2015 IEEE 54th Annual Conference on Decision and Control (CDC)
December 15-18, 2015. Osaka, Japan

Network Entropy: A System-Theoretic Perspective

Mathias Hudoba de Badyn, Airlie Chapman, and Mehran Mesbahi

***Abstract*— In this paper, we highlight the importance of two measures associated with networked dynamic systems, namely the loop entropy and the Kolmogorov-Sinai entropy, that quantify the notion of information content in the network. We then proceed to show connections between these measures and certain system theoretic properties that these networks exhibit for two classes of network dynamics. Throughout the paper, we also provide relevant bounds and insights into what these network measures quantify.**

I. INTRODUCTION

Networked dynamic systems are ubiquitous in many areas of engineering, and one can argue that they are of paramount importance for understanding complex phenomena in a wide range of disciplines including sociology (social networks [1]), physics (quantum networks [2], [3]), and biology (gene networks [4]).

As such, there has been a surge of research activities in distinct scientific areas with two fundamental questions at their core. Firstly, what is the role of the network and its structure in characterizing the behavior that these networked systems exhibit? Secondly, to what extent can this structure be used for the purpose of control?

Furthermore, it is desired to provide such a characterization with an eye towards large networks. Some of the such open problems are: to what extent can these characterizations be done efficiently as the number of nodes and edges in the network grow? Do the computational requirements corresponding to this characterization depend on the local network structure? Does the study of system measures lead to a conceptual framework to examine dynamic networks and their system theoretic properties [5]?

In this paper, we highlight the importance of two measures associated with the network that we believe have not been examined more thoroughly in the context of their system theoretic significance. These measures quantify a notion of "information content" in the network- and as such, they are referred to as the loop entropy and the Kolmogorov-Sinai (KS) entropy. Both measures have been extensively employed in the physics literature-however, we believe they should be more systematically studied in the networked dynamic systems due to their inherent connections with basic system theory of networks; a related work in this direction include that of Siami and Motee [6].

The research of the authors was supported by the U.S. Army Research Laboratory and the U.S. Army Research Office under contract number W911NF-13-1-0340, as well as by the University of Washington College of Engineering Dean's Fellowship. The authors are with the William E. Boeing Department of Aeronautics & Astronautics, University of Washington, Seattle. Emails: {hudomath, airliec, mesbahi}@uw.edu

The contribution of the paper is twofold. First, we examine the connection between loop entropy and noise propagation properties of networks. The loop entropy is defined separately for both the structure of the network, as well as for how the dynamics respond to a noise input. We show in Section III that these entropies are the same for the consensus dynamics. Secondly, we consider the KS entropy in Section IV in the context of networks evolving via their adjacency matrix. Some examples of entropies are given in Section V.

II. NETWORK MODEL

In this section, we provide the necessary background on the underlying model that will be considered in this paper.

A. Mathematical Notation and Graph Theory

The identity matrix is defined as I and e_i is the column vector with all zeros except a 1 in the ith position. A symmetric matrix A is *positive semi definite* if for all $v \in \mathbb{R}^n$, $v^T Av \geq 0$. The *positive semi-definite ordering* is defined on two symmetric matrices A, B such that $A \succeq B$ if $A - B$ is positive semi-definite. The matrix $A^{\{n-1\}}$ is a matrix constructed from A by removing the first row and column.

The fundamental algebraic structure that represents a network is the graph $\mathcal{G}(V, E, W)$ with a vertex set V of cardinality $|V| = n$, an edge set E of cardinality $|E| = m$, and a set of weights W on the edges. The edge set consists of pairs of vertices i and j that are called *adjacent* if $\{i, j\} \in E \subseteq V^2$, and this edge has weight $w_{ij} \in W_{\mathcal{G}}$. The number of adjacent nodes to node i is called the *degree*, and is denoted by d_i. The *degree matrix* is denoted as $\Delta = \mathrm{Diag}(d_i)$. The weights define an $m \times m$ *weight matrix* $W_{\mathcal{G}}$ with the lth edge $\{i, j\}$ having weight w_{ij} on the diagonal entry $(W_{\mathcal{G}})_{ll}$. The *adjacency matrix* A is defined such that a_{ij} has weight w_{ij} if there is an edge $\{i, j\}$ and 0 otherwise. The *graph Laplacian* $L_{\mathcal{G}}$ is given by $\Delta - A$.

The graph is said to be *connected* if there exists a path of edges from every node to every other node. The graph Laplacian is a positive semi-definite matrix, and the multiplicity of the zero eigenvalue denotes the number of connected components of the graph. The eigenvalues of the Laplacian are ordered such that $0 = \lambda_1 \leq \cdots \leq \lambda_n$. Similary, we order the eigenvalues of the adjacency matrix such that $\mu_1 \leq \cdots \leq \mu_n$.

B. Network Models

In this paper, we consider we consider four models. The first is the *controlled consensus dynamics* [7]:

$$\dot{x} = -L_{\mathcal{G}}x(t) + Bu(t) + G\omega(t), \tag{1}$$

where $\omega(t)$ is a zero-mean Gaussian random vector depicting a random disturbance to the network, $x(t)$ is the states of the nodes of the network, and $u(t)$ is the control input. For compactness of presentation, the time t indication on the states will be henceforth removed.

The *single integrator leader-follower dynamics* can be interpreted as a leader agent x_l which provides an input signal to the set of follower agents x_f via their adjacent edges. This allows the graph Laplacian to be partitioned into two sub-Laplacians for the leaders and followers:

$$L_{\mathcal{G}} = \left[\begin{array}{c|c} a & -B_{\mathcal{G}}^T \\ \hline -B_{\mathcal{G}} & \mathcal{A}_{\mathcal{G}} \end{array} \right] \tag{2}$$

with a perturbation matrix $B_{\mathcal{G}}$ encoding the connections between the leader and the followers:

$$\dot{x}_f = -\mathcal{A}_{\mathcal{G}} x_f + B_{\mathcal{G}} u + G\omega. \tag{3}$$

Here we take $G = I$.[1] The matrix $\mathcal{A}_{\mathcal{G}}$ is known as the Dirichlet matrix, or grounded Laplacian [8], [9]. It has the property of being strictly positive definite when the graph is connected, and hence is invertible unlike the original Laplacian $L_{\mathcal{G}}$.

Similarly, we can define the *double integrator dynamics* as,

$$\begin{aligned} \begin{bmatrix} \dot{x}_f \\ \ddot{x}_f \end{bmatrix} &= \begin{bmatrix} 0 & I \\ -\mathcal{A}_{\mathcal{G}} & -\mathcal{A}_{\mathcal{G}} \end{bmatrix} \begin{bmatrix} x_f \\ \dot{x_f} \end{bmatrix} + \begin{bmatrix} 0 \\ B_{\mathcal{G}} \end{bmatrix} u + \begin{bmatrix} 0 \\ G \end{bmatrix} \omega \\ &:= -\bar{A}_{\mathcal{G}} \begin{bmatrix} x_f \\ \dot{x_f} \end{bmatrix} + \bar{B}_{\mathcal{G}} u + \bar{G}\omega \end{aligned} \tag{4}$$

The *controlled adjacency dynamics* are given by

$$\dot{x} = Ax + Bu. \tag{5}$$

There has been a large body of work centered around finding connections between network structure and the system properties of the corresponding networked system. For example, it has been shown that symmetry structures in the graph topology undermine the controllability of the network [10]. Various notions of robustness measures have been introduced in an effort to describe the performance of networks, such as the convergence time $\tau = 1/\lambda_2$, and the utility of these measures often depends on the network topology [7], [5].

One quantity that captures system behavior with relation to system topology in other fields of science is that of entropy. Fundamentally, entropy is a measure that captures the randomness or disorder of a system. For example, a polymer can be modeled as a vector random walk where each iteration of the walk adds a monomer of some small length from tip-to-tail of the previous iteration [11]. If the polymer is extremely straight, its 'disorder' is very low and hence the entropy is small. On the other hand, a very disordered or loopy polymer has high entropy. The entropy of a polymer has a direct analytical effect on the dynamics of how the polymer behaves and evolves in a fluid at some temperature. In this paper, we explore two notions of entropy defined for a network, and show that for the leader-follower dynamics we can indeed link the topological disorder to control-theoretic properties of the network.

[1] This can be done without loss of generality since if G is non-identity, the effect of multiplying it with ω can be propagated into the covariance of the noise of ω.

III. Loop Entropy and Spanning Trees

In this section, we define the standard notion for the entropy of a network. Using a well-known result by Kirchoff, we show that this notion of entropy is related to the number of spanning trees in the graph representing the Laplacian. Although a similar line of inquiry has been pursued in [6] (for a different notion of entropy) and [7], there are several interpretations of this result with respect to leader-follower dynamics that have not been explored in the literature.

Definition 1: The *entropy* of a network $\mathcal{G}$ is given by the sum $S_{\mathcal{G}} = \sum_{i=2}^n \log \lambda_i$.

The non-zero eigenvalues of $L_{\mathcal{G}}$ are related to those of a submatrix $L_{\mathcal{G}}^{\{n-1\}}$ of $L_{\mathcal{G}}$. For clarity, we will refrain from connecting $L_{\mathcal{G}}^{\{n-1\}}$ to $\mathcal{A}_{\mathcal{G}}$ as seen in the partition in (2) until we complete the exposition of the properties of $S_{\mathcal{G}}$. We will make this connection at the end of the section. The matrix $L_{\mathcal{G}}^{\{n-1\}}$ is constructed by removing an arbitrary ith row and column of $L_{\mathcal{G}}$. The eigenvalues of $L_{\mathcal{G}}^{\{n-1\}}$ are strictly positive, and satisfy:

$$\begin{aligned} \frac{1}{n}\lambda_2(L_{\mathcal{G}}) \cdots \lambda_n(L_{\mathcal{G}}) &= \det L_{\mathcal{G}}^{\{n-1\}} \\ &= \lambda_1(L_{\mathcal{G}}^{\{n-1\}}) \cdots \lambda_{n-1}(L_{\mathcal{G}}^{\{n-1\}}). \end{aligned} \tag{6}$$

It is important to note that the eigenvalues of $L_{\mathcal{G}}^{\{n-1\}}$ are all positive because it is a Dirichlet matrix (or grounded Laplacian) of $L_{\mathcal{G}}$. It should now be clear that

$$S_{\mathcal{G}} = \textstyle\sum_{i=2}^n \log \lambda_i(L_{\mathcal{G}}) = \sum_{i=1}^{n-1} \log \lambda_i(L_{\mathcal{G}}^{\{n-1\}}) + \log n.$$

We can connect this definition to a well-known result by Kirchoff. First, recall that a *tree* is a connected undirected graph with no cycles. Furthermore, a *spanning tree* of a graph $\mathcal{G}$ is a tree containing every vertex of $\mathcal{G}$. Let $\tau(\mathcal{G})$ denote the number of unique spanning trees of $\mathcal{G}$. A well-known result by Kirchoff states that the $\det L_{\mathcal{G}}^{\{n-1\}} = \tau(\mathcal{G})$. The following observation is thus immediate.

Theorem 2: Let $L_{\mathcal{G}}$ be the unweighted network Laplacian for the consensus dynamics. The network entropy is given by $S_{\mathcal{G}} = \log(n\tau(\mathcal{G}))$.

Theorem 2 motivates a new paradigm in terms of interpreting the entropy measure of a networked system. We can now begin to view the entropy as being a fundamentally topological quantity. This formulation of entropy also has a very elegant analogy to Boltzmann entropy in statistical physics. Boltzmann defined the entropy of a physical system, say a gas, to be $S = k_b \log W_s$ where W_s is the number of microstates of the physical system and k_b is a constant [12]. The formula given in Theorem 2 is the number of possible paths a signal can take to propagate throughout the network. In this sense, we have a notion of systemic 'disorder' that captures the idea of enumerating possible states of the system, where in our case the quantity evolving the system is the control signal.

The system-theoretic interpretation motivating the examination of the determinant of the submatrix $L_{\mathcal{G}}^{\{n-1\}}$ in the entropy formula is the following. We take our system and designate a single node as the leader, and the rest of the nodes as followers. If we enumerate the leader node to correspond to the first vertex, then our Laplacian matrix is partitioned exactly according to:

$$L_{\mathcal{G}} = \left[\begin{array}{c|c} a & \delta^T \\ \hline \delta & L_{\mathcal{G}}^{\{n-1\}} \end{array}\right]$$

where $\mathcal{A}_{\mathcal{G}} = L_{\mathcal{G}}^{\{n-1\}}$ in the dynamics (3), and the formula of entropy is $S_{\mathcal{G}} = \log\det \mathcal{A}_{\mathcal{G}}$. In the next section, we will motivate another graph-centric notion of this entropy, and then discuss how these notions relate to graph controllability.

A. Counting Loops

In this section, we describe a method of associating the notion of entropy to a notion of loops in the network. We do this indirectly: we identify the Dirichlet matrix of a network Laplacian to a non-unique flow graph. For the purposes in this paper, we can consider a flow graph to be a directed graph with self-loops on each node. From this flow graph, the usual notion of the Coates Determinant is then examined; we will see that this will depend on the number of loops in the flow graph. This will shed light on a relation between the number of loops in the flow graph and the the number of spanning trees.

Definition 3: A *connection* C of a flow graph $\mathcal{G}$ is a subgraph of $\mathcal{G}$ with the following properties:

1) Each node of C is in $\mathcal{G}$
2) Each node of C has a single edge originating from it and terminating at it.

Definition 4: A *directed loop* of $\mathcal{G}$ is a subgraph of $\mathcal{G}$ whose edges $b_1, \ldots, b_l$ can be ordered such that

1) The tip of b_k is the origin of b_{k+1}, for $k = 1, \ldots, l-1$
2) The origin of b_1 is the tip of b_l.
3) Each node tip along the loop is encountered only once.

A fundamental result by Coates allows us to compute the determinant of $L_{\mathcal{G}}^{\{n-1\}}$ using a digraphic interpretation of this Dirichlet matrix. We can construct the flow graph corresponding to $L_{\mathcal{G}}^{\{n-1\}}$ as follows: Take each diagonal element $(L_{\mathcal{G}}^{\{n-1\}})_{ii}$ and identify it to a self-loop on the ith node with weight $(L_{\mathcal{G}}^{\{n-1\}})_{ii}$. Then, each node i is connected to node j in one direction with weight $(L_{\mathcal{G}}^{\{n-1\}})_{ij}$ and in the other direction with weight $(L_{\mathcal{G}}^{\{n-1\}})_{ji}$.

Lemma 5 (Coates Determinant): Let $\mathcal{G}^{\{n-1\}}$ be the flow graph associated with the Dirichlet matrix $L_{\mathcal{G}}^{\{n-1\}}$. Let the subscript ρ denote a connection of the flow graph $\mathcal{G}$. Then, let the total gain $C(\mathcal{G})_\rho$ of the connection ρ be the product of all the edge weights $(L_{\mathcal{G}}^{\{n-1\}})_{ij}$ in the connection ρ. Finally, let L_ρ be the number of directed loops in the connection ρ. Then, the determinant of $L_{\mathcal{G}}^{\{n-1\}}$ is given by:

$$\det L_{\mathcal{G}}^{\{n-1\}} = \Delta_c = (-1)^{n-1}\big(\textstyle\sum_\rho (-1)^{L_\rho} C(\mathcal{G}^{\{n-1\}})_\rho\big),$$

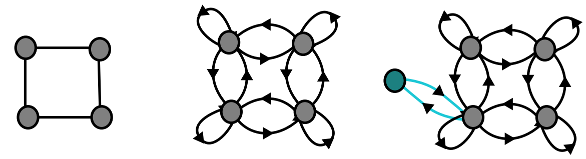

Fig. 1. The two flow graphs of the transfer function of a 4-cycle controlled from a single node.

where the sum is taken over all connections ρ.

Proof: A proof of this lemma is given in [13]. ■

This determinant formula is interesting because it depends on the number of directed loops in the flow graph - a topological feature.

Theorem 6: Let $L_{\mathcal{G}}$ be the Laplacian for the consensus dynamics; let $\mathcal{G}^{\{n-1\}}$ be the flow graph associated with the Dirichlet matrix $L_{\mathcal{G}}^{\{n-1\}}$. Then,

$$S_{\mathcal{G}} = -\log\left[(-1)^n \textstyle\sum_\rho (-1)^{L_\rho} C(\mathcal{G}^{\{n-1\}})_\rho\right]. \quad (7)$$

An application of the Coates determinant formula is for computing the transfer function of the controlled consensus dynamics (1). Assuming the system output is given as $y = B_{\mathcal{G}}^T x$, the transfer function for the network assumes the form,

$$T(s) = \begin{vmatrix} sI + L_{\mathcal{G}} & -B \\ B^T & 0 \end{vmatrix} / |sI + L_{\mathcal{G}}|. \quad (8)$$

To use the Coates determinant formula, we must create two flow graphs corresponding to the numerator and denominator of $H(s)$. The flow graph of the denominator will have self-loops at each ith node with gain $s + L_{\mathcal{G}_{ii}}$, and each edge $\{i, j\}$ of $L_{\mathcal{G}}$ will be replaced by a digon (pair of directed edges pointing in opposite directions) with gains $L_{\mathcal{G}_{ij}}$. The numerator will have an additional node for each column of B (corresponding to separate input channels) connected with a digon to each leader node with gains ± 1. An example is shown in Figure 1, where a 4-cycle is controlled,[2] with $B = [1, 0, 0, 0]^T$. The center and right flow graphs correspond to the denominator and numerator in (8), respectively. The Coates formula provides the coefficients of the polynomials in the transfer function, which in this case is

$$T(s) = \frac{4 + 10s + 6s^2 + s^3}{16s + 20s^2 + 8s^3 + s^4} = \frac{(2+s)(2+4s+s^2)}{s(2+s)^2(4+s)}.$$

B. Loop Entropy and Network Gramians

In this section, we highlight a result that connects the loop entropy to controllability properties of the consensus dynamics. Recalling the previous discussion highlighting the two interpretations of the loop entropy submatrix $L_{\mathcal{G}}^{\{n-1\}}$, without loss of generality let us assume that the loop entropy takes the form $S_{\mathcal{G}} = \log\det \mathcal{A}_{\mathcal{G}} = \sum_{i=1}^n \log \lambda_i(\mathcal{A}_{\mathcal{G}})$.

Definition 7 (Controllability Gramian): Consider the full single integrator consensus dynamics driven by white noise. Then, the *controllability Gramian* is the unique positive semi-definite matrix Σ satisfying

$$-\mathcal{A}_{\mathcal{G}}\Sigma - \Sigma \mathcal{A}_{\mathcal{G}}^T = -GG^T. \quad (9)$$

[2] This system is uncontrollable to illustrate the pole-zero cancellation.

The controllability Gramian describes the mapping from inputs to steady state behavior of the internal states. In the meantime, it turns out that there is an intimate relationship between the controllability Gramian and the loop entropy of the system.

Theorem 8 (Controllability Gramian and Loop Entropy): Let $\mathcal{A}_\mathcal{G}$ be the Dirichlet matrix of $L_\mathcal{G}$ in the full single integrator dynamics (3), let Σ be its controllability Gramian, and let $G = I$. Then, the loop entropy is given by $S_\mathcal{G} = -\log\left(2^n \det \Sigma\right)$.

Proof: Since $\mathcal{A}_\mathcal{G}$ is positive definite, the controllability Gramian satisfies (9), and hence is given by $\Sigma = \frac{1}{2}\mathcal{A}_\mathcal{G}{}^{-1}$ and so $(2^n \det \Sigma)^{-1} = \det \mathcal{A}_\mathcal{G}$. Taking the logarithm of both sides yields the result. ■

Theorem 9 (Loop Entropy for the Double Integrator): Let $\bar{\mathcal{A}}_\mathcal{G}$ be the dynamics matrix from the full double integrator dynamics (4):

$$\bar{\mathcal{A}}_\mathcal{G} = \begin{bmatrix} 0 & -I \\ \mathcal{A}_\mathcal{G} & \mathcal{A}_\mathcal{G} \end{bmatrix}. \tag{10}$$

Next, let $\bar{\Sigma}$ be the controllability Gramian satisfying

$$-\bar{\mathcal{A}}_\mathcal{G}\bar{\Sigma} - \bar{\Sigma}\bar{\mathcal{A}}_\mathcal{G}^T = -\bar{G}\bar{G}^T, \tag{11}$$

where $\bar{G} = [0,\ I]^T$ (4). Then, the loop entropy satisfies $S_\mathcal{G} = \log\left(2^{-2n} \det \bar{\Sigma}^{-1}\right)$.

Proof: The controllability Gramian for the double integrator dynamics is

$$\bar{\Sigma} = \frac{1}{2}\begin{bmatrix} I & 0 \\ 0 & \mathcal{A}_\mathcal{G}{}^{-1} \end{bmatrix}, \tag{12}$$

and a similar computation as from the previous proof yields the result that $S_\mathcal{G} = \left(2^{-2n} \det \bar{\Sigma}^{-1}\right)$. ■

The log determinant of the covariance matrix is proportional to the volume of the controllability ellipsoid. Recall that one interpretation of the controllability Gramian Σ is that it is the steady state covariance of $x(t)$ when the consensus dynamics is driven by white noise. Hence, when the steady state covariance of the state is small, the entropy is large, and vice versa. In this sense, a more disordered network is more robust to noise. Recalling our result linking the entropy to the number of spanning trees, a high entropy also corresponds to a large number of spanning trees, or a large number of possible paths noise can take as it propagates throughout the network.

Intuitively, one can think of the leader-node symmetries, described by Rahmani *et. al* [10], as being noise amplifiers, while disorder and large numbers of spanning trees act to dampen noise. Hence, when designing network topologies, it is beneficial to add edges between nodes such that entropy is maximized. In Section V, we illustrate this with some examples.

Next, we will discuss solutions to the Lyapunov equation when $G \neq I$. Choose G to be $\sum_{i\in\mathcal{S}} e_i e_i^T$ for some set of indices $\mathcal{S}$ corresponding to follower nodes. To do this, we define the covariance entropy.

Definition 10 (Covariance Entropy): Let $(-L_\mathcal{G}^{\{n-1\}}, G)$ be as before, and a controllable pair. Denote $\Sigma_\mathcal{S}$ to be the solution of the Lyapunov equation $-L_\mathcal{G}^{\{n-1\}}\Sigma_\mathcal{S} - \Sigma_\mathcal{S} L_\mathcal{G}^{\{n-1\}} = -GG^T$. Define the covariance entropy to be $M_\mathcal{S} = \log 2^n \det \Sigma_\mathcal{S}^{-1}$.

Clearly, when $G = I$, the covariance entropy is equal to the loop entropy. For a certain class of networks, the covariance entropy has a natural ordering property.

Theorem 11 (Covariance Entropy Ordering): Construct $L_\mathcal{G}^{\{n-1\}}$ as before, but with the stipulation that $(-L_\mathcal{G}^{\{n-1\}}, G)$ be controllable for any $G = \sum_{i\in\mathcal{S}} e_i e_i^T$ with indices $\mathcal{S}$ corresponding to follower nodes. Let $M_\mathcal{S}$ and $M_{\mathcal{S}\cup\{k\}}$ denote the covariance entropies with input edge sets $\mathcal{S}$ and $\mathcal{S}\cup\{k\}$, respectively. Then, $M_\mathcal{S} \geq M_{\mathcal{S}\cup\{k\}}$. In particular, the loop entropy minimizes the covariance entropy.

Proof: As before, choose G to be $\sum_{i\in\mathcal{S}} e_i e_i^T$ for some set of indices $\mathcal{S}$. Since $-L_\mathcal{G}^{\{n-1\}}$ is stable, it follows that the solution to the Lyapunov equation

$$-L_\mathcal{G}^{\{n-1\}}\Sigma_{\{i\}} - \Sigma_{\{i\}} L_\mathcal{G}^{\{n-1\}} = -GG^T \tag{13}$$

is given by $\Sigma_{\{i\}} = \int_0^\infty e^{-tL_\mathcal{G}^{\{n-1\}}} e_i e_i^T e_i e_i^T e^{-tL_\mathcal{G}^{\{n-1\}}} dt$ and is positive semi-definite and so $\Sigma_{\mathcal{S}\cup\{k\}} - \Sigma_\mathcal{S} = \Sigma_{\{k\}}$ is positive semi-definite. It follows that $\Sigma_{\mathcal{S}\cup\{k\}} \succeq \Sigma_\mathcal{S}$. Lastly, since the system is controllable, $\log 2^n \Sigma_\mathcal{S}^{-1} \succeq \log 2^n \Sigma_{\mathcal{S}\cup\{k\}}^{-1}$. ■

IV. Kolmogorov-Sinai Entropy

In this section, we examine the Kolmogorov-Sinai entropy that is related to the adjacency matrix spectrum. We will also compute this entropy for a variety of simple graphs, which will elucidate the topological features it capture. Using results from [14], we will derive bounds on the value of this entropy and show that the bound is sharp for both regular and highly-irregular graphs. At the end of this section, we will show the relation between the Kolmogorov-Sinai entropy and the stability of adjacency driven networks.

The Kolmogorov-Sinai entropy is a generalization of information-theoretic notion of the Shannon entropy on a network [15]. It characterizes the rate at which information is generated by the network. In the context of a network, information corresponds to a sequence of nodes visited by some Markov process on the network. The Kolmogorov-Sinai entropy is invariant under transformations that preserve the frequencies with which the network generates time-ordered sequences of nodes.

The stochastic process defining the information source is given by a Markov matrix M satisfying the properties

$$p_{ij} \geq 0,\ \ \textstyle\sum_j p_{ij} = 1,\ \text{and}\ \pi = \pi P \tag{14}$$

where π is the chain's stationary distribution. Consider the set of stochastic matrices M_A, where A is the adjacency matrix of the network graph, that have the property $a_{ij} = 0$ if and only if $p_{ij} = 0$. Let μ_n be the dominant eigenvalue of A with eigenvector v. Let M_A be a family of stochastic matrices with the property that $p_{ij} = 0$ if and only if $a_{ij} = 0$ in the network adjacency matrix. The *Kolmogorov-Sinai Entropy* defined presently is a measure that satisfies a variational principle analogous to that of the Gibbs principle in statistical mechanics [16]:

$$\log \mu_n = \sup_{P \in M_A} \left[- \sum_{ij} \pi_i p_{ij} \log p_{ij} + \sum_{ij} \pi_i p_{ij} \log a_{ij} \right].$$

The supremum over all admissible stochastic matrices is given by the unique matrix P satisfying $p^*_{ij} = \frac{a_{ij} v_j}{\mu_n v_i}$, and hence we have

$$\log \mu_n = - \sum_{ij} \pi_i p^*_{ij} \log p^*_{ij} + \sum_{ij} \pi_i p^*_{ij} \log a_{ij}. \quad (15)$$

By convention, $0 \log 0 = 0$. If the adjacency matrix is unweighted, in the sense that $a_{ij} = 1$ if there is an edge at $\{i, j\}$ and $a_{ij} = 0$ otherwise, then the term on the right hand side of (15) is zero. The term remaining is called the dynamical entropy of the Markov process, and is given by

$$H_P = - \sum_{i=1}^N \pi_i \sum_j p^*_{ij} \log p^*_{ij}. \quad (16)$$

The Kolmogorov-Sinai Entropy is defined as follows [15]:

Definition 12 (Kolmogorov-Sinai Entropy): The Kolmogorov-Sinai Entropy of a weighted adjacency matrix A is defined by

$$H = \log \mu_n - \sum_{ij} \pi_i p^*_{ij} \log a_{ij}, \quad (17)$$

where $p^*_{ij} = \frac{a_{ij} v_j}{\mu_n v_i}$.

For an unweighted graph, the right-hand side vanishes:

$$H = \log \mu_n. \quad (18)$$

In the next part of this section, we will give some bounds on the value of the KS entropy. For the remainder of the paper, we will assume unweighted adjacency matrices, and hence the KS-entropy is given by $H = \log \mu_n$.

A. Bounds on the Kolmogorov-Sinai Entropy

The Kolmogorov-Sinai entropy is difficult to compute in closed form for an arbitrary graph. Using results by Nikiforov [14], we can compute upper bounds on the Kolmogorov-Sinai entropy which is given in terms of topological quantities such as edge count and minimum degree. We can now recall the two theorems by Nikiforov.

Definition 13 (K_{p+1}-Free Graphs): The $\mathcal{G}$ is called $p+1$-*free* if it does not contain a complete subgraph on $p+1$ vertices.

Theorem 14 (Edge Bound): Let $\mathcal{G}$ be $p+1$-free. Then, the largest eigenvalue μ_n of the adjacency matrix of $\mathcal{G}$ satisfies $\mu_n \leq \sqrt{2 \frac{p-1}{p} |E|}$.

Theorem 15 (Minimal Degree Bound): Let $\mathcal{G}$ be $p+1$-free with minimal degree δ. Then, the largest eigenvalue μ_n of the adjacency matrix of $\mathcal{G}$ satisfies

$$\mu_n \leq \frac{\delta - 1}{2} + \sqrt{2|E| - n\delta + \frac{(\delta+1)^2}{4}}. \quad (19)$$

By the definition of the Kolmogorov-Sinai entropy, we immediately have the following theorem:

Theorem 16 (Bounds on KS-Entropy): Let $\mathcal{G}$ be $p+1$-free with minimal degree δ. Then, the Kolmogorov-Sinai entropy satisfies the two inequalities:

$$H \leq \frac{1}{2} \log \left[2 \frac{p-1}{p} |E| \right] \quad (20)$$

$$H \leq \log \left[\frac{\delta - 1}{2} + \sqrt{2|E| - n\delta + \frac{(\delta+1)^2}{4}} \right]. \quad (21)$$

Equation (21) is in fact tight for regular graphs, and also highly irregular graphs in the sense of maximizing the degree variance. For example, let $n \geq k$, and let $H_{n,k}$ be the n-node complement of the complete graph on $n-k$ nodes K_{n-k}. Then, the maximal adjacency eigenvalue of $H_{n,k}$ is $\mu_n(H_{n,k}) = \frac{k-1}{2} + \sqrt{\frac{(k-1)^2}{4} + k(n-k)}$.

Now note that $\delta = k$ and $2|E| = kn - k^2 - k$ and so we get equality in (21). A result by Bell [17] states that $H_{n,k}$ maximizes the degree variance over all n-vertex graphs with the same edge count as $H_{n,k}$, where the degree variance is given by $V = \frac{1}{n} \sum_{i=1}^n \left(d_i - 2\frac{|E|}{n} \right)^2$. Hence graphs that maximize the KS-entropy include the regular graphs and the highly irregular graphs (namely the ones that maximize the degree variance over graphs with fixed vertex and edge counts).

A lower bound on the maximum eigenvalue follows from the Perron-Frobenius theorem. Let $\langle d \rangle$ denote the average degree of the graph and let $d_{\max}$ denote its largest degree. Then, the following inequality holds [18], [19]:

$$\max\{\langle d \rangle, \sqrt{d_{\max}}\} \leq \mu_n \leq d_{\max}. \quad (22)$$

Hence, we can bound the entropy by

$$\max\{\log \langle d \rangle, 0.5 \log d_{\max}\} \leq H \leq \log d_{\max}. \quad (23)$$

Therefore increasing the average degree (say, by adding edges to every vertex) or increasing the maximum degree (adding many edges to a single vertex) will drive the entropy to increase from below.

B. Kolmogorov-Sinai Entropy of Adjacency-driven Networks

The system-theoretic interpretation of the Kolmogorov-Sinai entropy relates to the adjacency dynamics (5). Consider the dynamics where each node is driven by its neighbors as $\dot{x}_i = \sum_{\{i,j\} \in E} x_j$ on a connected network. In the matrix form, this is exactly the adjacency dynamics: $\dot{x} = Ax$.

Since $\mathrm{Tr} A = 0$ this dynamics is unstable as some eigenvalues are nonnegative. One way of stabilizing the system is to add a self-loop with positive weights to each node that drives the system to stability, as show in Figure 2. This corresponds to the dynamics $\dot{x}_i = \sum_{\{i,j\} \in E} x_j - \delta x_i$, which in matrix form is $\dot{x} = (A - \delta I)x$. This is equivalent to taking $B = I$ in the controlled adjacency dynamics (5) in with a feedback gain of $u = -\delta x$. The minimum value that δ has to take in order to drive the system to stability is precisely the largest eigenvalue of A. Hence, the system is stable for any $\epsilon > 0$ when $\dot{x} = (A - (\epsilon + \mu_n)I)\, x = \left(A - (\epsilon + e^H)I\right) x$.

V. Examples of Entropies for Simple Graphs

In Table I, we show some examples of the loop entropy for some simple graphs.

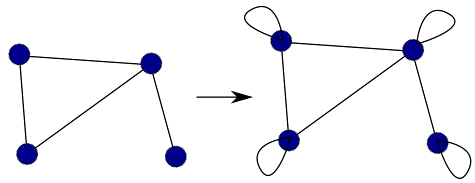

Fig. 2. Adding self-loops to stabilize the adjacency dynamics.

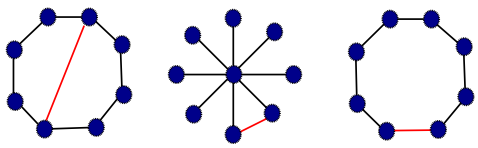

Fig. 3. **Left:** Adding a edge in a cycle increases entropy by adding spanning trees. **Middle:** Adding an edge to a star graph adds two spanning trees. **Right:** Creating a cycle out of a path adds $n-1$ spanning trees.

In Figure 3, we show how adding edges increases the entropy of some example graphs. A cycle of n nodes has n spanning trees since removing any one of the n edges creates a path on n nodes and hence a spanning tree. Turning a path into a cycle greatly increases entropy as one goes from one spanning tree to n spanning trees. Clearly, creating larger cycles in a network increasing entropy. In the star graph, connecting any two non-center nodes induces a 3-cycle, which increases the number of spanning trees by two. The star graph is already low-entropy, and it is hence robust against edge perturbations in the sense that adding edges will not significantly increase the number of spanning trees, say compared to joining the ends of a path graph.

TABLE I

EXAMPLES OF LOOP ENTROPIES

Graph	Spectrum of $L_{\mathcal{G}}$	Entropy $S_{\mathcal{G}}$
Cycle	$\left\{4\sin^2\left(\frac{\pi k}{2n}\right)\right\}$	$2\sum_{k=1}^{n-1}\log\left[2\sin\left(\frac{\pi k}{2n}\right)\right]$
Complete	$\{n, 0\}$	$(n-1)\log n$
Path	$\left\{4\sin^2\left(\frac{\pi k}{4n}\right)\right\}$	$\sum_{k=1}^{n-1}\log\left[2\sin\left(\frac{\pi k}{4n}\right)\right]$
Star	$\{0, 1, n\}$	$\log n$

Assuming an unweighted graph, we can write the Kolmogorov-Sinai entropy as $H = \log \mu_n$, where A is the adjacency matrix of the graph. Using results from the summary article by Brouwer *et al.* [20], we can write the entropy of some graph families, as shown in Table II.

TABLE II

EXAMPLES OF KS ENTROPIES

Graph	Spectrum of A	Entropy H
Cycle	$\left\{2\cos\left(\frac{2\pi j}{n}\right)\right\}$	$\log 2$
Complete	$\{n-1, -1\}$	$\log(n-1)$
Path	$\left\{2\cos\left(\frac{2\pi j}{n+1}\right)\right\}$	$\log\left[2\cos\left(\frac{\pi}{n+1}\right)\right]$
Star	$\{-\sqrt{n-1}, 0, \sqrt{n-1}\}$	$\log\sqrt{n-1}$
Regular	$\lvert\mu(A)\rvert \leq k$	$\log k$
Complete Bipartite	$\{-\sqrt{nm}, 0, \sqrt{nm}\}$	$\log\sqrt{nm}$

VI. Conclusion

In this paper, we examined how two notions of network entropy, namely loop entropy and Kolmogorov-Sinai entropy, provide useful measures for noise propagation and stability of networked systems. Along the way, we explored various bounds on these measures in relation to the network structure, and provided insights into what these measures quantify. Our future works will extend this point of view to evolving networks, and examine how various notions of network entropy capture distinct facets of complex dynamics that networked systems often exhibit.

ACKNOWLEDGMENT

MHdB would like to thank Paul Liu for useful conversations.

References

[1] G. Robins, P. Pattison, Y. Kalish, and D. Lusher, "An introduction to exponential random graph (p*) models for social networks," *Social networks*, vol. 29, no. 2, pp. 173–191, 2007.

[2] V. Vedral, A. Barenco, and A. Ekert, "Quantum networks for elementary arithmetic operations," *Physical Review A*, vol. 54, no. 1, p. 147, 1996.

[3] G. Mahler and R. Wawer, "Quantum networks: dynamics of open nanostructures," *VLSI Design*, vol. 8, no. 1-4, pp. 191–196, 1998.

[4] J. D. Allen, Y. Xie, M. Chen, L. Girard, and G. Xiao, "Comparing statistical methods for constructing large scale gene networks," *PLOS One*, vol. 7, no. 1, p. e29348, 2012.

[5] A. Chapman and M. Mesbahi, "Semi-autonomous consensus: network measures and adaptive trees," *IEEE Transactions on Automatic Control*, vol. 58, no. 1, pp. 19–31, 2013.

[6] M. Siami and N. Motee, "Schur-convex robustness measures in dynamical networks," in *American Control Conference (ACC)*, 2014, pp. 5198–5203.

[7] M. Mesbahi and M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*. Princeton University Press, 2010.

[8] P. Barooah and J. P. Hespanha, "Graph effective resistance and distributed control: Spectral properties and applications," in *45th IEEE Conference on Decision and Control*, 2006, pp. 3479–3485.

[9] S. Salsa, *Partial Differential Equations in Action: From Modelling to Theory*. Springer Science & Business Media, 2008.

[10] A. Rahmani, M. Ji, M. Mesbahi, and M. Egerstedt, "Controllability of multi-agent systems from a graph-theoretic perspective," *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 162–186, 2009.

[11] G. S. Chirikjian, "Modeling loop entropy," *Methods in enzymology*, vol. 487, p. 99, 2011.

[12] L. Boltzmann, *Lectures on Gas Theory*. Dover Publications, 2011.

[13] C. A. Desoer, "The optimum formula for the gain of a flow graph or a simple derivation of Coates' formula," *Proceedings of the IRE*, vol. 48, no. 5, pp. 883–889, 1960.

[14] V. Nikiforov, "Some inequalities for the largest eigenvalue of a graph," *Combinatorics, Probability & Computing*, vol. 11, no. 2, pp. 179–189, 2002.

[15] L. Demetrius and T. Manke, "Robustness and network evolution - an entropic principle," *Physica A: Statistical Mechanics and its Applications*, vol. 346, no. 3, pp. 682–696, 2005.

[16] L. Arnold, V. M. Gundlach, and L. Demetrius, "Evolutionary formalism for products of positive random matrices," *The Annals of Applied Probability*, pp. 859–901, 1994.

[17] F. K. Bell, "A note on the irregularity of graphs," *Linear Algebra and its Applications*, vol. 161, pp. 45–54, 1992.

[18] C. D. Godsil and G. Royle, *Algebraic Graph Theory*. Springer New York, 2001, vol. 207.

[19] D. Cvetković, P. Rowlinson, and S. Simić, "An introduction to the theory of graph spectra," *Cambridge-New York*, 2010.

[20] A. E. Brouwer and W. H. Haemers, *Spectra of Graphs*. Springer Science & Business Media, 2011.