchange to achieve the consensus; in the case of $|\cos \theta| \approx 1$, after rearranging of states A can be approximated by $I_2 \otimes J_{1,m}$, and $2m$ bits are sufficient. Anyway, for a $2m$ -th order system studied in this paper, $2m$ bits are enough to realize the consensus asymptotically, which is consistent with the conclusion for n -th order integrator systems [15].

V. NUMERICAL EXAMPLE

For simplicity we only show an example of $m = 2$. Consider a 5-node network with 4-th order dynamics, where the edges are generated randomly according to probability $P(i, j) \in \mathcal{E} = 0.5$ with 0-1 weights. The initial states are randomly chosen as $x_{ij}(0) \in (0, j)$, $i = 1, \dots, 5$, $j = 1, \dots, 4$. Given $\theta = \pi/3$, it is enough to use 3 bits of information exchange to realize the consensus, and we can compute $S_2 = \begin{bmatrix} -4/(3\sqrt{3}) & 2/\sqrt{3} & -4/(3\sqrt{3}) & 5/(3\sqrt{3}) \\ -1/3 & 1 & -1 & 2/3 \\ -1/\sqrt{3} & 1/\sqrt{3} & -1/\sqrt{3} & 0 \end{bmatrix}$ to construct the encoder and decoder respectively as (30)-(33). The communication topology is generated as in Figure 1 with $\lambda_2 = 0.8299$, and c_j 's are determined as in Remark 4.4 by choosing $h = \lambda_2$. Moreover, let $\varepsilon = 0.01$, $p_0 = 10$, $\gamma = 0.9975$ to satisfy the conditions in Theorem 4.1. From Figure 2 which depicts the trajectory of $\delta_{jmax}(t) = \{\delta_{nj}(t) : n = \arg \max_i |\delta_{ij}(t)|\}$, we can see that the consensus is achieved asymptotically.

VI. CONCLUDING REMARKS

In this paper, we explored the data rate problem for quantized consensus of a special kind of multi-agent systems. The dynamics of each agent is described by a $2m$ -th order real Jordan form consisting of m pairs of conjugate poles on the unit circle with single input, and only the first state can be measured. The encoding-decoding scheme was based on the observability matrix. Perturbation techniques were employed in the consensus analysis and the data rate analysis, and combinatorial techniques

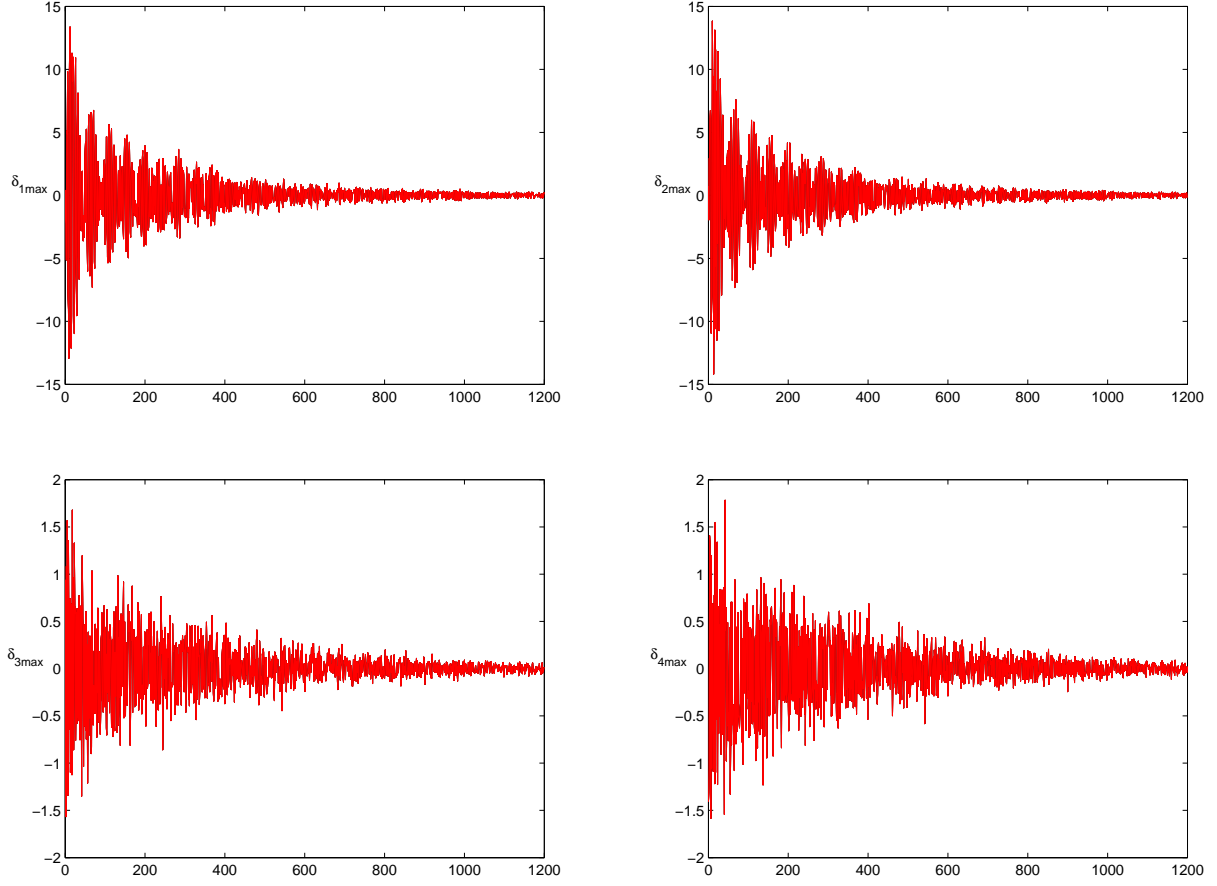


Fig. 2. Trajectories of disagreements $\delta_{jmax}(t)$

were used to explicitly obtain the data rate. The second-order case of $m = 1$ and higher-order cases of $m \geq 2$ were investigated separately. For the second-order case, we showed that at most 2 bits of information exchange suffice to achieve the consensus at an exponential rate, if the communication topology has a spanning tree. For the higher-order cases, consensus was achieved with at most $2m$ bits, provided that the undirected communication topology is connected. The exact number of bits for achieving consensus in both cases is an integer which increases from m to $2m$ when $|\cos \theta|$ increases from 0 to 1. The case of switching directed topology is still under investigation, and noisy communication channels will be considered in the future work. As for general unstable systems with poles outside the unit circle, perturbation techniques no longer apply and new methods need to be developed to serve the same purpose of stabilizing the dynamics of disagreements.

APPENDIX

Proof of Lemma 4.1 Here we mainly deal with the case of $m \geq 3$, since the proof can be slightly adapted if $m = 2$ and the modification will be pointed out accordingly. The characteristic equation of A_i can be computed as

$$\chi_i(\mu) = \det[(\mu I - Q)^m + \lambda_i K_m (\mu I - Q)^{m-1} + \cdots + \lambda_i K_2 (\mu I - Q) + \lambda_i K_1],$$

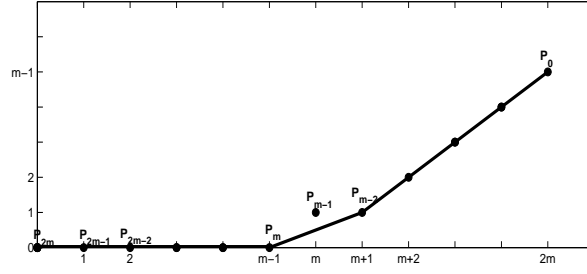


Fig. 3. Newton diagram

where $Q = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$ and $K_j = \begin{bmatrix} 0 & 0 \\ k_{2j-1} & k_{2j} \end{bmatrix}$ for $j = 1, \dots, m$. By employing (11) in the proof of Lemma 3.1, we rewrite $\chi_i(\mu)$ as

$$\begin{aligned} \chi_i(\mu) &= (\mu - e^{j\theta})^m (\mu - e^{-j\theta})^m + \frac{1}{2} [(\mu - e^{j\theta})^m + (\mu - e^{-j\theta})^m] \frac{\lambda_i}{2} \\ &\quad \left[\sum_{j=2}^m (-k_{2j-1}j + k_{2j})(\mu - e^{j\theta})^{j-1} + 2k_2 + \sum_{j=2}^m (k_{2j-1}j + k_{2j})(\mu - e^{-j\theta})^{j-1} \right] \\ &\quad + \frac{1}{2} [(\mu - e^{j\theta})^m - (\mu - e^{-j\theta})^m] \frac{\lambda_i}{2} \\ &\quad \left[\sum_{j=2}^m (k_{2j-1} + k_{2j}j)(\mu - e^{j\theta})^{j-1} + 2k_1 + \sum_{j=2}^m (k_{2j-1} - k_{2j}j)(\mu - e^{-j\theta})^{j-1} \right] \\ &= (\mu - e^{j\theta})^m [(\mu - e^{-j\theta})^m + \frac{\lambda_i}{2}(k_2 + k_{1j})] + \frac{\lambda_i}{2} (\mu - e^{-j\theta})^m \sum_{j=1}^m (\mu - e^{j\theta})^{j-1} (-k_{2j-1}j + k_{2j}). \end{aligned} \quad (52)$$

With A_i being real, we only need to focus on the perturbed roots around $e^{j\theta}$, which are denoted by $\mu = e^{j\theta} + \Delta\mu$. Noticing that $\mu - e^{-j\theta} = \mu - e^{j\theta} + 2\sin \theta j$, we substitute $\mu = e^{j\theta} + \Delta\mu$ into (52) and obtain

$$\chi_i(e^{j\theta} + \Delta\mu) = \sum_{n=1}^m a_{in}(\varepsilon)(\Delta\mu)^{m-n} + \sum_{n=0}^m (C(m, m-n)(2\sin \theta j)^{m-n} + O(\varepsilon))(\Delta\mu)^{m+n} \quad (53)$$

with the selection of k_{2j-1} and k_j in (37), where

$$a_{in}(\varepsilon) = \begin{cases} (2\sin \theta j)^{m-1} [2\sin \theta j(c_{2m} - c_{2m-1}j) + m(c_{2m-2} - c_{2m-3}j)]\varepsilon + o(\varepsilon), & n = 1; \\ (2\sin \theta j)^m (c_{2(m-n+1)} - c_{2(m-n)+1}j)\varepsilon^{n-1} + o(\varepsilon^{n-1}), & n = 2, \dots, m. \end{cases}$$

Now the Newton diagram [17] can be depicted as in Fig. 3, by first plotting points $P_{2m-j}(j, \alpha_{2m-j})$, $j = 0, \dots, 2m$ and then connecting the segments on the lower boundary of the convex hull of the above points, where α_{2m-j} is the leading exponent of ε in the coefficient of $(\Delta\mu)^{2m-j}$. The slopes of the two non-horizontal segments are $1/2$, 1 respectively, implying that $\Delta\mu$ has the following two forms of expansions:

$$\Delta\mu = \mu_1 \varepsilon^{\frac{1}{2}} + \mu_2 \varepsilon^\beta + o(\varepsilon^\beta), \quad \mu_1 \neq 0; \quad (54a)$$

$$\Delta\mu = v_1 \varepsilon + o(\varepsilon), \quad v_1 \neq 0. \quad (54b)$$

Substituting (54a) into (53) and finding the coefficients of the term $\varepsilon^{m/2}$, it yields that $\mu_1^m (2j \sin \theta)^m + \mu_1^{m-2} \frac{\lambda_i}{2} (2j \sin \theta)^m (c_{2m-2} - c_{2m-3}j) = 0$, and thus

$$\mu_1 = \pm \sqrt{\frac{\lambda_i}{2} \sqrt{c_{2m-3}^2 + c_{2m-2}^2} e^{j\frac{\alpha}{2}}}, \quad (55)$$

where $\alpha = \arg(c_{2m-3}j - c_{2m-2})$. Moreover, to determine μ_2 and β , we substitute (54a) into (53) again and find the lowest order term as

$$\begin{aligned} &m\mu_1^{m-1} \mu_2 (2j \sin \theta)^m \varepsilon^{(m-1)/2} \varepsilon^\beta \\ &+ \frac{\lambda_i}{2} \mu_1^{m-1} (2j \sin \theta)^m (c_{2m} - c_{2m-1}j) \varepsilon^{(m+1)/2} \\ &+ \frac{\lambda_i}{2} (m-2) \mu_1^{m-3} \mu_2 (2j \sin \theta)^m (c_{2m-2} - c_{2m-3}j) \varepsilon^{(m-1)/2} \varepsilon^\beta \\ &+ \frac{\lambda_i}{2} \mu_1^{m-3} (2j \sin \theta)^m (c_{2m-4} - c_{2m-5}j) \varepsilon^{(m+1)/2} = 0, \end{aligned} \quad (56)$$

which implies $\beta = 1$ and $\mu_2 = \frac{\lambda_i}{4} (c_{2m-1}j - c_{2m}) + \frac{(c_{2m-5}j - c_{2m-4})}{2(c_{2m-3}j - c_{2m-2})}$. In the form of (54a), the module of μ is determined as

$$|\mu|^2 = \mu \bar{\mu} = 1 + 2\operatorname{Re}(\mu_1 e^{-j\theta}) \varepsilon^{\frac{1}{2}} + (|\mu_1|^2 + 2\operatorname{Re}(\mu_2 e^{-j\theta})) \varepsilon + o(\varepsilon), \quad (57)$$

with $\text{Re}(\mu_1 e^{-j\theta}) = \pm \sqrt{\frac{\lambda_i}{2} \sqrt{c_{2m-3}^2 + c_{2m-2}^2} \cos(\frac{\alpha}{2} - \theta)}$. In order that $|\mu| < 1$ with sufficiently small ε , we must have $\frac{\alpha}{2} - \theta = \frac{\pi}{2}$, and hence it suffices to let $|\mu_1|^2 + 2\text{Re}(\mu_2 e^{-j\theta}) < 0$. Combining these arguments gives rise to a sufficient condition as

$$-c_{2m-3}/c_{2m-2} = \tan(2\theta + \pi), \quad c_{2m-3}^2 + c_{2m-2}^2 \neq 0; \quad (58a)$$

$$\frac{\lambda_i}{2} \sqrt{c_{2m-3}^2 + c_{2m-2}^2} + \frac{\lambda_i}{2} (c_{2m-1} \sin \theta - c_{2m} \cos \theta) + \text{Re}[\frac{c_{2m-5}j - c_{2m-4}}{c_{2m-3}j - c_{2m-2}} e^{-j\theta}] < 0. \quad (58b)$$

With $c_{2m-3} = -\sin 2\theta$ and $c_{2m-2} = \cos 2\theta$ satisfying (58a), (58b) is equivalent to $\lambda_i R_m + H < 0$. When $m = 2$, μ only takes the form of (54a) and $\mu_2 = \frac{\lambda_i}{4} (c_{2m-1}j - c_{2m})$, leading to the sufficient condition $R_2 < 0$ for $|\mu| < 1$.

On the other hand, substituting (54b) into (52) and finding the coefficients of the term ε^{m-1} , we obtain the equation (39). Similarly, the module of μ with the form (54b) is determined by

$$|\mu|^2 = 1 + 2\text{Re}(\vartheta_1 e^{-j\theta})\varepsilon + o(\varepsilon) \quad (59)$$

and it suffices to let $\text{Re}(\vartheta_1 e^{-j\theta})$ to be negative such that $|\mu| < 1$ with sufficiently small ε . For prescribed c_{2m-3} and c_{2m-2} , the roots of (39) can be assigned arbitrarily such that $\text{Re}(\vartheta_1 e^{-j\theta}) < 0$ with $m-2$ distinct ϑ_1 ; after determining c_{2m-4} and c_{2m-5} , (58b) can always be satisfied by properly chosen c_{2m} and c_{2m-1} since $\frac{\lambda_i}{2} (c_{2m-1} \sin \theta - c_{2m} \cos \theta)$ can be assigned to any number. In summary, the proof is completed. ■

Proof of Lemma 4.2 As in the last proof, we only focus on the case of $m \geq 3$ which essentially includes the case of $m = 2$. For A_i , we are to find the following Jordan decomposition:

$$A_i = A + \sum_{j=1}^{m-1} A_{ij} \varepsilon^j = R_i \tilde{A}_i R_i^{-1}, \quad (60)$$

where \tilde{A}_i is a diagonal matrix consisting of $2m$ different eigenvalues determined in Lemma 4.1. To find an appropriate R_i and the corresponding R_i^{-1} , we first determine the Jordan basis of the unperturbed matrix A . The Jordan chain corresponding to the eigenvalue $\mu_0 = e^{j\theta}$ is given by

$$u_{m-1} \xrightarrow{A - \mu_0 I} u_{m-2} \xrightarrow{A - \mu_0 I} \dots \xrightarrow{A - \mu_0 I} u_1 \xrightarrow{A - \mu_0 I} u_0,$$

where $u_j = e_{2j+1} + j e_{2j+2}$, $j = 0, \dots, m-1$ and $e_n \in \mathbb{R}^{2m}$ denotes the vector with a 1 in the n -th coordinate and 0's elsewhere. Similarly, the Jordan chain corresponding to the eigenvalue $\bar{\mu}_0 = e^{-j\theta}$ is given by

$$\bar{u}_{m-1} \xrightarrow{A - \bar{\mu}_0 I} \bar{u}_{m-2} \xrightarrow{A - \bar{\mu}_0 I} \dots \xrightarrow{A - \bar{\mu}_0 I} \bar{u}_1 \xrightarrow{A - \bar{\mu}_0 I} \bar{u}_0.$$

Hence the two Jordan chains of A can be rearranged as $R_0 = (u_0 \ \bar{u}_0 \ \dots \ u_{m-1} \ \bar{u}_{m-1}) = I_m \otimes P$ with $P = \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix}$. With A_i being real, once we obtain the eigenvectors corresponding to the m different perturbed eigenvalues around μ_0 , the other eigenvectors can be obtained by taking conjugates. Hence we only need to find the eigenvectors corresponding to the m different perturbed eigenvalues around μ_0 .

The eigenvectors corresponding to the m perturbed eigenvalues around μ_0 have the following form of Puiseux series [19]:

$$\begin{aligned} \mu_{in} &= \mu_0 + \sum_{k=1}^{\infty} \mu_{ink} \varepsilon^{k/2}, \quad u_{in} = u_{in0} + \sum_{k=1}^{\infty} u_{ink} \varepsilon^{k/2}, \quad n = 1, 2; \\ \mu_{in} &= \mu_0 + \sum_{k=1}^{\infty} \vartheta_{ink} \varepsilon^k, \quad u_{in} = u_{in0} + \sum_{k=1}^{\infty} u_{ink} \varepsilon^k, \quad n = 3, \dots, m, \end{aligned}$$

where $\mu_{i11} = \sqrt{\frac{\lambda_i}{2} \sqrt{c_{2m-3}^2 + c_{2m-2}^2} e^{j\frac{\alpha}{2}}}$, $\mu_{i21} = -\mu_{i11}$ and $\vartheta_{in1} = \vartheta_{n1}$, $n = 3, \dots, m$ have been defined in Lemma 4.1. Substituting μ_{in}, u_{in} into the equation $A_i u_{in} = \mu_{in} u_{in}$ respectively, and collecting coefficients of equal powers of ε ; moreover, noticing the fact that $A_{ij} u_k = 0$, $j = 1, \dots, m-2$; $k = 0, \dots, m-2-j$, where A_{ij} has been defined in (60) and imposing the normalization condition as $v_{m-1}^T u_{in} = 1$, where $v_{m-1}^T = \frac{1}{2} \begin{bmatrix} 1 & -j & 0 & \dots & 0 \end{bmatrix}$ is the left associated eigenvector of A with respect to the eigenvalue μ_0 such that $v_{m-1}^T u_0 = 1, v_{m-1}^T u_1 = \dots = v_{m-1}^T u_{m-1} = 0$, the m eigenvectors can be obtained as:

$$\begin{aligned} u_{in} &= u_0 + \sum_{k=1}^{m-1} \varepsilon^{k/2} (\mu_{i1n}^k u_k + u'_{ikn}) + o(\varepsilon^{m/2}), \quad n = 1, 2; \\ u_{in} &= u_0 + \sum_{k=1}^{m-1} \varepsilon^k (\vartheta_{1n}^k u_k + u'_{ikn}) + o(\varepsilon^m), \quad n = 3, \dots, m, \end{aligned}$$

where $u'_{i1n} = 0$ and $u'_{ikn} \in \text{span}\{u_1, \dots, u_{k-1}\}$, $k = 2, \dots, m-1$ for $n = 1, \dots, m$.

Letting $R_i = \begin{bmatrix} u_{i1} & \bar{u}_{i1} & \dots & u_{im} & \bar{u}_{im} \end{bmatrix}$, we are to investigate the magnitude of each entry in R_i^{-1} by adjoint method. Therefore we need to find the order of $\det R_i$ and the corresponding cofactor, both of which can be expressed as Puiseux series. The following facts should be mentioned before the calculation:

1). Determinant is a multi-linear function of column vectors, and it vanishes when two or more columns coincide.

2). There exist two types of series in the columns of R_i , and we categorize u_{i1}, u_{i2} and their conjugates for type I, the others for type II.

With these facts, we can see that the lowest degree can be obtained by taking out terms with $\varepsilon^{(m-2)/2}u_{m-2}$ and $\varepsilon^{(m-1)/2}u_{m-1}$ respectively from u_{i1} and u_{i1} , terms with $u_0, \varepsilon u_1, \dots, \varepsilon^{m-3}u_{m-3}$ respectively from u_{i3}, \dots, u_{im} , as well as the corresponding conjugates from $\bar{u}_{i1}, \dots, \bar{u}_{im}$, and calculated by $2(0+1+\dots+m-3+\frac{m-2}{2}+\frac{m-1}{2})=m^2-3m+3$. Moreover,

$$\begin{aligned} |\det R_i| &= |\mu_{i11}^{m-1}\mu_{i21}^{m-2}-\mu_{i11}^{m-2}\mu_{i21}^{m-1}|^2|\det V_0|^2|\det R_0|\varepsilon^{m^2-3m+3}(1+o(1)) \\ &= |\mu_{i11}^{m-2}\mu_{i21}^{m-2}(\mu_{i11}-\mu_{i21})|^2|\det V_0|^22^m\varepsilon^{m^2-3m+3}(1+o(1)), \end{aligned} \quad (61)$$

where $V_0 = V(\vartheta_{31}, \dots, \vartheta_{m1}) = \begin{bmatrix} 1 & \vartheta_{31} & \dots & \vartheta_{31}^{m-3} \\ 1 & \vartheta_{41} & \dots & \vartheta_{41}^{m-3} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \vartheta_{m1} & \dots & \vartheta_{m1}^{m-3} \end{bmatrix}$ is a Vandermonde matrix of order $m-2$.

On the other hand, we need to determinate the order of the cofactor $C_{s,t}^{(i)}$ of the (s, t) entry, and we illustrate it by calculating $C_{1,1}^{(i)}$ with $m=3$. After deleting the first column u_{i1} , we delete the first row and use the same notations u_1, u_2 and e_2, \dots, e_6 . Now R_i has been reduced to a square matrix $R_{1,1}^{(i)}$ consisting of the following 5 columns:

$$\begin{aligned} a_1 &= -je_2 + \varepsilon^{1/2}\bar{\mu}_{i11}\bar{u}_1 + \varepsilon\bar{\mu}_{i11}^2\bar{u}_2 + O(\varepsilon)\bar{u}_1 + O(\varepsilon^2), \\ a_2 &= je_2 + \varepsilon^{1/2}\mu_{i21}u_1 + \varepsilon\mu_{i21}^2u_2 + O(\varepsilon)u_1 + O(\varepsilon^2), \quad a_3 = \bar{a}_2, \\ a_4 &= je_2 + \varepsilon\vartheta_{31}u_1 + \varepsilon^2\vartheta_{31}^2u_2 + O(\varepsilon^2)u_1 + O(\varepsilon^2), \quad a_5 = \bar{a}_4. \end{aligned}$$

Consequently the order of $C_{1,1}^{(i)}$ is found in such a way: take out terms with $\varepsilon^{1/2}\bar{u}_1, \varepsilon\bar{u}_2$ respectively from a_1, a_3 , terms with e_2 from a_5 , terms with εu_2 from a_2 , terms with εu_1 from a_4 . Now that a_1, a_3, a_5 jointly contribute the same degree of $\frac{1}{2}(m^2-3m+3)$ as $\bar{u}_{i1}, \bar{u}_{i2}, \bar{u}_{i3}$, we are left to choose terms with u_1 and u_2 respectively from a_2 and a_4 . The above can be conducted similarly for calculating the order of $C_{1,1}^{(i)}$ when $m > 3$, and actually for every cofactor. Moreover, by the symmetry of conjugates, $C_{2k-1,2n-1}^{(i)}, C_{2k-1,2n}^{(i)}, C_{2k,2n-1}^{(i)}, C_{2k,2n}^{(i)}$ have an identical order. So we only focus on $C_{2k-1,2n-1}^{(i)}$ below. Reminded by the case of $C_{1,1}^{(i)}$ when $m=3$, we suffice to choose linearly independent terms with a lowest sum of degrees from the modified columns u_{ij} for $j \neq n$, where u_{k-1} has been subtracted from each column. Recall that in finding the order of $|\det R_i|$, terms with $\varepsilon u_0, \dots, \varepsilon^{m-3}u_{m-3}$ from type II columns are first selected, and then terms with $\varepsilon^{(m-2)/2}u_{m-2}, \varepsilon^{(m-1)/2}u_{m-1}$ from type I columns. Such a method still applies in finding the order of cofactors, and we conclude that $C_{2k-1,2n-1}^{(i)}$ has the lowest order for fixed n if and only if $k=m$. In other words, for any row in $\text{adj}R_i$, the entries at the $2m-1$ -th and $2m$ -th column exclusively have the lowest order when compared with other entries at the same row. To be detailed,

$$\begin{aligned} &|C_{2m-1,2n-1}^{(i)}|, |C_{2m-1,2n}^{(i)}|, |C_{2m,2n-1}^{(i)}|, |C_{2m,2n}^{(i)}| \\ &= \begin{cases} 2^{m-1}|\det V_0|^2|\mu_{i11}^{m-2}\mu_{i21}^{m-2}(\mu_{i11}-\mu_{i21})||\mu_{i11}^{m-2}| \\ \cdot \varepsilon^{m^2-3m+3-(m-1)/2}(1+o(1)), & n=1, 2; \\ 2^{m-1}|\det V_0||\det V_n||\mu_{i11}^{m-2}\mu_{i21}^{m-2}(\mu_{i11}-\mu_{i21})| \\ \cdot |\mu_{i11}^{m-3}\mu_{i21}^{m-3}(\mu_{i11}-\mu_{i21})|\varepsilon^{m^2-3m+3-(m-2)}(1+o(1)), & n=3, \dots, m, \end{cases} \end{aligned} \quad (62)$$

where $V_n = V(\vartheta_{31}, \dots, \vartheta_{n-1,1}, \vartheta_{n1}, \dots, \vartheta_{m1})$ is a Vandermonde matrix of order $m-3$. Together with (61) it yields that by $\mu_{i11} = -\mu_{i21}$

$$\begin{aligned} &|R_i^{-1}(2n-1, 2m-1)|, |R_i^{-1}(2n, 2m-1)|, |R_i^{-1}(2n-1, 2m)|, |R_i^{-1}(2n, 2m)| \\ &= \begin{cases} \frac{1}{4|\mu_{i11}|^{m-1}}\varepsilon^{-(m-1)/2}(1+o(1)), & n=1, 2; \\ \frac{1}{2|\mu_{i11}|^2}|\det V_n||\det V_0|^{-1}\varepsilon^{-(m-1)}(1+o(1)), & n=3, \dots, m; \end{cases} \\ &|R_i^{-1}(2n-1, k)|, |R_i^{-1}(2n, k)| \\ &= \begin{cases} o(\varepsilon^{-(m-1)/2}), & n=1, 2; \\ o(\varepsilon^{-(m-1)}), & n=3, \dots, m; \end{cases} \quad \text{for } 1 \leq k \leq 2m. \end{aligned} \quad (63)$$

In the meanwhile, the following holds for $1 \leq j \leq m$:

$$\begin{aligned} &|R_i(2j-1, 2n-1)|, |R_i(2j-1, 2n)|, |R_i(2j, 2n-1)|, |R_i(2j, 2n)| \\ &= \begin{cases} |\mu_{i11}^{j-1}|\varepsilon^{(j-1)/2} + O(\varepsilon^{j/2}), & n=1, 2; \\ |\vartheta_{n1}^{j-1}|\varepsilon^{j-1} + O(\varepsilon^j), & n=3, \dots, m; \end{cases} \end{aligned} \quad (64)$$

Combining (60), (63) and (64) we can obtain

$$\begin{aligned} & |\xi_{s,2j-1}| \\ & \leq \rho_i^s \|\xi\| \left[\sum_{n=1}^m |R_i(2j-1, 2n-1)| (|R_i^{-1}(2n-1, 2m-1)| + |R_i^{-1}(2n-1, 2m)|) \right. \\ & \quad \left. + \sum_{n=1}^m |R_i(2j-1, 2n)| (|R_i^{-1}(2n, 2m-1)| + |R_i^{-1}(2n, 2m)|) \right] (1 + o(1)) \\ & \leq \rho_i^s \|\xi\| \left[2 \left[\sum_{n=1}^2 \frac{|\mu_{i11}^{j-1}|}{4|\mu_{i11}^{m-1}|} \varepsilon^{(j-m)/2} + \sum_{n=3}^m \frac{|\vartheta_{n1}^{j-1}|}{2|\mu_{i11}^{m-1}|} \frac{|\det V_n|}{|\det V_0|} \varepsilon^{j-(m-1)} \right] (1 + o(1)) \right]. \end{aligned}$$

and the conclusion follows by noticing that $j - (m-1) < \frac{j-m}{2}$ for $j < m-2$, $j - (m-1) = \frac{j-m}{2}$ for $j = m-2$ and $j - (m-1) > \frac{j-m}{2}$ for $j = m-1, m$, as well as $|\mu_{i11}| = \sqrt{\lambda_i/2}$, $\frac{|\det V_n|}{|\det V_0|} = \prod_{3 \leq k \leq m, k \neq n} |\vartheta_{k1} - \vartheta_{n1}|$. ■

Proof of Lemma 4.3 The proof of Lemma 4.3 relies on the following combinatorial identity.

Lemma A.1: [21] Let $f(t) = \sum_{k=0}^{\infty} f_k t^k$ be a formal power series [20]. Then the following rule holds if $b = 0$ and $f(t)$ is a polynomial:

$$\sum_k C(n + ak, m + bk) z^{m+bk} f_k = [t^m] (1 + zt)^n f(t^{-b}(1 + zt)^a),$$

where $[t^m]g(t)$ denotes the extraction of the coefficient of t^m from the formal power series $g(t)$.

Now let us return to the proof. Denoting

$$\begin{aligned} v(\theta) &= \cos \theta A^{2m-1}(1, \cdot) + \sin \theta A^{2m-1}(2, \cdot) + A^{2m-1}(3, \cdot) \\ &= [v_1(\theta) \ v_2(\theta) \ \dots \ v_{2m-1}(\theta) \ v_{2m}(\theta)] \end{aligned}$$

and recalling $S = A^{2m-1} \mathcal{O}^{-1}$, the original equation is equivalent to $l(\theta) \mathcal{O} = v(\theta)$. Direct computation shows that the entries of $v(\theta)$ are given by

$$\begin{aligned} v_{2j-1}(\theta) &= C(2m, j-1) \cos(2m-j+1)\theta, \\ v_{2j}(\theta) &= C(2m, j-1) \sin(2m-j+1)\theta, \quad j = 1, \dots, m; \end{aligned}$$

and the entries of \mathcal{O} are given by

$$\begin{aligned} \mathcal{O}(k, 2j-1) &= C(k-1, j-1) \cos(k-j)\theta, \\ \mathcal{O}(k, 2j) &= C(k-1, j-1) \sin(k-j)\theta, \quad k = 1, \dots, 2m, \quad j = 1, \dots, m. \end{aligned}$$

As a result, the equation $l(\theta) \mathcal{O} = v(\theta)$ is equivalent to the following m equations:

$$\sum_{k=0}^{2m-1} l_k(\theta) C(k, h) e^{j(k-h)\theta} = C(2m, h) e^{j(2m-h)\theta}, \quad h = 0, 1, \dots, m-1, \quad (65)$$

or equivalently

$$\sum_{k=0}^{2m} l_k(\theta) C(k, h) e^{j(k-h)\theta} = 0, \quad h = 0, 1, \dots, m-1, \quad (65')$$

if we let $l_{2m}(\theta) = -1$. Noticing that $2 \cos \theta = e^{j\theta} + e^{-j\theta}$, we substitute the expression of $l_k(\theta)$ into the left-hand side of the above h -th equation, and expand it into a power series of $e^{j\theta}$ as $\sum_{k=0}^{2m} l_k(\theta) C(k, h) e^{j(k-h)\theta} = \sum_{w=0}^m \alpha_{w,h} e^{j(2w-h)\theta}$, with

$$\alpha_{w,h} = \sum_{k=0}^w \sum_{j=w-k}^{m-k} C(m, j) C(m-j, k) C(j, w-k) C(j+2k, h) (-1)^{j+2k-1}.$$

Therefore, if we can show that $\alpha_{w,h}$ for $w = 0, 1, \dots, m$ and $h = 0, 1, \dots, m-1$, then the prescribed $l(\theta)$ is a solution of (65), and by the nonsingularity of \mathcal{O} it is also unique.

We first transform $\alpha_{w,h}$ as follows. By remembering that

$$\begin{aligned} C(m, j) C(m-j, k) C(j, w-k) &= C(m, k) C(m-k, j) C(j, w-k) \\ &= C(m, k) C(m-k, w-k) C(m-w, j-(w-k)) \end{aligned}$$

and letting $s = j - (w-k)$, it is clear that

$$\alpha_{w,h} = \sum_{k=0}^w C(m, k) C(m-k, w-k) (-1)^{k+w-1} \sum_{s=0}^{m-w} C(m-w, s) C(s+w+k, h) (-1)^s.$$

Now we claim that

- 1). $\sum_{s=0}^{m-w} C(m-w, s) C(s+w+k, h) (-1)^s = C(w+k, h-(m-w)) (-1)^{m-w}$, $w = 0, 1, \dots, m$,
- 2). $\sum_{k=0}^w C(m, k) C(m-k, w-k) C(w+k, h-(m-w)) (-1)^{m+k-1} = 0$, $h = 0, 1, \dots, m-1$,

and the proof of the first part is completed by combining these two claims.

1). Let $f_s = (-1)^s C(m-w, s)$ and $f(t) = \sum_s f_s t^s = (1-t)^{m-w}$. Applying Lemma A.1, we have

$$\begin{aligned} & \sum_{s=0}^{m-w} C(m-w, s) C(s+w+k, h) (-1)^s \\ &= \sum_{s=0}^{m-w} C(w+k+1+s, h+0+s) \cdot 1 \cdot f_s \\ &= [t^h] (1+t)^{w+k} f(1+t) \\ &= [t^h] (1+t)^{w+k} (-t)^{m-w} \\ &= (-1)^{m-w} [t^{h-(m-w)}] (1+t)^{w+k} \\ &= (-1)^{m-w} C(w+k, h-(m-w)), \end{aligned}$$

which establishes the first claim.

2). For the second claim,

$$\begin{aligned} & \sum_{k=0}^w C(m, k) C(m-k, w-k) C(w+k, h-(m-w)) (-1)^{m+k-1} \\ &= (-1)^{m-1} \sum_{k=0}^m C(m, k) (-1)^k C(m-k, w-k) C(w+k, h-(m-w)) \\ &= (-1)^{m-1} \sum_{k=0}^m [t^k] (1-t)^m [v^{m-w}] (1+v)^{m-k} [u^{h-(m-w)}] (1+u)^{w+k} \\ &= (-1)^{m-1} [v^{m-w}] (1+v)^m [u^{h-(m-w)}] (1+u)^w \sum_{k=0}^m [t^k] \left(\frac{1+u}{1+v}\right)^k (1-t)^m \\ &= (-1)^{m-1} [v^{m-w}] (1+v)^m [u^{h-(m-w)}] (1+u)^w \left(1 - \frac{1+u}{1+v}\right)^m \\ &= (-1)^{m-1} [v^{m-w}] (1+v)^m [u^{h-(m-w)}] (1+u)^w \frac{(v-u)^m}{(1+v)^m} \\ &= (-1)^{m-1} [v^{m-w}] [u^{h-(m-w)}] (1+u)^w (v-u)^m \\ & \xrightarrow{r=h-(m-w)} (-1)^{m-1} [v^{m-w}] [u^r] (1+u)^w (v-u)^m \\ &= (-1)^{m-1} [v^{m-w}] \sum_{k=0}^r C(w, r-k) C(m, k) v^{m-k} (-1)^k \\ &= (-1)^{m-1} [v^{k-w}] \sum_{k=0}^r C(w, r-k) C(m, k) (-1)^k, \end{aligned}$$

where t, v, u are indeterminates. Noticing that $r \geq w \Leftrightarrow h-(m-w) \geq w \Leftrightarrow h \geq m$ is contradictory to $h = 0, 1, \dots, m-1$, we have $r < w$ and $k \leq r < w$, which suggests the vanishing of the last equation in the above, and the proof for the first part is complete.

As for the second part, by noting that the exponents of $\cos \theta$ in $l_k(\theta)$ are even when k is an even number, while the exponents are odd when k is an odd number, it can be noted that the sign of each term in $l_k(\theta)$ is the same. Therefore we obtain

$$\begin{aligned} \sum_{k=0}^{2m-1} |l_k(\theta)| &= \sum_{k=0}^{2m-1} \sum_{h=0}^{\lfloor k/2 \rfloor} C(m, k-2h) C(m-(k-2h), h) |2 \cos \theta|^{k-2h} \\ &= \sum_{j=0}^m \beta_j |\cos \theta|^j, \end{aligned}$$

with $\beta_0 = 2^m - 1$ and $\beta_j = 2^j C(m, j) \sum_{h=0}^{m-j} C(m-j, h) = 2^m C(m, j)$ for $j = 1, \dots, m$, and the conclusion follows directly. ■

Acknowledgement The authors would like to thank Dr. Shuai Liu for his valuable suggestions.

REFERENCES

- [1] Z. Qiu, L. Xie, and Y. Hong, "Data rate for quantized consensus of high-order multi-agent systems with poles on the unit circle," in *53rd IEEE Conference on Decision and Control*, Dec. 2014, pp. 3771–3776.
- [2] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and Cooperation in Networked Multi-Agent Systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.
- [3] Y. Cao, W. Yu, W. Ren, and G. Chen, "An Overview of Recent Progress in the Study of Distributed Multi-Agent Coordination," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 1, pp. 427–438, Feb. 2013.
- [4] S. Knorn, Z. Chen, and R. Middleton, "Overview: Collective Control of Multi-agent Systems," *IEEE Transactions on Control of Network Systems*, vol. PP, no. 99, pp. 1–1, 2015.
- [5] A. Kashyap, T. Başar, and R. Srikant, "Quantized consensus," *Automatica*, vol. 43, no. 7, pp. 1192–1203, July 2007.
- [6] A. Nedic, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis, "On distributed averaging algorithms and quantization effects," *IEEE Transactions on Automatic Control*, vol. 54, no. 11, pp. 2506–2517, 2009.
- [7] R. Carli, F. Bullo, and S. Zampieri, "Quantized average consensus via dynamic coding/decoding schemes," *International Journal of Robust and Nonlinear Control*, vol. 20, no. 2, pp. 156–175, Jan. 2010.
- [8] T. Li, M. Fu, L. Xie, and J.-F. Zhang, "Distributed Consensus With Limited Communication Data Rate," *IEEE Transactions on Automatic Control*, vol. 56, no. 2, pp. 279–292, Feb. 2011.
- [9] Keyou You and Lihua Xie, "Network Topology and Communication Data Rate for Consensusability of Discrete-Time Multi-Agent Systems," *IEEE Transactions on Automatic Control*, vol. 56, no. 10, pp. 2262–2275, Oct. 2011.
- [10] Y. Meng, T. Li, and J.-F. Zhang, "Coordination Over Multi-Agent Networks With Unmeasurable States and Finite-Level Quantization," *arXiv:1505.03259 [cs, math]*, May 2015.
- [11] Q. Zhang and J.-F. Zhang, "Quantized Data-Based Distributed Consensus under Directed Time-Varying Communication Topology," *SIAM Journal on Control and Optimization*, vol. 51, no. 1, pp. 332–352, Jan. 2013.
- [12] D. Li, Q. Liu, X. Wang, and Z. Yin, "Quantized consensus over directed networks with switching topologies," *Systems & Control Letters*, vol. 65, pp. 13–22, Mar. 2014.
- [13] A. Olshevsky, "Consensus with Ternary Messages," *SIAM Journal on Control and Optimization*, vol. 52, no. 2, pp. 987–1009, Jan. 2014.

- [14] Tao Li and Lihua Xie, "Distributed Coordination of Multi-Agent Systems With Quantized-Observer Based Encoding-Decoding," *IEEE Transactions on Automatic Control*, vol. 57, no. 12, pp. 3023–3037, Dec. 2012.
- [15] Z. Qiu, L. Xie, and Y. Hong, "Quantized Leaderless and Leader-Following Consensus of High-Order Multi-Agent Systems With Limited Data Rate," *IEEE Transactions on Automatic Control*, vol. 61, no. 9, pp. 2432–2447, Sept. 2016.
- [16] Wei Ren and R. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Transactions on Automatic Control*, vol. 50, no. 5, pp. 655–661, May 2005.
- [17] A. P. Seyranian and A. A. Mailybaev, *Multiparameter stability theory with mechanical applications*. World Scientific, 2003, vol. 13.
- [18] C.-T. Chen, *Linear system theory and design*. Oxford University Press, Inc., 1995.
- [19] H. Baumgärtel, *Analytic perturbation theory for matrices and operators*. Springer, 1985, vol. 15.
- [20] R. Sprugnoli, "An introduction to mathematical methods in combinatorics," *Dipartimento di Sistemi e Informatica Viale Morgagni*, 2006.
- [21] —, "Riordan array proofs of identities in Gould's book," *Published electronically at <http://www.dsi.unifi.it/resp/GouldBK.pdf>*, 2007.