

# SCIENTIFIC REPORTS



OPEN

## Entropy Evolution in Consensus Networks

Shuangshuang Fu<sup>1</sup>, Guodong Shi<sup>2</sup>, Ian R. Petersen<sup>2</sup> & Matthew R. James<sup>2</sup>

We investigate the evolution of the network entropy for consensus dynamics in classical and quantum networks. We show that in the classical case, the network differential entropy is monotonically non-increasing if the node initial values are continuous random variables. While for quantum consensus dynamics, the network's von Neumann entropy is in contrast non-decreasing. In light of this inconsistency, we compare several distributed algorithms with random or deterministic coefficients for classical or quantum networks, and show that quantum algorithms with deterministic coefficients are physically related to classical algorithms with random coefficients.

How agreement emerging among a group of agents interacting each other is an intriguing subject in various research disciplines<sup>1–5</sup>. The fundamental idea lies in that, cooperative decisions can lead to consensus in node states throughout a network even each node can only interact with a few neighboring nodes. Inspired by this, consensus control over networks has been systematically studied in the past decade<sup>6,7</sup>, becoming one of the foundational blocks for engineering solutions to distributed coordination problems in multi-agent systems<sup>8</sup>. Recent work<sup>9,10</sup>, further developed consensus dynamics for quantum networks, where each node corresponds to a qubit<sup>11</sup>. The concepts regarding the network density matrix as statistical ensembles of pure quantum states, reaching a quantum consensus were systematically developed<sup>9</sup>, and it has been shown that a quantum consensus can be reached with the help of quantum swapping operators for both continuous-time and discrete-time dynamics<sup>9,12</sup>. In fact, the two categories of dynamics over classical and quantum networks can be put together into a group-theoretic framework<sup>10</sup>, and quantum consensus dynamics can in fact be equivalently mapped into certain parallel classical dynamics over disjoint subsets of the entries of the network density matrix<sup>12</sup>. This line of research on consensus dynamics is related to the work on quantum walks over complex networks<sup>13,14</sup>, where associated with the network Laplacian there is a Hermitian operator defining the evolution of the network state in a quantum space.

Despite of their algorithmic consistency from a high level between classical and quantum consensus dynamics<sup>10,12</sup>, it is worth investigating the two categories of physical processes from an information perspective. Classical consensus dynamics is realized by nodes observing their neighbors' states (or, their relative states) and then taking real-time feedback. On the other hand, such precise state observation among different components in a quantum network is proven to be impossible<sup>11</sup>, and quantum consensus dynamics is realized via nodes interacting not directly, but with the help of local environments which by themselves are quantum subsystems. Tracing out these local environments, the state evolution of the quantum network follows a master equation with the same attraction nature<sup>12</sup>.

To this end, we investigate the evolution of the network entropy for consensus dynamics in classical and quantum networks. We show that in classical consensus dynamics, the network network differential entropy is monotonically non-increasing if the node initial values are continuous random variables with proper density functions. While for quantum consensus dynamics, the network's von Neumann entropy is in contrast non-decreasing. These observations suggest that the two types of consensus schemes have different physical footings. Then, we compare several gossiping algorithms with random or deterministic coefficients for classical or quantum networks and present novel convergence conditions for gossiping algorithms with random coefficients. The result shows that quantum gossiping algorithms with deterministic coefficients are physically consistent with classical gossiping algorithms with random coefficients.

### Results

Consider a network with  $N$  nodes. The nodes are indexed in the set  $V = \{1, \dots, N\}$ , and their interconnections are described by a connected undirected graph  $G = (V, E)$ , where each element in  $E$  is an unordered pair of nodes in  $V$ . The graph  $G$  is an abstraction to the components (given by  $V$ ) and the structure (given by  $E$ ) of the information flow among the components.

<sup>1</sup>School of Mathematics and Physics, University of Science and Technology Beijing, Beijing, 100083, China. <sup>2</sup>Research School of Engineering, The Australian National University, Canberra, 0200, Australia. Correspondence and requests for materials should be addressed to G.S. (email: [guodong.shi@anu.edu.au](mailto:guodong.shi@anu.edu.au))

Received: 18 April 2016

Accepted: 3 April 2017

Published online: 08 May 2017

**Classical Networks.** In classical scenarios, each node  $i \in V$  holds a real-valued state at time  $t$ , denoted  $X_i(t)$ , representing opinions of a peer in social networks, or signals measured at a sensor in engineering networks. The network state is represented by the vector  $X(t) = (X_1(t) \dots X_N(t))^T$ . The Laplacian of the graph  $G$  is an  $N \times N$  matrix defined by  $L_G = D_G - A_G$ , where  $A_G$  is the adjacency matrix of  $G$  with  $[A_G]_{ij} = 1$  if  $\{i, j\} \in E$  and  $[A_G]_{ij} = 0$  otherwise, and  $D_G$  is a diagonal matrix with the  $i$ th diagonal entry given by the degree of node  $i$ ,  $d_i = \sum_j [A_G]_{ij}$ . In continuous-time settings, DeGroot's type of node interactions lead to a time evolution of  $X(t)$  described by the following differential equation (7)

$$\frac{d}{dt}X(t) = -L_G X(t). \tag{1}$$

The initial time is set to be 0 and each node  $i$  holds an initial value  $X_i(0)$ . During (1) the nodes' states are mixed in such a way that at each time  $t$ , the direction of the movement of  $X_i(t)$  always points to the interior of the convex combination of its neighbors' current states. In the end an average consensus is reached in the sense that (7)

$$X(\infty) := \lim_{t \rightarrow \infty} X(t) = \mathbf{1}\mathbf{1}^T X(0)/N$$

where  $\mathbf{1}$  is the  $N \times 1$  vector with all entries being 1.

The Shannon entropy is a fundamental measure of uncertainty of a random variable<sup>15</sup>. The entropy  $H(Z)$ , of a discrete random variable  $Z$  with alphabet  $\mathcal{Z}$  is defined as  $H(Z) := -\sum_{z \in \mathcal{Z}} p(z) \log p(z)$ . Here  $\log$  is to the base 2 and  $p(\cdot)$  is the probability mass function. The differential entropy  $h(Z)$  of a continuous random variable  $Z$  with density  $f(z)$  is defined as  $h(Z) := -\int_{\mathcal{S}} f(z) \log f(z) dz$ , where  $\mathcal{S}$  is the support of  $Z$ . Then the following result holds.

**Theorem 1** *Let the  $X_i(0)$  be continuous random variables. Then  $h(X(t)) \geq h(X(s))$  for all  $0 \leq t \leq s$  along the system (1).*

When the  $X_i(0)$  are independent and identically distributed (i.i.d.) discrete random variables, the Shannon entropy of  $X(t)$ ,  $H(X(t))$ , is invariant for all  $t \geq 0$  along the system (1) since there is a one-to-one mapping between  $X(t + \varepsilon)$  and  $X(t)$  for any  $t > 0$  and  $\varepsilon > 0$  from the proof of Theorem 1. We also note that the entropy at individual nodes, either  $h(X_i(t))$  or  $H(X_i(t))$ , can certainly admit non-monotone trajectories. Particularly, if the  $X_i(0)$  are i.i.d. Gaussian with mean  $\mu$  and variance  $\sigma^2$ , we know

$$h(X_i(0)) = \frac{1}{2}[\log(2\pi e\sigma^2)], \quad h(X_i(\infty)) = \frac{1}{2} \log(2\pi e\sigma^2/N)$$

since  $X_i(\infty)$  is also a Gaussian random variable but with mean  $\mu$  and variance  $\sigma^2/N$ . On the other hand, if the  $X_i(0)$  are i.i.d. Bernoulli random variables with mean  $p \in (0, 1)$ ,  $NX_i(\infty)$  obeys binomial distribution and therefore

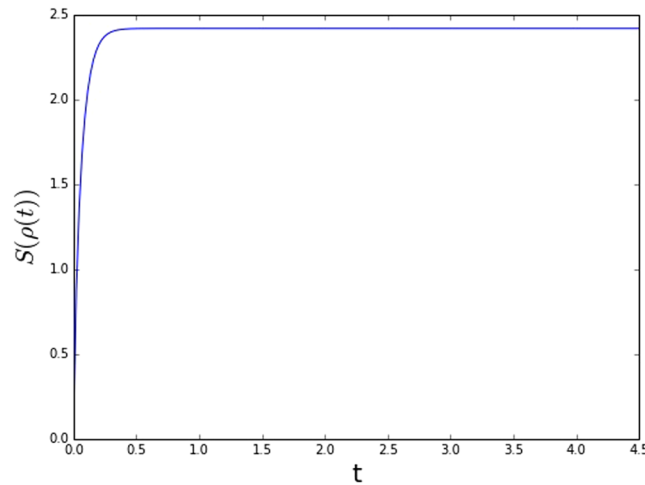
$$H(X(0)) = N[p \log p^{-1} + (1 - p) \log(1 - p)^{-1}], \quad H(X(\infty)) \simeq \frac{1}{2} \log(2\pi e N p(1 - p)) + O\left(\frac{1}{N}\right).$$

**Quantum Networks.** In a simple quantum network, each node  $i \in V$  represents a qubit (quantum bit). The state space associated with any isolated quantum system is a complex vector space with inner product, i.e., a Hilbert space  $\mathcal{H}$ . The system is completely described by its state vector, which is a unit vector in the system's state space and often denoted by  $|\psi\rangle \in \mathcal{H}$  (known as the Dirac notation). The state space of a composite quantum system is the tensor product of the state space of each component system. Two quantum systems with state spaces  $\mathcal{H}_A$  and  $\mathcal{H}_B$ , respectively, form a composite system with state space  $\mathcal{H}_A \otimes \mathcal{H}_B$ , where  $\otimes$  stands for tensor product. If the two quantum systems are isolated respectively with states  $|\psi_A\rangle \in \mathcal{H}_A$  and  $|\psi_B\rangle \in \mathcal{H}_B$ , the composite system admits a state  $|\psi_A\rangle \otimes |\psi_B\rangle$ . Quantum states can also be described by a positive (i.e., positive semi-definite) Hermitian density operator  $\rho$  over the Hilbert space. A quantum state  $|\psi\rangle \in \mathcal{H}$ , induces a linear operator, denoted  $|\psi\rangle\langle\psi|$ , by  $|\psi\rangle\langle\psi|(|x\rangle) = (\langle\psi|x\rangle)|\psi\rangle$  where  $\langle\psi|x\rangle$  is the inner product equipped by the Hilbert space  $\mathcal{H}$  with  $\langle\psi|$  being the dual vector of  $|\psi\rangle$ . In this way the operator  $\rho = |\psi\rangle\langle\psi|$  describes the quantum system at state  $|\psi\rangle$ . Density operators provide a convenient description of *mixed states* as ensembles of pure states: A quantum system in state  $|\psi_i\rangle$  with probability  $p_i$  can be described by  $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ . Any positive and Hermitian operator with trace one defines a proper density operator describing certain quantum state, and vice versa.

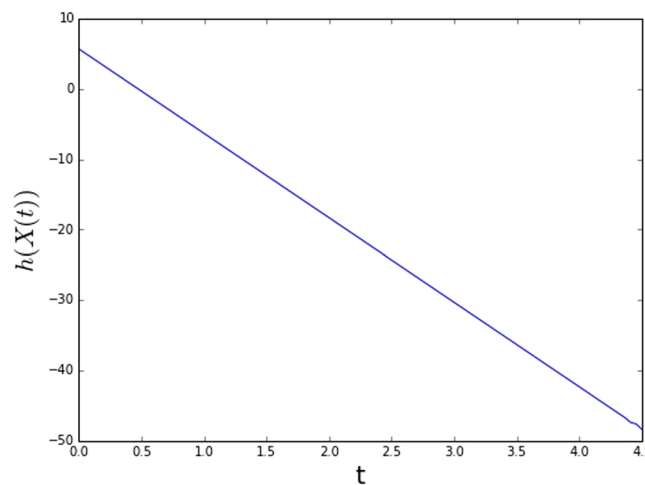
The state of each qubit is represented by a density operator over the two-dimensional Hilbert space  $\mathcal{H}$ , and the network state corresponds to a density operator over  $\mathcal{H}^{\otimes N}$ , the  $N$ th tensor product of  $\mathcal{H}$ . Let  $\rho(t)$  be the network density operator at time  $t$ . The *swapping operator* between qubits  $i$  and  $j$ , denoted as  $U_{ij}$ , is defined by  $U_{ij}(q_1 \otimes \dots \otimes q_i \otimes \dots \otimes q_j \otimes \dots \otimes q_n) = q_1 \otimes \dots \otimes q_j \otimes \dots \otimes q_i \otimes \dots \otimes q_n$  for all  $q_i \in \mathcal{H}$ ,  $i = 1, \dots, n$ . In other words, the swapping operator  $U_{ij}$  switches the information held on qubits  $i$  and  $j$  without changing the states of other qubits. Continuous-time quantum consensus process can be specified by a master equation<sup>11</sup>

$$\frac{d}{dt}\rho(t) = \sum_{(j,k) \in E} (U_{jk}\rho(t)U_{jk} - \rho(t)). \tag{2}$$

For the system (2), there holds<sup>9,12</sup>,



**Figure 1.** The evolution of network entropy for classical network dynamics.



**Figure 2.** The evolution of network entropy for quantum network dynamics.

$$\rho(\infty) := \lim_{t \rightarrow \infty} \rho(t) = \frac{1}{N!} \sum_{\pi \in \mathfrak{P}} U_{\pi} \rho(0) U_{\pi}^{\dagger}$$

where  $\mathfrak{P}$  is the permutation group over  $V$ , and  $U_{\pi}$  represents the quantum permutation operator induced by  $\pi \in \mathfrak{P}$  satisfying  $U_{\pi}(q_1 \otimes \dots \otimes q_n) = q_{\pi(1)} \otimes \dots \otimes q_{\pi(n)}$  for all  $q_i \in \mathcal{H}$  with  $i = 1, \dots, n$ .

As a natural generalization of the Shannon entropy, for a quantum-mechanical system described by a density matrix  $\rho$ , the von Neumann entropy is defined as ref. 11  $S(\rho) = -\text{tr}(\rho \log \rho)$ , where  $\text{tr}(\cdot)$  is the trace operator. The following result holds.

**Theorem 2** For the system (2),  $S(\rho(t))$  is a non-decreasing function over  $[0, \infty)$ .

**Example.** Let  $G$  be a complete graph with 4 nodes. For the classical case, we take the  $X_i(0)$  as an i.i.d. standard Gaussian random variable. For the quantum case, we take the initial density matrix as  $\rho(0) = |01 + -\rangle\langle 01 + -|$ . The evolution of the differential entropy and the von Neumann entropy with the classical and quantum consensus dynamics is plotted, respectively, in Figs 1 and 2.

**Algorithmic and Physical Equivalence.** The above results reveal that, the network entropy in general decreases with classical consensus dynamics, but increases with quantum consensus dynamics. This appears to be surprising noticing their consistencies pointed out in refs 9 and 12. However, although the systems (1) and (2) can be formally united from an algorithmic point of view (cf. ref. 12),  $X(t)$  represents a random variable in the classical world, while  $\rho(t)$  is a probability mass function by its definition. We now provide a *physical* perspective to explain the observations in Theorem 1 and Theorem 2 by investigating a serial of classical or quantum gossiping algorithms with random or deterministic coefficients.

A random gossiping process is defined as follows. Consider  $N$  nodes in the set  $V$  with an underlying interaction graph  $G$  which is undirected and connected. Time is sequenced by  $k = 0, 1, \dots$ . At time  $k$ , a node  $i$  is first drawn with probability  $1/N$ , and then node  $i$  selects another node  $j$  who shares a link with node  $i$  in the graph  $G$  with probability  $1/\text{deg}(i)$ . Here  $\text{deg}(i)$  is the degree of node  $i$  in the graph  $V$ . In this way, a random pair  $\{i, j\}$  is selected. Additionally, let  $b_k, k = 0, 1, \dots$  be a sequence of i.i.d. Bernoulli random variables with mean  $1/2$ , which are also independent of any other possible randomness.

- In the classical case, each node  $i$  holds a real-valued state  $X_i(k) \in R$  at time  $k$ . Their initial states,  $X_1(0), \dots, X_N(0)$ , are assumed to be  $N$  (not necessarily independent, continuous or discrete) random variables over a common underlying probability space. The marginal probability (mass or density) distribution of node  $X_i(k)$  is denoted as  $p_k^i(\cdot)$ . When the pair  $\{i, j\}$  is selected at time  $k$ , only the two selected nodes update their values and we consider the following algorithms.

[A1] (Classical Gossiping with Deterministic Coefficients<sup>16</sup>) Node  $i$  and  $j$  update their values as

$$X_i(k + 1) = X_j(k + 1) = \frac{1}{2}X_i(k) + \frac{1}{2}X_j(k). \tag{3}$$

[A2] (Classical Gossiping with Random Swapping<sup>17</sup>) Node  $i$  and  $j$  update their values as

$$X_i(k + 1) = b_k X_i(k) + (1 - b_k) X_j(k); \quad X_j(k + 1) = (1 - b_k) X_i(k) + b_k X_j(k). \tag{4}$$

- In the quantum case, each node  $i$  represents a qubit and  $\rho(k)$  is the network density matrix at time  $k$ . When the qubit pair  $\{i, j\}$  is selected at time  $k$ , we correspondingly consider the following algorithms.

[AQ1] (Quantum Gossiping with Deterministic Coefficients<sup>9</sup>) The quantum network updates its density matrix as

$$\rho(k + 1) = \frac{1}{2}\rho(k) + \frac{1}{2}U_{ij}\rho(k)U_{ij}^\dagger. \tag{5}$$

[AQ2] (Quantum Gossiping with Random Swapping<sup>10</sup>) Node  $i$  and  $j$  update their values as

$$\rho(k + 1) = b_k \rho(k) + (1 - b_k) U_{ij} \rho(k) U_{ij}^\dagger. \tag{6}$$

We state a few immediate facts for the algorithms [A1], [A2], [AQ1], and [AQ2].

- (i) The evolution of  $\mathbf{E}\{X(k)\}$  is exactly the same along with the algorithms [A1] and [A2]. Similar conclusion holds also for the algorithms [AQ1] and [AQ2].
- (ii) Algorithms [A1] and [AQ1] are *algorithmically equivalent*, in the sense that [AQ1] can be divided into a set of parallel algorithms in the form of [A1] over disjoint entries of  $\rho(t)$  (see ref. 12 for a thorough treatment via vectorizing  $\rho(t)$ ). Similarly, the algorithms [A2] and [AQ2] are algorithmically equivalent.
- (iii) Algorithms [A2] and [AQ1] are *physically equivalent*, in the sense that for a sequence of underlying random variables  $X(k)$  evolving along [A2], their joint probability mass/density function, denoted  $f_k(x_1, \dots, x_N)$  (which is exactly the physical interpretation of the density matrix  $\rho(k)$ ) will evolve in the form of [AQ1] (cf., ref. 10):

$$f_{k+1}(x_1, \dots, x_N) = \frac{1}{2}f_k(x_1, \dots, x_i, \dots, x_j, \dots, x_N) + \frac{1}{2}f_k(x_1, \dots, x_j, \dots, x_i, \dots, x_N) \tag{7}$$

if the pair  $\{i, j\}$  is selected at time  $k$ .

Recall that a Markov chain is *irreducible* if there exists a positive probability that the chain transits from one state to another in some finite steps between any pair of states; *aperiodic* if the greatest common divisor of the minimal steps admitting a positive probability from any state to itself is exactly one; *ergodic* if it is both aperiodic and irreducible<sup>18</sup>. We present the following result establishing the limiting behaviors of the algorithm [A2], which is consistent with the observations of the entropy evolution in Theorems 1 and 2 as well as the point (iii) above.

**Theorem 3** For the algorithm [A2], the following statements hold. (i)  $\{X(k)\}_{k=0}^\infty$  forms an ergodic Markov chain given  $X(0)$ ; (ii)  $\lim_{k \rightarrow \infty} p_k^i(\cdot) = \sum_{i=1}^N p_0^i(\cdot)/N$ , where the convergence is exponentially fast under the distance induced by  $\ell^1$  (for  $X(0)$  given by discrete random variables) or  $\mathcal{L}^f$  (for continuous  $X(0)$  norms); (iii) the network entropy (continuous or discrete) is non-decreasing.

Note that Theorem 1 indeed shows that the network entropy along the algorithm [A1] is non-increasing in the limit for certain classes of initial distributions. It is reasonable to believe that such non-increasing entropy holds in general for the algorithm [A1], although establishing a rigorous proof would be challenging. In the algorithms [A1] and [A2], the randomness of  $X(k)$  comes from three resources: initial random draws  $X(0)$ , random pair selections, and external randomness in the updates (the Bernoulli random variables  $b_k$ ). It is clear that we can assume a deterministic pair sequence in their updates, all the above conclusions in Theorems 1 and 3 continue to hold as long as this deterministic sequence visits each pair in the graph sufficiently often. It now becomes clear that it is the external randomness in [A2] driven by  $\{b_k\}$  that causes the increasing of network entropy.

One can also consider the case in a gossiping process when two selected node  $i$  and  $j$  update their values by (Classical Gossiping with Random Coefficients)

$$[A1'] X_i(k + 1) = X_j(k + 1) = b_k X_i(k) + (1 - b_k) X_j(k). \tag{8}$$

From the second Borel-Cantelli Lemma (e.g., Theorem 2.3.6. in ref. 18), that almost surely,  $X_i(k)$  reaches a common value for all  $i \in V$  in finite time along the algorithm [A1']. Interestingly, it is easy to see that the evolution of the  $p_k^i(\cdot)$  is the same along the algorithms [A1'] and [A2].

The scheme of the algorithms [A2] was briefly discussed in Section 6.2 of ref. 10, which is also a form of gossiping algorithms with unreliable but perfectly dependent link communications studied in ref. 17 with mixing coefficient one. Here Theorem 3 advances the previous understandings by showing that the algorithm [A2] defines an ergodic Markov chain for any given initial condition as well as presenting the detailed convergence properties of the marginal distribution functions for both discrete and continuous  $X(0)$ . Moreover, we assume that the mean of the  $b_k$  is 1/2 just for the ease of presentation. It is clear from the proof that Theorem 3 holds for arbitrary  $\mathbf{E}\{b_k\} \in (0,1)$ . The ergodicity plays an essential role in the convergence of the marginal distributions: the case with  $\mathbf{E}\{b_k\} = 0$  fails because  $X(k)$  is no longer aperiodic; the case with  $\mathbf{E}\{b_k\} = 1$  fails because  $X(k)$  is no longer irreducible.

### Methods

This section provides the proofs of Theorems 1, 2 and 3.

#### Proof of Theorem 1

The solution  $X(t)$  of the system (1) is  $X(t) = e^{-tL_G}X(0)$ . We take  $\varepsilon > 0$  and compare  $h(X(t + \varepsilon))$  with  $h(X(t))$ . From

$$X(t + \varepsilon) = e^{-\varepsilon L_G}X(t),$$

we know (Theorem 8.6.4<sup>15</sup>)

$$h(X(t + \varepsilon)) = h(X(t)) + \log|\det(e^{-\varepsilon L_G})|.$$

Since  $L_G$  is the Laplacian of a connected undirected graph  $G$ ,  $L_G$  has a unique zero eigenvalue, and all non-zero eigenvalues of  $L_G$  are positive<sup>8</sup>. Consequently, all eigenvalues of  $e^{-\varepsilon L_G}$  are positive and no larger than one, which yields that  $|\det(e^{-\varepsilon L_G})| \leq 1$ . This proves  $h(X(t + \varepsilon)) \leq h(X(t))$ . Since  $\varepsilon$  is chosen arbitrarily, Theorem 1 holds.  $\square$

#### Proof of Theorem 2

Fix  $s \geq 0$ . Define a set  $\Sigma_s = \text{co}(U_\pi \rho(s) U_\pi^\dagger : \pi \in \mathfrak{P})$ , where  $\text{co}(\cdot)$  stands for the convex hull. It is straightforward to see that  $U_{jk} \rho U_{jk}^\dagger \in \Sigma_s$  if  $\rho \in \Sigma_s$ . As a result,  $\Sigma_s$  is an invariant set of the system (2) in the sense that  $\rho(t) \in \Sigma_s$  for all  $t \geq 0$  as long as  $\rho(0) \in \Sigma_s$ . Therefore, for the system (2) and for any  $\varepsilon > 0$ , there exist  $m_\pi(\varepsilon) \geq 0$ ,  $\pi \in \mathfrak{P}$  with  $\sum_{\pi \in \mathfrak{P}} m_\pi(\varepsilon) = 1$  such that

$$\rho(s + \varepsilon) = \sum_{\pi \in \mathfrak{P}} m_\pi(\varepsilon) U_\pi \rho(s) U_\pi^\dagger. \tag{9}$$

Recalling that the von Neumann entropy  $S(\rho)$  is a concave function of  $\rho$ , and that  $S(\rho) = S(U\rho U^\dagger)$  for any unitary operator  $U$ , we conclude that

$$S(\rho(s + \varepsilon)) = S\left(\sum_{\pi \in \mathfrak{P}} m_\pi(\varepsilon) U_\pi \rho(s) U_\pi^\dagger\right) \geq \sum_{\pi \in \mathfrak{P}} m_\pi(\varepsilon) S(U_\pi \rho(s) U_\pi^\dagger) = S(\rho(s)) \tag{10}$$

for any  $\varepsilon > 0$  and  $s \geq 0$  in light of the fact that  $U_\pi$  is unitary for all  $\pi \in \mathfrak{P}$ . This proves that  $S(\rho(t))$  is a non-decreasing function and Theorem 2 holds.  $\square$

#### Proof of Theorem 3

- (i) First of all it is clear that  $\{X(k)\}_{k=0}^\infty$  is Markovian from its definition. Recall that  $\mathfrak{P}$  is the  $N$ 'th permutation group. We denote the permutation matrix associated with  $\pi \in \mathfrak{P}$  as  $M_\pi$ . In particular, the permutation matrix associated with the swapping between  $i$  and  $j$  is denoted as  $M_{ij}$ . The state transition of  $\{X(k)\}_{k=0}^\infty$  along the algorithm A2 can be written as  $\mathbf{P}(M_{ij} X(k) | X(k)) = (1/\deg(i) + 1/\deg(j))/N$  for  $\{i, j\} \in E$ . Since the graph  $G$  is connected, the swapping permutations defined along the edges of  $G$  form a generating set of the permutation group  $\mathfrak{P}$ . Consequently, given  $X(0)$ , the set  $\{M_\pi X(0), \pi \in \mathfrak{P}\}$  is the state space of  $X(k)$ , which contains at most  $N!$  elements. Finally it is straightforward to verify that for any given  $X(0)$ ,  $X(k)$  is irreducible (due to the fact that any two states in the above state space can reach each other by a finite number of swapping permutations along the edges of  $G$ ) and aperiodic (due to the fact that there is a positive probability that  $X(k) = X(k + 1)$  when  $b_k = 1$ ), and therefore forms an ergodic Markov chain.
- (ii) The statement is in fact a direct consequence from the ergodicity of  $X(k)$ . We however need to be a bit more careful since we assume that  $X(0)$  takes value from an arbitrary (not necessarily discrete) probability space and the  $X(0)$  are not necessarily independent. We denote the state transition matrix for  $X(k)$  as  $P \in \mathbb{R}^{N \times N}$ . We calculate  $p_k^i(\cdot)$  from basic probability equality  $\mathbf{P}(A) = \sum_{s=1}^m \mathbf{P}(A | C_s)$  under  $\sum_{i=1}^N \mathbf{P}(C_i) = 1$  and

$P(C_i \cap C_j) = 0$ , and then immediately obtain  $p_k^i(\cdot) = \sum_{s=1}^N e_s^T P^k e_i p_0^s(\cdot)$ , where  $e_i$  is the unit vector with the  $i$ 'th entry being one. It is clear that the above calculation does not rely on  $X(0)$  being discrete or continuous, and  $p_k^i(\cdot)$  represents probability mass or density function wherever appropriate. From the definition of the algorithm A2,  $P$  is a symmetric matrix and the ergodicity of  $X(k)$  leads to  $\lim_{k \rightarrow \infty} P^k = \mathbf{1}\mathbf{1}^T/N$  at an exponential rate.

(iii) This observation follows directly from (7) and the concavity of (continuous or discrete) entropy.

The desired conclusion thus follows.

## Conclusions

We have investigated the evolution of the network entropy for consensus dynamics in classical or quantum networks. In the classical case, the network differential entropy is monotonically non-increasing with continuous initial random values. For quantum consensus dynamics, the network's von Neumann entropy is on the contrary non-decreasing. This observation can be easily generalized to balanced directed graphs<sup>19</sup>. In light of this inconsistency, we also compared several gossiping algorithms with random or deterministic coefficients for classical or quantum networks, and showed that quantum gossiping algorithms with deterministic coefficients are physically consistent with classical gossiping algorithms with random coefficients.

## References

1. Wilde, M. M. *Quantum information theory* (Cambridge University Press, 2013).
2. Conradt, L. & Roper, T. J. Consensus decision making in animals. *Trends in Ecology & Evolution* **20**, 449–456, doi:10.1016/j.tree.2005.05.008 (2005).
3. DeGroot, M. H. Reaching a consensus. *Journal of the American Statistical Association* **69**, 118–121, doi:10.1080/01621459.1974.140137 (1974).
4. Tsitsiklis, J. N. *Problems in Decentralized Decision Making and Computation*. (Ph.D. thesis, Dept. of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Boston, MA, 1984).
5. Vicsek, T., Czirok, A., Jacob, E. B., Cohen, I. & Schochet, O. Novel type of phase transitions in a system of self-driven particles. *Physical Review Letters* **75**, 1226–1229, doi:10.1103/PhysRevLett.75.1226 (1995).
6. Jadbabaie, A., Lin, J. & Morse, A. S. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Trans. Autom. Control* **48**, 988–1001, doi:10.1109/TAC.2003.812781 (2003).
7. Olfati-Saber, R. & Murray, R. M. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Trans. Autom. Control* **49**, 1520–1533, doi:10.1109/TAC.2004.834113 (2004).
8. Mesbahi, M. & Egerstedt, M. *Graph Theoretic Methods in Multiagent Networks*. (Princeton University Press, 2010).
9. Mazzarella, L., Sarlette, A. & Ticozzi, F. Consensus for quantum networks: from symmetry to gossip iterations. *IEEE Trans. Autom. Control* **60**, 158–172, doi:10.1109/TAC.2014.2336351 (2015).
10. Mazzarella, L., Ticozzi, F. & Sarlette, A. From consensus to robust randomized algorithms: A symmetrization approach. *quant-ph, arXiv 1311.3364* (2013).
11. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information*. (10th Edition, Cambridge University Press, 2010).
12. Shi, G., Dong, D., Petersen, I. R. & Johansson, K. H. Reaching a quantum consensus: master equations that generate symmetrization and synchronization. *IEEE Trans. Autom. Control* **61**, 374–387 (2016).
13. Faccin, M., Johnson, T., Biamonte, J., Kais, S. & Migdal, P. Degree distribution in quantum walks on complex networks. *Phys. Rev. X* **3**, 041007, doi:10.1103/PhysRevX.3.041007 (2013).
14. Faccin, M., Migdal, P., Johnson, T. H., Bergholm, V. & Biamonte, J. D. Community detection in quantum complex networks. *Phys. Rev. X* **4**, 041012, doi:10.1103/PhysRevX.4.041012 (2014).
15. Cover, T. M. & Thomas, J. A. *Elements of Information Theory*. (1st Edition, New York: Wiley-Interscience, 1991).
16. Boyd, S., Ghosh, A., Prabhakar, B. & Shah, D. Randomized gossip algorithms. *IEEE Trans. Information Theory* **52**, 2508–2530, doi:10.1109/TIT.2006.874516 (2006).
17. Shi, G., Johansson, M. & Johansson, K. H. Randomized gossiping with unreliable communication: dependent and independent node updates. *The 51st IEEE Conference on Decision and Control*, 4846–4851 (2012).
18. Durrett, R. *Probability: Theory and Examples*. (Duxbury advanced series, Third Edition, Thomson Brooks/Cole, 2005).
19. Shi, G., Fu, S. & Petersen, I. R. Consensus of Quantum Networks with Directed Interactions: Fixed and Switching Structures. *IEEE Transactions on Automatic Control* **62**, 549–56, doi:10.1109/TAC.2016.2590503 (2017).

## Author Contributions

S.F. and G.S. conceived and designed initial idea of the study. S.F., G.S., I.R.P. and M.R.J. contributed to generalization of the original idea and guided technical direction of the research. S.F. and G.S. conducted the analysis and wrote the manuscript. All authors reviewed the manuscript.

## Additional Information

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017