12-11-2021

Entropy analysis of Boolean network reduction according to the determinative power of nodes

Matthew J. Pelz

Mihaela T. Velcsov

ENTROPY ANALYSIS OF BOOLEAN NETWORK REDUCTION ACCORDING TO THE DETERMINATIVE POWER OF NODES

MATTHEW PELZ AND MIHAELA T. MATACHE

Department of Mathematics, University of Nebraska at Omaha
Omaha, NE 68182-0234, USA
mpelz@unomaha.edu, dmatache@unomaha.edu

Abstract. Boolean networks are utilized to model systems in a variety of disciplines. The complexity of the systems under exploration often necessitates the construction of model networks with large numbers of nodes and unwieldy state spaces. A recently developed, entropy-based method for measuring the determinative power of each node offers a new method for identifying the most relevant nodes to include in subnetworks that may facilitate analysis of the parent network. We develop a determinative-power-based reduction algorithm and deploy it on 36 network types constructed through various combinations of settings with regards to the connectivity, topology, and functionality of networks. We construct subnetworks by eliminating nodes one-by-one beginning with the least determinative node. We compare entropy ratios between these subnetworks and the parent network and find that, for all network types, the change in network entropies (sums of conditional node entropies) follows a concave down decreasing curve, and the slightest reductions in network entropy occur with the initial reductions which eliminate the nodes with the least determinative power. Comparing across the three network characteristics, we find trends in the rates of decrease in the entropy ratios. In general, the decline occurs more slowly in networks with degree values assigned from a power-law distribution and canalyzing functions of higher canalization depth. We compare results of the determinative-power-based reduction with those of a randomized reduction and find that, in forming subnetworks with maximal network entropy, the determinative-power-based method performs as well as or better than the random method in all cases. Lastly, we compare findings based on this conditional-entropy-based calculation of network entropy with those of an alternative calculation using simple sums of (independent) node entropies to demonstrate the vast differences resulting from the two approaches.

Keywords: Boolean networks, biological information theory, mutual information, determinative power, Shannon entropy, network reduction, numerical simulations

1. Introduction

The application of mathematical network modeling techniques in the analysis of dynamical real-world systems has expanded with the growth of computational capacity. The benefits of Boolean network modeling and other mathematical approaches are enticing. Such models are more contained and often cleaner, safer, and less expensive to implement than physical experiments. Yet the reliance on network models introduces a couple of substantial challenges. First, informed network initialization in terms of connectivity, topology, and functionality is crucial. Inappropriate choices with regards to any of these settings could render a simulation irrelevant. Second, despite advances in computational power, large networks continue to push the limits of our capacities. Consequently, the identification of methods for simplifying networks while retaining their essential characteristics is a pertinent area of research. In this study, we explore a method for reducing large networks to their most relevant nodes. We demonstrate a reduction algorithm based on a previously developed measure termed *determinative power* and deploy the algorithm on a variety of network types to evaluate impacts under various conditions.

Researchers in multiple disciplines have found Boolean networks to be useful in modeling complex systems. Developed most famously through the work of Stuart Kauffman [1], Boolean networks feature interconnected nodes, which take on binary values such as *0* or *1*, *on* or *off*, or *true* or *false*, along with rules that define how nodes process information from one or more inputs upon repeated iterations. **The definitions in Sections 1 and 2 are taken from Pentzien et al. [9].**

Definition 1. *A **Boolean network** includes a set $\{X_1, X_2, \ldots, X_N\}$ of N nodes and a set $\{f_1, f_2, \ldots, f_N\}$ of N functions with one or more inputs from among the N nodes, such that each node X_i takes a binary value based on the function f_i governing its behavior upon iteration. If $X_i = 0$ the node is considered to be* off *and if $X_i = 1$ the node is considered to be* on*.*

Definition 2. *Each node X_i of a Boolean network has a set $\{X_{i_1}, X_{i_2}, \ldots, X_{i_K}\}, 1 \leq i_1 < i_2 < \ldots < i_K \leq N$ of K associated inputs. The value of K is the node's **connectivity**.*

Then, for each $i \in \{1, 2, \ldots, N\}$, the node X_i is assigned the Boolean function $f_i(X_{i_1}, X_{i_2}, \ldots, X_{i_K})$. Assuming synchronous updating for all network nodes, i.e. all

nodes update at the same time, then for each node X_i and iteration or time step t, $X_i(t+1) = f_i(X_{i_1}(t), X_{i_2}(t), \ldots, X_{i_K}(t))$, using the notation $X_i(t)$ to denote the state of the node X_i at time t. Then the vector $(X_1(t), X_2(t), \ldots, X_N(t))$ represents the state of the network at time t.

Boolean approaches have been adopted in modeling genetic regulatory networks [2–5], signal transduction in cells [6], and neural networks [7]. In a study of the gene expression network of the *Drosophila* embryo, Albert et al. [8] found that Boolean models represented the real-world network as accurately as models that permitted more complicated, non-binary values.

The study of Boolean networks tends to focus on behavioral patterns in four network characteristics. First, the size of the network, designated by the number of nodes N, provides an upper limit on its scale and overall complexity. Second, the connectivity K has substantial impacts on the dynamics of the network, particularly in regards to the periodicity and stability of orbits. The value of K can be fixed so that each node has the same number of inputs or it can represent the average of a variable connectivity. A third network characteristic, topology, is especially important for networks in which K varies. Each node has associated *in-degree* and *out-degree* attributes describing the number of inputs the node receives (in) and the frequency with which it serves as an input for other nodes (out). The topology of a network is described by the distributions of degree values, and the differences among these distributions substantially impact the network dynamics. The fourth characteristic governing network behavior is functionality. Boolean functions may be homogeneous across all nodes of the network, meaning that all nodes obey the same function, or they may vary.

One recent project modeling signal transduction in a fibroblast cell exemplifies the computational burdens that emerge in large Boolean networks [9]. The model incorporates 130 nodes and, consequently, the state space has 2^{130} configurations. Every configuration could serve as a potentially interesting set of initial conditions, one which we might like to investigate by iterating the network hundreds or even thousands of times. Characterization of the behavior of this network according to associated mathematical functions would benefit from an investigation of the truth tables of the nodes, which unfortunately grow in size exponentially as in-degree values increase. As researchers strive to make models as realistic

as possible, network sizes and maximum in-degree values are certain to increase. We can easily foresee a 1000-node model in which several nodes have in-degrees of at least 20 and truth tables with over a million entries each.

Researchers manage the challenges inherent in large, complex network models by utilizing analytical and simulative tools that do not require a comprehensive accounting of network dynamics for every configuration in the state space. Multiple strategies involve the reduction of the network through the elimination of nodes. With each node removed from the parent network, the size of the state space of the resulting subnetwork is halved. In this context, the development of criteria for selecting nodes for elimination is a central concern. One approach relies on the identification and elimination of *frozen* nodes, those that reach a fixed value and remain there for any number of network iterations, along with nodes that serve as inputs only for frozen nodes [10–12]. Related strategies involve the removal of *leaf* nodes, those with zero outputs, and *mediator* nodes, those with exactly one input and one output [13–15]. The common goal behind these approaches is the summarization of a network according to its attractors and stable states [16].

A second, complementary strategy for network reduction emphasizes the role that nodes play in information propagation. This strategy also has ramifications for network stability as research has supported the conjecture that information propagation and stability are intertwined [17]. Heckel et al. [19] built upon work by Ribeiro [20] to develop a method for quantifying each node's determinative power based on measures of Shannon entropy. Matache and Matache [21] used this method to analytically show that knowledge of the states of nodes with the highest determinative power reduces the entropy of the parent network significantly. Other research projects [22, 23] have focused on the reducibility of networks not restricted to Boolean structures. In addition, the entropy of the relevant components of the network which are comprised of relevant nodes that eventually influence each other's state was used as a measure of uncertainty of the future behavior of a random state of the network by Krawitz and Shmulevich [24, 25]. Shreim et al. [26] suggest that the basin entropy may be a relevant and robust signifier of criticality, irrespective of the specific dynamics. The authors of [26] find that both classical random Boolean networks and asynchronous random Boolean networks exhibit a basin entropy that increases with system size only for critical networks. This suggests that superuniversal features exist for

the fluctuations in the structure of the state space of random Boolean networks, and these are invariant with respect to the specific dynamics.

In the present study, we explore relationships between information and subnetwork size for various network types. If we generate subnetworks by eliminating nodes with the least determinative power, what are the ratios between the network entropies of the subnetworks and those of the parent networks? We will label these ratios "sub-to-parent entropy ratios." As we eliminate nodes with the least determinative power, generating smaller and smaller subnetworks, how rapidly do sub-to-parent entropy ratios decline for different types of networks? If the ratios are declining gradually then the subnetwork can capture most of the variability intrinsic to the parent network.

To explore these questions, we implement the determinative-power-based reduction algorithm of [9] on 36 different network types. We compare results with respect to each of the three network characteristics, connectivity, topology, and functionality, and identify trends suggestive of patterns in information retention. We compare sub-to-parent entropy ratios resulting from this determinative-power-based reduction algorithm with similar ratios resulting from a randomized reduction process. We find that, in all cases, the slightest decreases in sub-to-parent entropy ratios occur with the first node eliminations and that the entropy ratios decrease more slowly when using the determinative-power-based reduction method, thus preserving better informational and variability attributes of the network. In general, the decrease is most gradual in networks with topologies assigned by power-law distributions and outputs determined by canalyzing functions. The value of average connectivity K appears to be significant only in networks with topologies generated from Poisson distributions, in which cases lower K values are associated with more gradual decreases in sub-to-parent entropy ratios. Finally, we compare these findings, generated through a conditional-entropy-based calculation of network entropy, with an alternate calculation based on an aggregation of the node entropies within a network. We find that the shapes of the sub-to-parent entropy ratio curves change drastically between the two methods, suggesting that the choice of entropy measurement is pivotal in assessing the extent to which node reduction impacts information measures.

The paper is structured as follows. In Section 2 we review the mathematical background for the determinative power of nodes. Then, in Section 3, we describe the choices regarding

the network characteristics. The construction of the subnetworks and computational details are provided in Section 4, which is followed by the results of the numerical simulations in Section 5. We conclude with a summary of the findings and directions for further research in Section 6.

We find that for all network types, initial node reductions produce the smallest change in sub-to-parent network entropy. This result provides evidence that determinative power can be a useful tool in identifying subnetworks reflective of parent networks.

2. Determinative Power

The concept of determinative power quantifies the importance of each node in governing the state of the network as a whole. It offers an understanding of the extent to which a reduction in uncertainty of the state of a particular node reduces the uncertainty of the state of the network. Foundations of the determinative power measure date to 2008, when Ribeiro et al. [20] explored how concepts from the field of information theory could be employed in the calculation of each node's contributions to uncertainty in the network. Their work focused on *Shannon entropy*, referred to here simply as *entropy*, along with related concepts *conditional entropy* and *mutual information*.

Definition 3. *Denote as X a discrete random variable and as $p(x) = P(X = x)$ its associated probability mass function. In the general case, the* ***entropy*** *of X is defined as*

$$H(X) = -\sum_{x} p(x) \log_2 p(x)$$

If we confine our consideration of entropy to its application to Boolean networks, where X_i represents a node that can take only binary values, we can offer a more precise definition:

$$H(X_i) = -p(0) \log_2 p(0) - p(1) \log_2 p(1)$$

If $p(0) = 0$ or 1, then $H(X_i) = 0$ by definition.

Definition 4. *Consider two discrete random variables X and Y. The* ***conditional entropy*** *of Y given knowledge of X is defined as*

$$H(Y|X) = -E\left[\log_2 P(Y|X)\right] = -\sum_{x,y} P(Y = y, X = x) \log_2 P(Y = y|X = x).$$

Using knowledge of the conditional entropies, we can calculate the reduction in the uncertainty of nodes given knowledge of the states of their inputs.

Definition 5. *The **mutual information** (MI) of a random variable Y with respect to another random variable X is the reduction in the uncertainty of Y given knowledge of X.*

$$MI(Y;X) = H(Y) - H(Y|X)$$

For example, if X is the only input for Y, then

$$MI(Y;X) = H(Y).$$

If X is not an input for Y, then

$$MI(Y;X) = 0.$$

In principle, mutual information is a measure of the "gain of information", or the determinative power of X over Y. Mutual information is useful in the analysis of real-world complex networks, as demonstrated by a 2013 study of a siesmicity model. [18]

In 2013, Heckel et al. [19] adapted the work or Ribeiro et al. [20] on entropy, conditional entropy, and mutual information to define determinative power as a measure of the mutual information of X with respect to all nodes of the network.

Definition 6. *The **determinative power** (DP) of a node X_j is the reduction of uncertainty of the network given knowledge of the state of X_j. More precisely, the DP of the node X_j is given by a summation of the MI over all outputs of node X_j, namely*

$$DP(j) = \sum_{i=1}^{n} MI(X_i; X_j).$$

Heckel et al. [19] applied determinative power analysis to the regulatory network of *E. coli* cells and found that knowledge of a small subset of nodes with the highest determinative power reduces the uncertainty and the size of the state space significantly. Matache and Matache [21] confirmed the general finding analytically. In 2018 Pentzien et al. [9] examined 36 models created in the *Cell Collective* network modeling platform (www.cellcollective.org, [27, 28]). They calculated determinative power and performed biological function analysis for each network. The authors found that "a large fraction of the most determinative nodes are essential and involved in crucial biological functions." The finding established

a connection between theoretical determinative power computations and the behavior of real-world networks.

3. Network Types

The behavior of networks depends to a great extent on the interactions of three characteristics: connectivity, topology, and functionality. In studying random Boolean networks, in which functionality and topology are randomly generated to approximate real-world phenomena, Kauffman [29] found that differences in the values of K have substantial impacts on network stability. Wherever $3 \leq K \leq N$, networks exhibit chaotic behavior. Near $K = 2$, networks exhibit some order and demonstrate robustness against perturbations. Such networks are considered to exist in the *critical* phase. Due in part to their robustness, biological systems are believed to operate in the critical phase [17]. In the current study, we explore networks with K values equaling or approximating 2, 3, and 4. In each case, K represents the average of a variable connectivity rather than a fixed degree value. Examination of these three values is sufficient to identify a possible trend in how choices of K impact sub-to-parent entropy ratios.

The topology of a network–the arrangement of its connections–can vary widely in different network types. Previous research has identified and applied three topologies useful in creating Boolean models of real-world systems. The Watts-Strogatz or "small-world" model incorporates the localized clustering that tends to occur in social networks [30]. The Barabási-Albert or "scale-free" model mimics the occurrence of "hubs"–nodes with very high degree values–that characterizes many real networks including the world wide web [31, 32]. A third topology, the Holme-Kim model, combines elements of the small-world and scale-free models [33]. These three topologies are defined primarily by specific network-wiring algorithms and by the resulting distribution of degree values. The Watts-Strogatz model yields degree values with a Poisson distribution while the Barabási-Albert and Holme-Kim models share a power-law distribution.

Recently, Wacker et al. [34] provided an in-depth correlation analysis between determinative power and two topological measures, the clustering coefficient and the betweenness centrality of nodes. The authors found that for smaller size networks, determinative power seems to be strongly correlated to the betweenness centrality and weakly or very weakly

correlated to the clustering coefficient, depending on the underlying topology. The Barabási-Albert and Holme-Kim topologies lead to fairly similar results, while the Watts-Strogatz networks generate a different type of behavior. Apparently, the existence of hubs in the network may be more important than the clustering of the nodes as far as determinative power is concerned. The work in [34] has inspired the choices of topologies in this research.

To explore a broad variety of network types we assign degree values from uniform, Poisson, or power-law distributions. The latter two cases serve as approximations of the Watts-Strogatz model (Poisson) and the Barabási-Albert and Holme-Kim models (power-law). All networks constructed and examined are *directed*, meaning that the input-output relationships between nodes are not necessarily reciprocal. To simplify the network initialization process, we stipulate that the in-degree value must equal the out-degree value for each node. For example, if a node has five inputs it must have five outputs. Assignments of the inputs and outputs are performed by separate random processes. We also stipulate that each node has at least one input and that self-inputs are not permissible.

With regards to functionality, research indicates that two categories of unate functions are common in naturally occurring systems [19]. First, canalyzing functions, in which the value of one particular input can determine the output regardless of the values of other inputs, are believed to be essential in biological systems [29]. Nested canalyzing functions (NCFs), a subcategory of functions in which multiple inputs have the potential to determine the output according to their position in an input hierarchy, appear to be particularly common [35,36]. In NCFs, the depth d of canalyzation is a distinguishing parameter. For example, consider a node with four inputs. If two of those inputs are canalyzing, then $d = 2$ and the function is considered to be *partially nested canalyzing*. Alternatively, if $d = 4$, then the function is *fully nested canalyzing*.

Research has found that while network stability increases as d increases, only a few degrees of depth are necessary to mimic the behavior in real-world systems [35, 36]. With this in mind, our study explores canalyzing functions of depths 1 and 2. In our $d = 1$ function, one canalyzing input in an *on* position turns the output node *on* in the next iteration. Where $d = 2$, the function is nested canalyzing. A first canalyzing input in the *on* position turns the output node *on*. If the first input is *off*, a second canalyzing input in the *on* position turns the output node *on*. If the number of inputs is greater than d

and no canalyzing inputs are in the *on* position, then our function determines the output value according to an unbiased random function analagous to a coin flip. These choices simplify our approach and make computations manageable. Future research will consider other types of canalyzing functions.

A second function type explored here concerns thresholds or "switches." Wittman et al. [37] argue that threshold functions are the "least common denominator" of biologically meaningful update rules. These functions toggle the output if the combined input values reach a specified threshold. A function may have both canalyzing and threshold characteristics. For instance, consider a node that has two inputs, either of which will toggle an output of *on* if its input is *on*. The function associated with this node can be seen as having both canalyzing depth of 2 and as having a threshold of 0.5. There are other cases in which the principles of canalyzation and thresholds diverge. Inputs in threshold functions do not strictly adhere to a hierarchy that characterizes canalyzing functions. The present study explores two functions with threshold values of 0.5 and 0.25. If the proportion of a node's inputs in the *on* position is above the threshold, the node turns to *on* in the next iteration.

Our study includes 36 networks featuring every combination of the parameter settings outlined above, including three K values, three topologies, and four functions. Because the total number of network types increases exponentially with each additional parameter setting, we limit the choice of settings to include only those most likely to reveal suggestive patterns. We hold network size N constant at 100 across all cases.

4. Network Models and Reduction Algorithm

For all networks listed in Tables 1-A and 1-B, network initialization, reduction, calculation, and visualization was completed using the R programming language.[1] In this section, we provide a step-by-step explanation of the initialization, calculation, and reduction procedures. The visualization and discussion of the results are included in Section 5.

4.1. **Initialization.** The initialization process is consistent across all 36 network types. For each type, we fix parameter settings and create 50 networks of size $N = 100$ and each network is wired randomly per the specified topology setting. A summary matrix with 50

[1]In addition to base R, the project utilized the following packages: actuar, BoolNet, data.table, dplyr, ggplot2, permute, poisson, poweRlaw. reshape2, tibble, and tidyr

columns and 100 rows is associated with each network type. The summary matrix stores the sub-to-parent entropy ratios calculated after each node elimination.

The entropy and mutual information calculations performed during the project rely on data found in the truth tables that define functional logic. To facilitate simultaneous calculations across all nodes of the network, truth tables for each node are "stacked" to form a single truth table matrix. If we denote as m the largest degree value of any node, then the

TABLE 1-A. The first half of the list of 36 studied network types. The network types represent all combinations for three selected settings for K, three selected topologies, and four selected functions (two canalyzing and two threshold).

Network Type	**Average K**	**Topology Distribution**	**Function**
1	2	uniform	threshold = 0.5
2	2	uniform	threshold = 0.25
3	2	uniform	canalyzing d = 2
4	2	uniform	canalyzing d = 1
5	2	power-law	threshold = 0.5
6	2	power-law	threshold = 0.25
7	2	power-law	canalyzing d = 2
8	2	power-law	canalyzing d = 1
9	2	Poisson	threshold = 0.5
10	2	Poisson	threshold = 0.25
11	2	Poisson	canalyzing d = 2
12	2	Poisson	canalyzing d = 1
13	3	uniform	threshold = 0.5
14	3	uniform	threshold = 0.25
15	3	uniform	canalyzing d = 2
16	3	uniform	canalyzing d = 1
17	3	power-law	threshold = 0.5
18	3	power-law	threshold = 0.25

TABLE 1-B. The second half of the list of 36 studied network types. The network types represent all combinations for three selected settings for K, three selected topologies, and four selected functions (two canalyzing and two threshold).

Network Type	Average K	Topology Distribution	Function
19	3	power-law	canalyzing d = 2
20	3	power-law	canalyzing d = 1
21	3	Poisson	threshold = 0.5
22	3	Poisson	threshold = 0.25
23	3	Poisson	canalyzing d = 2
24	3	Poisson	canalyzing d = 1
25	4	uniform	threshold = 0.5
26	4	uniform	threshold = 0.25
27	4	uniform	canalyzing d = 2
28	4	uniform	canalyzing d = 1
29	4	power-law	threshold = 0.5
30	4	power-law	threshold = 0.25
31	4	power-law	canalyzing d = 2
32	4	power-law	canalyzing d = 1
33	4	Poisson	threshold = 0.5
34	4	Poisson	threshold = 0.25
35	4	Poisson	canalyzing d = 2
36	4	Poisson	canalyzing d = 1

truth table matrix will have $m + 2$ columns: one column identifying the node by index (1, 2,...,100), m columns specifying the values of various input nodes, and one column defining the output based on the setting of the function parameter. By stacking truth tables in this manner, we technically associate with each node an m x m truth table and, consequently, the stacked truth table matrix has m x 100 rows. However, m is the *maximum* degree value and the nodes with degree values (connectivities) less than m do not use a truth table of

this size. If we denote as n the degree value for any node such that $n \leq m$, then for each truth table matrix column associated with input indices $n+1, n+2, ..., m$, an input value of NA is assigned in all rows. The same procedure is performed in rows $n+1, n+2, ..., m$ so that each node has non-NA input values only in its associated n x n truth table.

In each of the four function types explored in the project, values in each row of the output column are generated through calculations involving input values located in the same row. For example, in the case of the 0.25 threshold function, if the mean of the input values in a row is greater than or equal to 0.25, the output column in that row is set equal to 1. All calculations ignore NA entries. We describe in detail this initialization method since it is preferred for its facilitation of iterative calculations and summations of conditional entropies across entire networks.

4.2. **Entropy calculations.** By stacking truth tables for all nodes into a single truth table matrix, we simplify the administration of the formulas presented in section 2 to produce determinative power measurements. To find mutual information for any relevant node pairing $\{X_i; X_j\}$, where X_j is an input to X_i, we need to know six probabilities, each of which is readily available via the stacked truth table matrix.

$$P(X_i = 1) \qquad \text{and} \qquad P(X_i = 0)$$

$$P(X_i = 1|X_j = 0) \qquad \text{and} \qquad P(X_i = 0|X_j = 0)$$

$$P(X_i = 1|X_j = 1) \qquad \text{and} \qquad P(X_i = 0|X_j = 1)$$

Of course, not all of these values need to be computed from the truth tables due to the basic rules of probability.

We also assume that the networks are ergodic, that is all states of the network are equally likely. Thus, any initial state of the network used in simulations has the same probability of being chosen from the collection of all possible states. To find the mutual information quantities used in the determinative power we employ the following formula from [21] and [9]:

$$MI(X_i; X_j)$$

$$= h\left(\sum_{x \in \text{supp}\, f_i} p(x)\right) - P(X_j = 1)h\left(\sum_{x \in \text{supp}\, f_i} P(X = x|X_j = 1)\right)$$

$$-P(X_j = 0)h\left(\sum_{x\in \operatorname{supp} f_i} P(X = x|X_j = 0)\right)$$

where $X = (X_1, X_2, \ldots, X_N)$ with states $x = (x_1, x_2, \ldots, x_N) \in \{0, 1\}^N$, and $\operatorname{supp} f_i = \{x : f_i(x) = 1\}$. Although $f_i(x)$ in this formula is considered a function of the state of the entire network, only the relevant inputs are actually of interest in the computations, namely $\{x_{i_1}, x_{i_2}, \ldots, x_{i_K}\}, 1 \leq i_1 < i_2 < \ldots < i_K \leq N$ as in Definition 2.

A companion matrix contains topology information linking the input indices in the truth table matrix to the input nodes designated randomly according to the distribution parameter setting. Using the companion matrix, we calculate the determinative power of X_j by summing the mutual information $MI(X_i; X_j)$ across all i. We store the determinative power for each node and arrange them in ascending order so that the most determinative node is in the 100th row.

Before initializing the node reduction process, we calculate the network entropy of the parent network by adapting the chain rule for entropy, provided by Cover and Thomas [38] as $H(X_1, X_2, ..., X_n) = \sum_{i=1}^{N} H(X_i|X_{i-1}, ..., X_1)$. The authors in [38] show that "the entropy of a collection of random variables is the sum of conditional entropies." For the Boolean network context, in which the network nodes serve as the collection of random variables, we take into account the one-step conditional entropy corresponding to the one-step iteration of the network, or equivalently, to the Boolean functions associated with the nodes. We modify the formula as follows:

Definition 7. *The* ***network entropy*** *is given by*

$$H(X_1, X_2, ..., X_n) = \sum_{i=1}^{N}\sum_{j=1}^{N} H(X_i|X_j).$$

4.3. **Reduction.** Using a list of node indices ordered by determinative power, we reduce the network one node at a time beginning with the least determinative node. Denote the node to be removed as X_a. We "knock out" X_a, meaning that we freeze it in the *off* position, by redefining certain entries in the truth table. This approach follows the procedure used in *Cell Collective* for analysis of biological networks. First, we redefine the function associated with X_a so that the output is 0 in all cases. Next, wherever X_a serves as an input for another node, we define the output as NA because these cases are no longer possible.

TABLE 2. Examples of the node reduction process in simplified network. The stacked truth table matrix on the left defines the logic for each of three nodes in the network. In the truth table matrix on the right, X_2 is knocked out, meaning that all of its output values are set to 0 and wherever it serves as an input with value 1, the output is set to NA.

Node Index	**Input 1**	**Input 2**	**Output**
1	0	NA	0
1	1	NA	1
2	0	0	0
2	1	0	1
2	0	1	1
2	1	1	1
3	0	0	0
3	1	0	1
3	0	1	1
3	1	1	1

Node Index	**Input 1**	**Input 2**	**Output**
1	0	NA	0
1	1	NA	1
2	0	0	0
2	1	0	**0**
2	0	1	**0**
2	1	1	**0**
3	0	0	0
3	1	0	1
3	0	1	**NA**
3	1	1	**NA**

Table 2 demonstrates how this node reduction is represented in the stacked truth table matrix. In this example, a simplified network uses a 0.5 threshold function to define outputs. In a companion topology matrix (not shown), we declare that

$$X_1(t+1) = f_1(X_3(t)), \qquad X_2(t+1) = f_2(X_1(t), X_3(t)), \qquad X_3(t+1) = f_3(X_1(t), X_2(t)).$$

Notice in the truth table on the left that the Input 2 columns for X_1 are defined as NA because this node only has a single input. In the matrix on the right, we knock out X_2. The bolded entries show where changes are made as a result. All outputs in the X_2 truth table are set to 0. And because the topology matrix defines X_2 as the second input for X_3, wherever Input 2 equals 1 in the X_3 truth table, the output is set to NA.

After altering the truth table matrix, we recalculate conditional entropies across all network nodes and sum the values as demonstrated in section 4.2 to find the entropy of the subnetwork. We divide the subnetwork entropy by the parent network entropy to find a

sub-to-parent entropy ratio. We store this ratio in the summary matrix. We repeat the knock-out procedure on the node with the least determinative power of all remaining nodes until no nodes remain. Where two nodes have equal determinative power values, they are ordered randomly. The truth table matrix is never reset to its original state, so all alterations to the matrix and consequent changes in ratios of entropies are cumulative. With each reduction, the resulting entropy ratio is stored in the summary matrix. We obtain a sequence $R(N), R(N-1), \ldots, R(1)$ of sub-to-parent entropy ratios starting with the parent network and ending with the network reduced to the most determinative node. We repeat this procedure for each network scenario under consideration.

To account for the variation between simulations with various random initial states we consider a Monte Carlo approach. We repeat the initialization, calculation, and reduction procedures 50 times for each network type and all ratios of entropies are added to one summary matrix. The columns of the summary matrix correspond to the 50 networks generated and the rows correspond to the 100 reductions performed on each network. More precisely, given a generic network type, we construct the matrix

$$R = \begin{pmatrix} R_1(N) & R_2(N) & \ldots & R_{50}(N) \\ R_1(N-1) & R_2(N-1) & \ldots & R_{50}(N-1) \\ \vdots & \vdots & & \vdots \\ R_1(1) & R_2(1) & \ldots & R_{50}(1) \end{pmatrix}$$

where $R_u(v)$ is the sub-to-parent entropy ratio corresponding to the random initial state $u \in \{1, 2, \ldots, 50\}$ and the subnetwork with $v \in \{1, 2, \ldots, N\}$ nodes. As part of the Monte Carlo procedure, we average the row values to find the mean sub-to-parent entropy ratios associated with each subnetwork size, that is

$$\bar{R} = \left(\frac{1}{50} \sum_{k=1}^{50} R_k(N) \;\; \frac{1}{50} \sum_{k=1}^{50} R_k(N-1) \;\ldots\; \frac{1}{50} \sum_{k=1}^{50} R_k(1) \right)^T$$

where T stands for transposed. We plot the mean values. After gathering figures for all network types, we compare trends in entropy ratios across the various network settings.

To assess more thoroughly the utility of the determinative-power-based reduction algorithm described above, we prepare a control data set using a random reduction procedure for each network type. We follow the same initialization, calculation, and reduction steps

with only one exception. Rather than selecting knock-out nodes according to their determinative power, we choose them at random. Comparison of the resulting ratio plots offers insights concerning how the impacts of determinative-power-based reduction vary for different network types.

Finally, we repeat the entire procedure replacing the conditional-entropy-based calculation of network entropy with a node-entropy-based measure. This means that, rather than summing conditional entropies across the entire network, we simply calculate the entropies of each node and sum them across the network assuming the nodes act independently. More precisely, we define the aggregated network entropy as follows.

Definition 8. *The* ***aggregated network entropy*** *is given by*

$$AH(X_1, X_2, ..., X_n) = \sum_{i=1}^{N} H(X_i).$$

The results constitute the aggregated node entropies of the network. Just as we did with the conditional-entropy-based network entropy measure, we compare the aggregated node entropies of the parent network with those of each subnetwork to create a sequence of entropy ratios which we then graph to identify patterns across network settings. We compare patterns found in the aggregated node entropy graphs with those found in the network entropy graphs to identify commonalities and discrepancies between the two approaches.

We note here that for computational purposes we have limited our analysis to a maximum network size of $N = 100$, yet the reduction algorithm represents basically an exercise of creating subnetworks iteratively, and could be considered a study of the impact of network size despite de fact that the determinative power of nodes is inherited from the previous parent network.

5. Results

In total, the research included initialization and reduction of 144 networks: 36 measured by network entropy and reduced according to determinative power, 36 measured by network entropy (Definition 7) and reduced randomly, 36 measured by aggregated node entropy (Definition 8) and reduced according to determinative power, and 36 measured by aggregated node entropy and reduced randomly. We begin by focusing on networks measured by network entropy. The plots reveal that as reductions increase from 0 to 100, corresponding

to a decrease of subnetwork size from 100 to 0, sub-to-parent entropy ratios for all network types and both reduction methods follow a curve with a shape that is either generally concave down decreasing or nearly diagonal. Figure 1 demonstrates the curve shapes using examples from two network types.

Across all network types, the slightest changes in ratios of entropies tend to occur with the initial node reductions. However, as we see in Figure 1, the slope of the curves varies for different network types. By comparing the curves corresponding to each setting across the three network characteristics of connectivity, topology, and functionality, we gain insights into how relationships between entropy and subnetwork size differ for various network types when using the determinative-power-based reduction method.

5.1. **Connectivity.** Figure 2 shows compares the shapes of entropy ratios curves with relation to K across all tested combinations of topologies and functions. The plots show that for networks with topologies generated from uniform and power-law distributions, changes in the value of K have minimal effect on sub-to-parent entropy ratios since the graphs are very close to each other. In networks with topologies generated from Poisson distributions, however, lower K values are associated with higher sub-to-parent entropy ratios. In this case, for subnetworks of all sizes, the highest sub-to-parent entropy ratios are found in $K = 2$ networks. Put another way, when comparing subnetworks of similar sizes, $K = 2$ subnetworks have greater entropy ratios than those with higher K values. The difference is greatest between reductions 25 and 75, corresponding to subnetworks with between 25 and 75 nodes.

Another interesting observation in Figure 2 are the inflection points where the graphs change concavity and potential plateaus are occurring as seen in the plots associated with power-law-distributed topologies. Each of the 12 curves associated with power-law distributions reveals a similar structure with two inflection points. The phenomenon is most easily recognized in the plots associated with switch functions. Examination of the summary matrices associated with each network types reveals that the networks with power-law-distributed topologies go through periods in which their sub-to-parent entropy ratios are "frozen" across multiple network reductions (plateaus). This means that even as nodes are reduced, the sum of conditional entropies across the network remains unchanged. This would happen whenever the mutual information between a reduced node and all other nodes

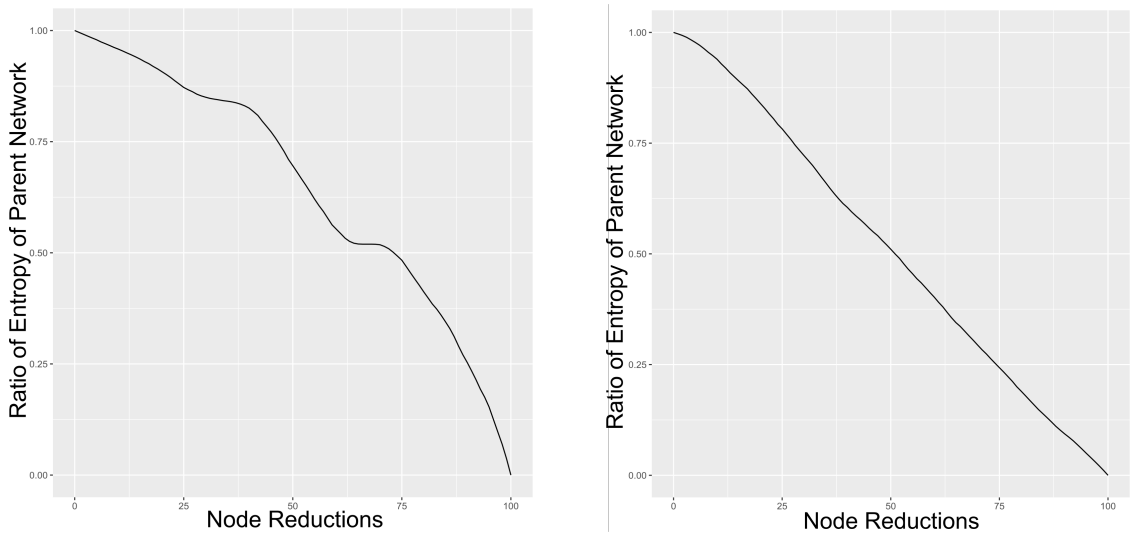

FIGURE 1. Example plots of sub-to-parent entropy ratios showing the characteristic curve shapes formed as subnetwork size decreases. Each curve represents average ratio values for 50 networks. The plot on the left, representing a network (type 5 in Table 1-A) with an average connectivity of $K = 2$, degree values drawn form a power-law distribution, and a using a homogeneous threshold = 0.5 function, depicts a concave decreasing down curve. The plot on the right, representing a $K = 4$ network drawn from a Poisson distribution and adhering to a canalyzing function of depth 1 (type 36 in Table 1-B), shows a curve closer to the diagonal. Both entropy ratio curves are generated using the network entropy measure of Definition 7.

of the subnetwork equals 0, so that no further information is gained or lost with the reduction of that node. Note that these inflection points occur between the 40th and 80th reductions, meaning that each of the relevant node reductions involves a node with initial determinative power measurements high enough to rank it between 20 and 60 among all 100 nodes of its respective network. Thus, the node's logic almost certainly contributes to the sum of conditional entropies of the parent network (the network entropy). However, by the time the node is reduced, preceding node reductions have reduced these conditional entropies to 0, and the node has no effect on subnetwork entropy. Additional research should

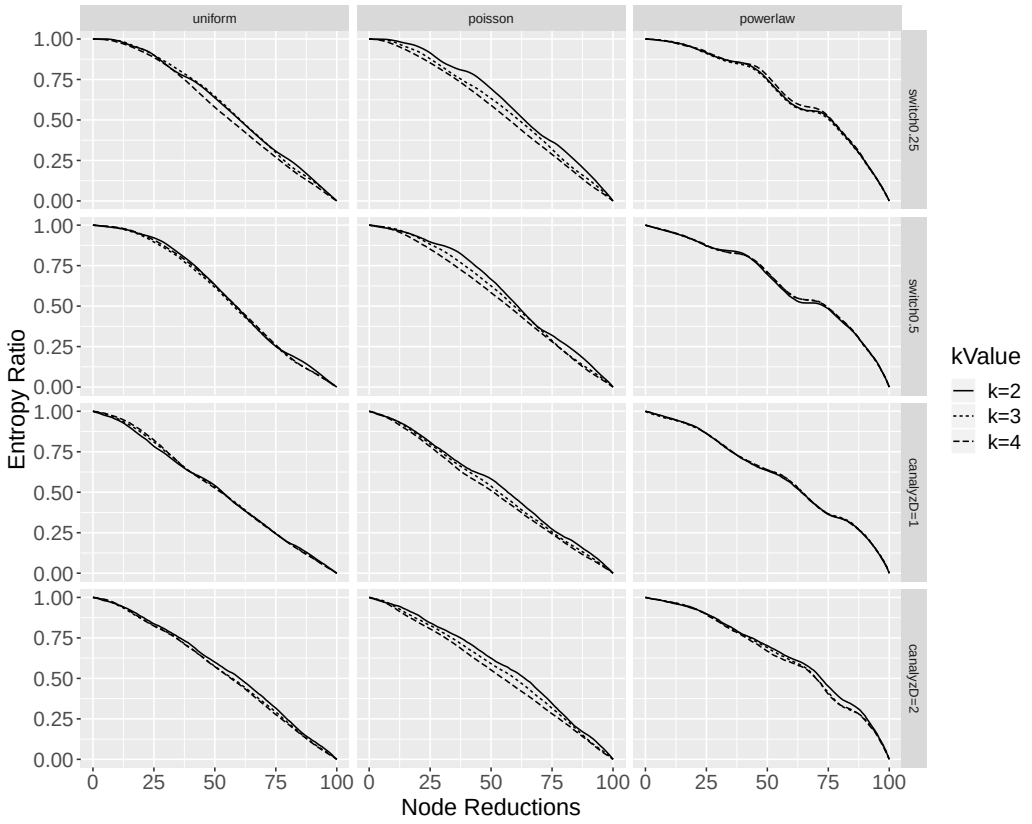

FIGURE 2. A plot of sub-to-parent entropy ratios for all 36 tested network types, illustrating trends relative to K values. For networks with Poisson-distributed topologies, the decline in entropy ratios becomes more gradual as K decreases. For networks with uniform and power-law-distributed topologies, the choice of K does not appear to affect the shape of the entropy ratio curve. All entropy ratio curves are generated using the network entropy measure of Definition 7.

focus on these inflection points to determine if they are a general feature of determinative-power-based reduction of power-law-distributed networks or are instead an artifact of our network initialization process. The existence of inflection points in sub-to-parent entropy ratio plots of networks generated from specific wiring algorithms, such as those dictated in the Barabási-Albert topology, may indicate that such inflection points illustrate general features of networks with power-law-distributed topologies.

5.2. **Topology.** Figure 3 has a format similar to that of Figure 2, but in this case we prioritize the visualization of patterns with respect to topology distribution. Each of the 12 subplots corresponds to a particular combination of K value and function settings. Curves associated with topology distribution are differentiated by line type. We observe that in all cases, networks with degree values drawn from a power-law distribution exhibit more gradual declines in entropy ratios than networks relying on either uniform or Poisson distributions. In all four $K = 2$ networks, the decline of entropy ratios is slightly more gradual when degree values are drawn from Poisson distributions than when drawn from uniform distributions. In $K = 4$ networks and $K = 3$ networks, separations between the Poisson and uniform curves are not as apparent.

The explanation for differences in entropy ratios curves illustrated in Figure 3 is intuitive. In networks with out-degree values generated from power-law distributions, a small number of nodes hold a relatively large proportion of the network's total number of outgoing links. With each outlink comes the potential for added determinative power. Indeed, Pentzien et al. [9] showed a fair correlation between out-degree links and determinative power. Furthermore, because we have set the out-degree and in-degree values as equal for each node, hubs have very large truth tables and state spaces. As the network is reduced, an increasing proportion of inputs for any given node begins to freeze and state spaces become smaller. For low-degree nodes with small state spaces, the freezing of input nodes can reduce the uncertainty of the output very quickly. High-degree nodes, which are more numerous in Power-Law-distributed networks, are more likely to retain some degree of uncertainty even after a few inputs are frozen. The apparent distinction between the Poisson and uniform curves where $K = 2$ or $K = 3$ owes to a similar principle. At small values of K, the Poisson distribution is skewed and resembles a power-law distribution.

5.3. **Functionality.** Figures 4 and 5 concern entropy ratios with regards to the function setting. To more clearly exhibit trends, we create separate facet plots to distinguish between threshold and canalyzing functions. In Figure 4, the grid of nine subplots includes every combination of the three degree-distribution settings and three K values. In each subplot, threshold functions with threshold values of 0.25 and 0.5 are distinguished by linetype. In networks with power-law-distributed topologies, networks using a 0.25 threshold exhibit a more gradual decline in entropy ratios than those using a 0.5 threshold. Exploratory tests

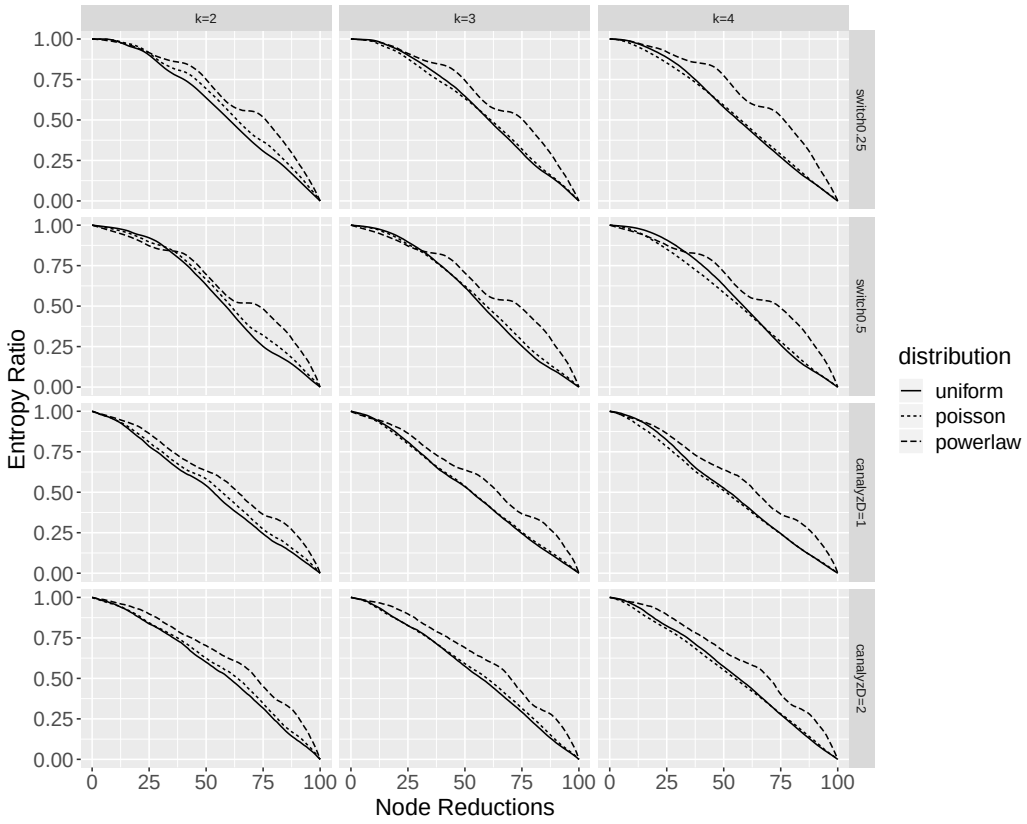

FIGURE 3. A plot of sub-to-parent entropy ratios for all 36 tested network types, showing trends relative to topology distribution settings. For all network types, the decline in the ratios of entropies is most gradual when degree values are assigned from a power-law distribution. All entropy ratio curves are generated using the network entropy measure of Definition 7.

regarding a 0.75 threshold function, not included in this visualization, revealed a curve very similar to that of the 0.25 threshold function.

Figure 5 displays trends resulting from changes in canalyzing depth. In all nine combinations of connectivities and topologies, entropy ratios in networks with $d = 2$ functions tend to be greater than those with $d = 1$ functions across all subnetwork sizes. The difference appears to be largest for $K = 2$ networks and networks with power-law-distributed topologies. Figure 5 suggests a possible link between canalyzing depth and information propagation in that our determinative-power-based reduction algorithm, which prefers to retain nodes

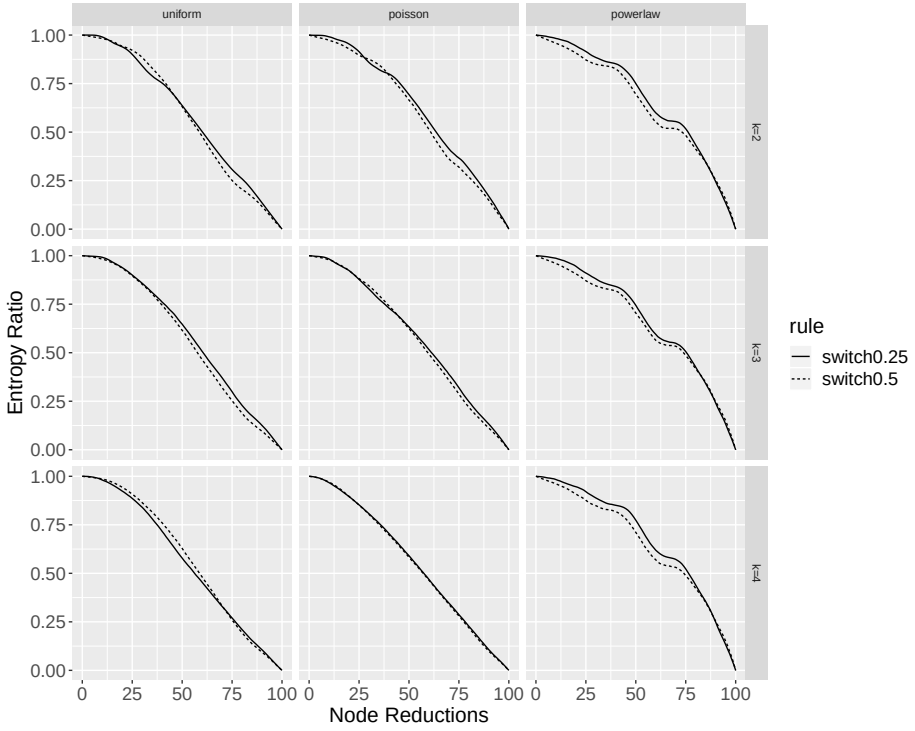

FIGURE 4. A plot of sub-to-parent entropy ratios for 18 tested network types showing trends relative to choice of specification of threshold function. In networks with power-law-distributed topologies, curves corresponding to 0.25 thresholds decline more gradually than those corresponding to 0.5 thresholds. All entropy ratio curves are generated using the network entropy measure of Definition 7.

involved in numerous and high mutual-information pairings, sustains higher entropy ratios in networks with higher canalyzing depth.

5.4. **Summary of Network Entropy Trends.** Across all 36 studied networks, topology appears to be the most significant factor supporting higher sub-to-parent entropy ratios. In particular, networks with power-law-distributed topologies reveal consistently higher entropy ratios. In networks with canalyzing functions, canalyzing depth appears to also be a consistent factor, with increasing canalyzing depth leading to higher entropy ratios in subnetworks of all sizes. Elsewhere, combinations of network settings yield interesting results. Lower K values lead to higher entropy ratios only in networks with Poisson-distributed

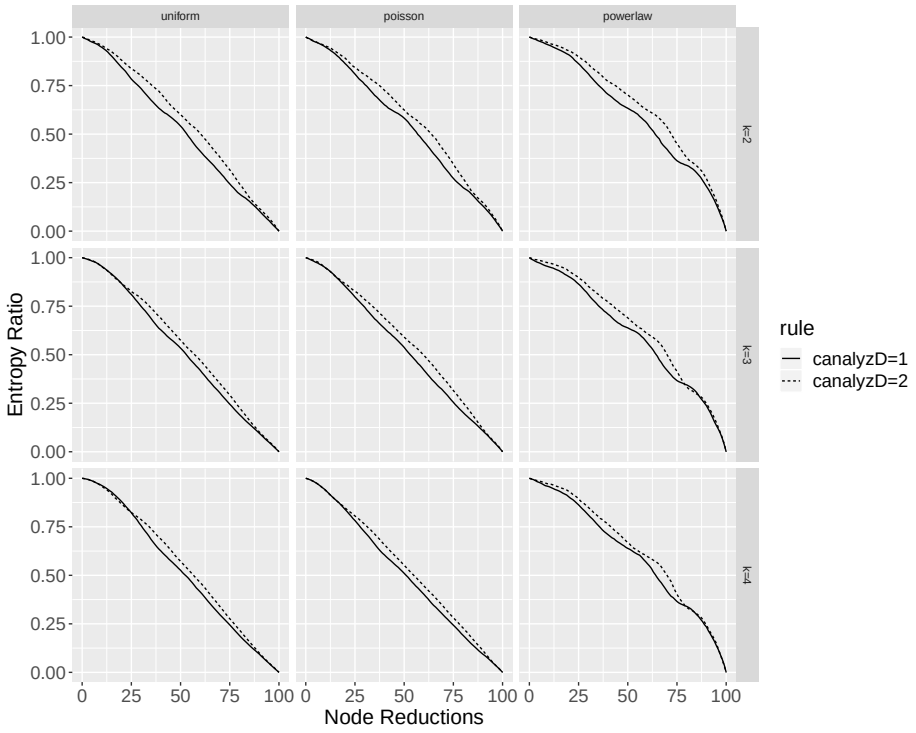

FIGURE 5. A plot of sub-to-parent ratios of entropies for 18 tested network types showing trends relative to depth of canalyzation. For $K = 2$ networks, the decline in the ratios of entropies becomes more gradual as canalyzing depth increases. All entropy ratio curves are generated using the network entropy measure of Definition 7.

topologies. In networks with threshold functions, the choice of threshold appears to matter only in networks with power-law-distributed topologies.

5.5. **Comparison of Reduction Methods.** Next, we repeat all initialization, calculation, and reduction steps with one change to randomize the reduction process. In the determinative-power-based method, the algorithm selects the remaining node with the least determinative power. In the random method, the algorithm selects a node randomly, irrespective of determinative power or any other measured attribute. Comparison of the two reduction methods reveals an unambiguous and consistent trend. Across all network types, sub-to-parent entropy ratios resulting from determinative-power-based reduction are greater

than those ratios resulting from a random reduction. This suggests that determinative-power-based reduction may be a useful approach for any case in which network entropy and information retention are prioritized during network simplification.

Inspection of the sub-to-parent entropy curves associated with each network type is informative. The relative sizes of the disparities between curve shapes highlight the instances in which determinative-power-based reduction is particularly advantageous. Figure 6 illustrates curves for two network types. Subplots on the left pertain to networks with settings of $K = 2$, power-law distribution, and threshold = 0.5 (type 5 in Table 1-A). The network visualized on the right has settings of $K = 4$, Poisson distribution, and canalyzing depth = 1 (type 36 in Table 1-B). Plots on the top row show the entropy curves yielded by the respective reduction methods for both network types. We see here that in both cases the decline in ratios is more gradual using the determinative-power method. The distance between the two curves, however, differs substantially. Where the difference is greatest, the determiative-power-based reduction method yields particularly higher sub-to-parent entropy ratios for the specified network type. The bottom row offers an alternative visualization showing determinative-power-based reduction entropy ratios relative to random reduction entropy ratios.

We quantify differences between the two reduction methods for various network types by measuring the Euclidean distance between the associated vectors of entropy ratios. We recall that the Euclidean distance between vectors u and v of length N is given by

$$\sqrt{\sum_{i=1}^{N}(u_i - v_i)^2}.$$

Tables 3-A and 3-B sort network types by Euclidean distance, allowing us to easily identify the types for which determinative-power-based reduction is particularly advantageous over random reduction. We observe that of the 13 network types with the largest Euclidean distances, 12 have power-law-distributed topologies. There does not appear to be a similar ordering between network types with Poisson-distributed and uniform-distributed topologies. Trends in functionality among the top network types are apparent, though not as blatant. The three network types with power-law-distributed topologies and threshold = 0.5 functions are at the top of the list and six of the top seven network types use threshold functions rather than canalyzing functions. All of the eight network types with threshold

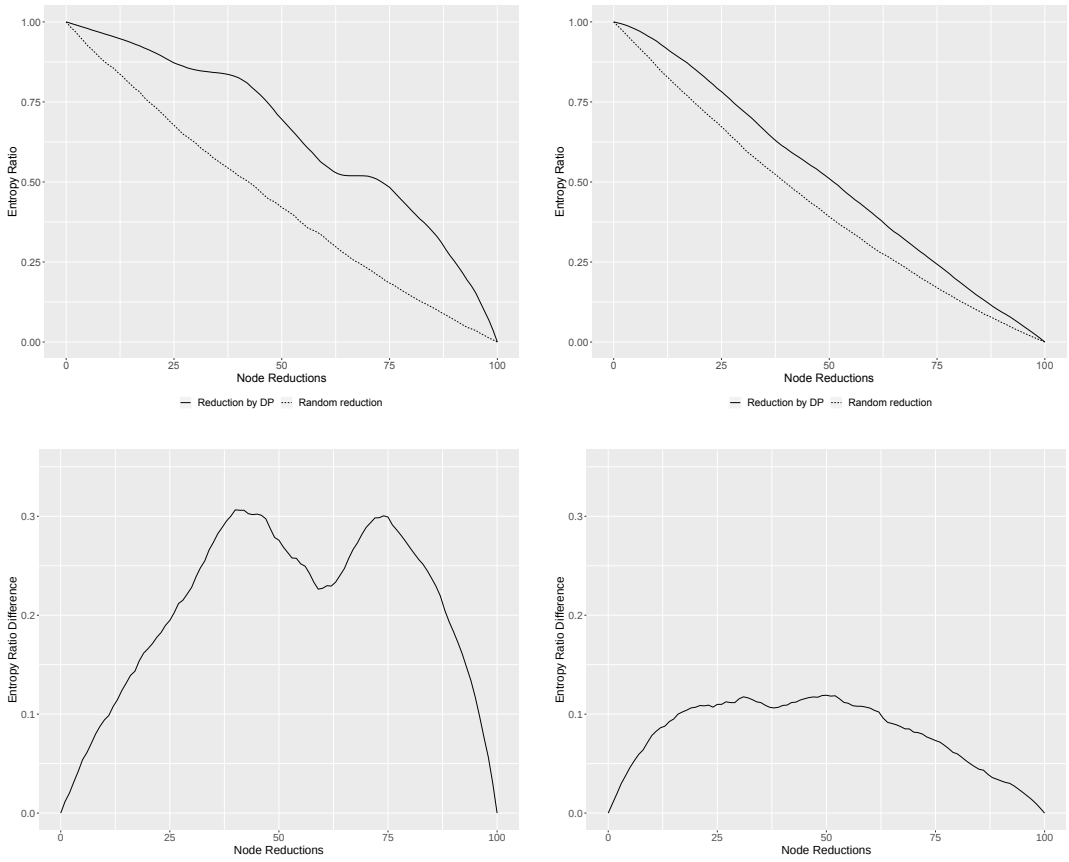

FIGURE 6. A comparison of determinative-based network reduction and random reduction for two example network types. Upper left: entropy ratio curves for $K = 2$, power-law-distributed, threshold = 0.5 networks (type 5 in Table 1-A). Upper right: entropy ratio curves for $K = 4$, Poisson-distributed, canalyzing depth = 1 networks (type 36 in Table 1-B). Bottom: plots of the difference between the curves in the plots above. All entropy ratio curves are generated using the network entropy measure of Definition 7.

= 0.5 functions appear in the top half of the table. As previously discussed in this section, figures 4 and 5 did not show exceptionally high sub-to-parent entropy ratios associated with threshold = 0.5 functions, so the ascendance of those networks in regards to this Euclidean distance measure likely demonstrates that the entropy ratios generated by the random reduction method are particularly low. There are no apparent trends associated with K

TABLE 3-A. Distances between determinative-power-based reduction entropy ratios and random reduction entropy ratios, sorted by Euclidean distance. In these networks, calculations of network entropy relied upon the conditional entropies of all node-node pairs in the network.

Network Type	**Euclidean Distance**	**Average K**	**Topology Distribution**	**Function**
5	5.0282	2	power-law	threshold = 0.5
29	5.0120	4	power-law	threshold = 0.5
17	4.9800	3	power-law	threshold = 0.5
30	4.2614	4	power-law	threshold = 0.25
7	4.2431	2	power-law	canalyzing d = 2
18	4.1586	3	power-law	threshold = 0.25
6	4.0553	2	power-law	threshold = 0.25
20	3.6876	3	power-law	canalyzing d = 1
19	3.5984	3	power-law	canalyzing d = 2
31	3.4831	4	power-law	canalyzing d = 2
8	3.3272	2	power-law	canalyzing d = 1
9	3.1238	2	Poisson	threshold = 0.5
32	3.1141	4	power-law	canalyzing d = 1
25	2.9166	4	uniform	threshold = 0.5
21	2.7868	3	Poisson	threshold = 0.5
1	2.7062	2	uniform	threshold = 0.5
13	2.4637	3	uniform	threshold = 0.5
33	2.2538	4	Poisson	threshold = 0.5

values, so connectivity does not appear to play a strong role in deciding which networks are best suited for determinative-power-based reduction.

5.6. **Aggregated Node Entropy Measure.** Finally, we compare the results described above, which relied upon conditional entropies to measure total network entropy, with a second set of results relying upon the alternative measure of aggregated node entropies

TABLE 3-B. Continuation of distances between determinative-power-based reduction entropy ratios and random reduction entropy ratios, sorted by Euclidean distance. In these networks, calculations of network entropy relied upon the conditional entropies of all node-node pairs in the network.

Network Type	**Euclidean Distance**	**Average K**	**Topology Distribution**	**Function**
10	2.0608	2	Poisson	threshold = 0.25
14	2.0022	3	uniform	threshold = 0.25
11	1.9156	2	Poisson	canalyzing d = 2
23	1.5032	3	Poisson	canalyzing d = 2
22	1.4587	3	Poisson	threshold = 0.25
27	1.3954	4	uniform	canalyzing d = 2
12	1.3807	2	Poisson	canalyzing d = 1
3	1.2747	2	uniform	canalyzing d = 2
15	1.2699	3	uniform	canalyzing d = 2
28	1.2345	4	uniform	canalyzing d = 1
26	1.1650	4	uniform	threshold = 0.25
16	1.1383	3	uniform	canalyzing d = 1
24	1.0816	3	Poisson	canalyzing d = 1
2	1.0238	2	uniform	threshold = 0.25
34	0.9858	4	Poisson	threshold = 0.25
35	0.9661	4	Poisson	canalyzing d = 2
4	0.8801	2	uniform	canalyzing d = 1
36	0.7817	4	Poisson	canalyzing d = 1

(Definition 8). We repeat the procedures described in Section 4, including all initialization, calculation, and reduction steps. And once again, for each of the 36 network types, we perform the procedures two times: first using the determinative-power-based reduction method and then using the random reduction. To calculate the aggregated node entropies of all parent networks and subnetworks, we calculate the entropy for each of the $N = 100$ nodes and sum. The sub-to-parent entropy ratios represent the proportional relationships between

the aggregated node entropies of subnetworks and their corresponding parent networks. We rleate these findings with those from the network entropy approach to gain insights into the impacts of the choice of entropy measurement. We find substantial differences between the two approaches. The most notable differences are the shape of the resulting sub-to-parent entropy ratio curves, the impact of K values in determining the rate of entropy ratio decline, and the scale of the differences between determinative-power-reduction entropy curves and random-reduction entropy curves.

In contrast to the entropy ratio curve shapes resulting from the network entropy approach of Definition 7, which were either concave down decreasing or nearly diagonal, all of

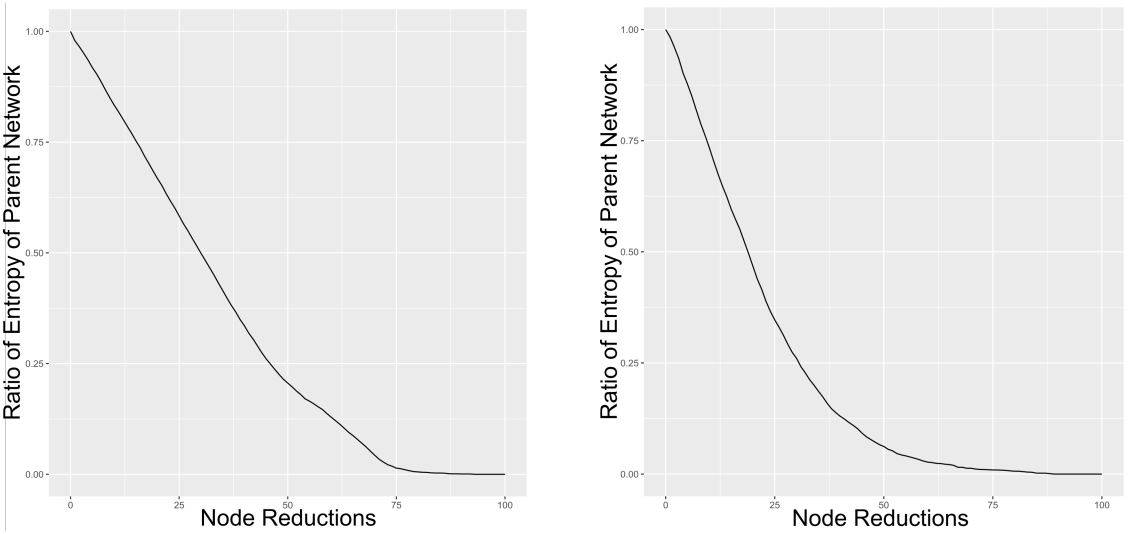

FIGURE 7. Example plots of sub-to-parent ratios of aggregated node entropies, showing the characteristic curve shapes formed as subnetwork size decreases. Compare the plots with those in Figure 1, which features curves of the same networks measured instead by network entropy. Where Figure 1 exhibits concave down decreasing curves, here the curves are concave up decreasing. All entropy ratio curves are generated using the network entropy measure (Definition 7) and exhibit similar shapes. The differences between plots in Figures 1 and 7 suggest that sub-to-parent entropy ratios are higher for subnetworks of all sizes when considering network entropy (Definition 7) rather than aggregated node entropy (Definition 8).

the curve shapes resulting from the aggregated node entropy approach are concave up decreasing. Figure 7 illustrates two such curves for comparison with Figure 1, which features plots of the same network types. The drastic differences in the shape of the curves indicate that the choice of entropy measurement has profound impacts on the apparent results of network reduction. As explored earlier in this section, network entropy calculations indicate that the initial node reductions, which remove the nodes with the least determinative power, generally have minimal impacts on the entropy ratio of the resulting subnetwork. In the aggregated-node-entropy approach, plots suggest the opposite. Initial reductions are accompanied by the greatest decreases in sub-to-parent entropy ratios.

In analysis of the sub-to-parent entropy ratio curves associated with network entropy, comparison of the relative rates of change provided key insights as to how network settings impact the values of the ratios. Some of those patterns are also found in the aggregated-node-entropy ratio curves. For instance, power-law-distributed topologies and increasing canalyzation depth both tend to yield high sub-to-parent entropy ratios in both cases. The most conspicuous divergence between the curves generated by the network entropy measure (Definition 7) and those generated by the aggregated node entropy measure (Definition 8) concerns the impact of the connectivity setting. When using the network entropy measurement (Definition 7), the value of K appeared to have a significant impact only in networks with Poisson-distributed topologies like in Figure 8. Per the aggregated node entropy approach (Definition 8), decreasing K values lead to higher sub-to-parent entropy ratios in all network types. The effects of decreases in K are greatest in networks with canalyzing functions and either Poisson or power-law-distributed topologies.

A third eye-catching difference between reduction informed by the network entropy of Definition 7 as opposed to reduction informed by aggregated node entropy of Definition 8 lies in the comparison between reduction methods. Differences between the sub-to-parent network entropy ratio curves of determinative-power-based and random reduction methods, exemplified and Figure 6 and listed in Tables 3-A and 3-B, are much greater than the differences in aggregated node entropy ratio curves of the same network types. Using network entropy, the Euclidean distances between the two ratio curves range from 0.7817 to 5.0282 as the last and the first rows in Tables 3-A and 3-B. Aggregated node entropy, however, produces Euclidean distances ranging from 0.2241 to 1.9780. Additionally, while

concerning network entropy the ratios generated by determinative-power-based reduction are greater than those generated by random reduction for subnetworks of all sizes, in the case of aggregated node entropy the two reduction curves intersect for all network types. The intersection point occurs between the 46th and 92nd node reduction. Thereafter, subnetworks generated by random reduction tend to have higher sub-to-parent entropy ratios.

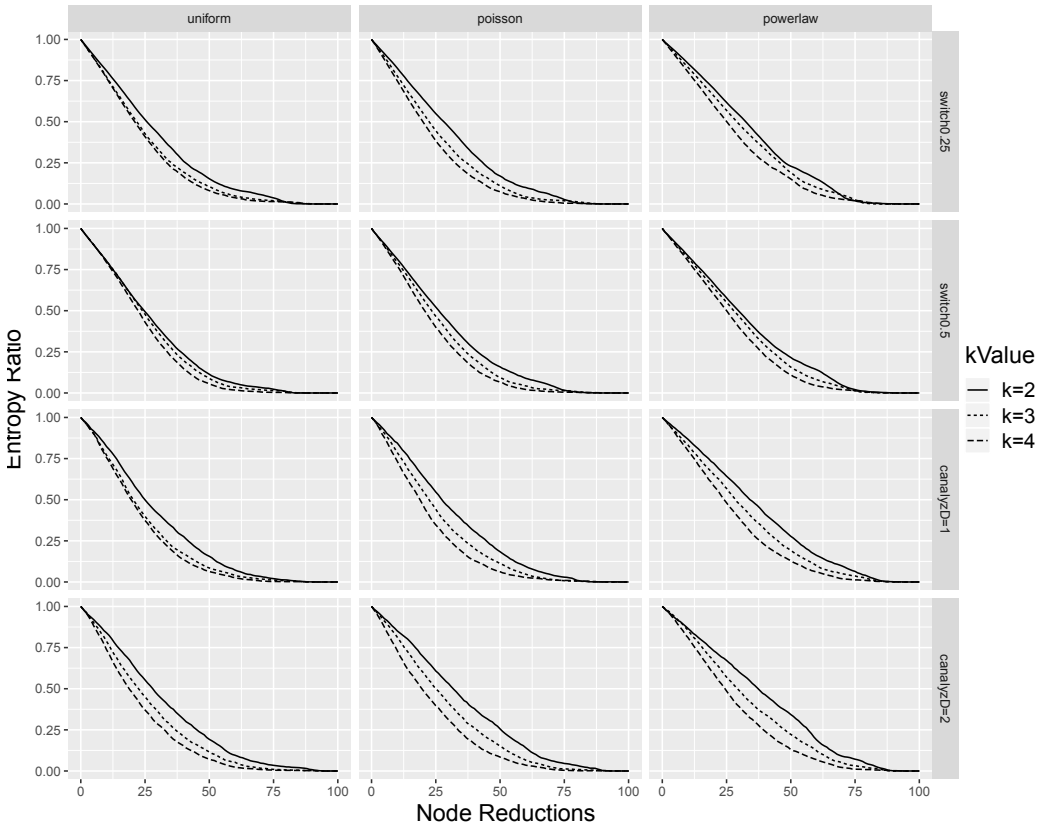

FIGURE 8. A plot of sub-to-parent entropy ratios for all 36 tested network types measured by aggregated node entropy, illustrating trends relative to K values. Compare to Figure 2, which plots the same features according to the network entropy measurement, which is based on a calculation of conditional entropies. Here, there is a distinct trend in which decreasing K values result in higher sub-to-parent entropy ratios for subnetworks of all sizes.

6. Conclusion

Our research explores the application of the determinative power of nodes in the reduction of Boolean network models, which can be large and computationally burdensome. We develop a determinative-power-based reduction algorithm and test it on 36 network types with various settings of connectivity, topology, and functionality. We evaluate the generated subnetworks according to their network entropies as ratios of the network entropies of their corresponding parent networks. We plot curves to investigate changes in sub-to-parent entropy ratios as subnetworks decrease in size. We compare curves generated by the determinative-power-based reduction algorithm with those generated by a randomized reduction process. Lastly, we repeat the entire procedure using an alternate measurement of aggregated node entropy.

The research yields three primary conclusions. First, using the determinative-power-based reduction, for all 36 network types the initial node reductions effect the least change in sub-to-parent network entropy ratios. This supports previous research [9, 19, 21] indicating that determinative power is a usefulness tool in identifying subnetworks. Second, for all network types, determinative-power-based reduction "outperform" random reduction in that the sub-to-parent network entropy ratio curves generated via determinative power reliably have higher values for subnetworks of all sizes. Third, simulation results differ substantially when using an aggregated node entropy measurement in lieu of network entropy. This suggests that an accounting of the conditional entropies associated with a network can drastically affect our understanding of the network's behavior. In this paper, we emphasize the network entropy approach, based on conditional entropy, because we feel it represents a more nuanced view of network complexity and information-processing characteristics. It is possible that for larger networks the two approaches would not generate the differences seen in this paper due to diminishing local correlations.

The shapes of sub-to-parent entropy ratio curves for different network type suggest that in some network types, subnetworks have a particularly high proportion of the parent's network entropy. The highest ratios are found in networks with topologies drawn from power-law distributions. An increase in canalyzation depth also appears to lead to higher sub-to-parent entropy ratios.

Networks with power-law-distributed topologies also show the greatest difference in performance between the determinative-power-based and random reduction methods, suggesting that some degree of the improved overall performance of these networks in the maintenance of relatively high entropy ratios owes to interactions between their topology and determinative power measurements of their nodes. The greatest differences between the two reduction methods involve networks with power-law-distributed-topologies and threshold = 0.5 functions.

A useful next step in assessing the utility of a determinative-power-based reduction method would involve a comparison between its performance and that of a stable-states-based reduction method relying on the removal of frozen, leaf, and mediator nodes. Our conjecture is that the two methods would have similar performance as both would tend to target leaf nodes and nodes which serve as inputs only for frozen nodes. However, there are cases in which the methods diverge. A mediator node, irrelevant to the stable states of a network, may have an amount of determinative power that could result in its inclusion in relatively small subnetworks. Conversely, the determinative power method may identify a node that is not a frozen, leaf, or mediator node but nonetheless has little impact on the dynamics of the network. An analysis of how these reduction strategies actually differ in node selection would be of interest.

Additionally, future research could focus on determinative-power-based network types more explicitly related to real-world phenomena. Specifically, topology and function settings could be more finely tuned through the incorporation of small-world and scale-free wiring algorithms or the implementation of heterogeneous functionality. One useful complication of these simulations would involve the de-coupling of in-degree and out-degree values. While it is intuitive that nodes with high out-degree values tend to have higher determinative power, Epstein and Bazzan [39] identified cases in which "simpler" nodes–those with fewer inputs–are generally more influential. Perhaps nodes with low in-degree values and high out-degree values tend to have the most determinative power.

Pentzien et al. [9] performed determinative power analysis on many biological models found on the *Cell Collective* platform, with intriguing results. A categorization of the models on the platform according to topology and functionality would offer a helpful understanding of the relationships between determinative power and network settings in real-world systems.

7. ACKNOWLEDGMENTS

MP was supported by the Kerrigan Research Minigrant Program at the University of Nebraska at Omaha. MTM was supported by the University of Nebraska at Omaha Maury and Nancy Lipton Professorship of Mathematics.

References

[1] S.A. Kauffman, *Metabolic stability and epigenesis in randomly constructed genetic nets*, Journal of Theoretical Biology **22(3)** (1969), 437-467, DOI 10.1016/0022-5193(69)90015-0.

[2] S.A. Kauffman, *The origins of order*, Oxford University Press, 1993.

[3] K. Klemm and S. Bornholdt, *Stable and unstable attractors in Boolean networks*, Physical Review E **72** (2000), 055101, DOI 10.1103/PhysRevE.72.055101.

[4] I. Shmulevich, E.R. Dougherty, and W. Zhang, *From Boolean to probabilistic Boolean networks as models for genetic regulatory networks*, Proceedings of the IEEE **90(11)** (2002), 1778-1792, DOI 10.1109/JPROC.2002.804686 .

[5] I. Shmulevich and S.A. Kauffman, *Activities and sensitivities in Boolean network models*, Physical Review Letters **93(4)** (2004), 048701, DOI 10.1103/PhysRevLett.93.048701.

[6] T. Helikar, J. Konvalina, J. Heidel, and J.A. Rogers, *Emergent decision-making in biological signal transduction networks*, Proceedings of the National Academy of Sciences of the United States of America **105(6)** (2008), 1913-1918, DOI 10.1073/pnas.0705088105.

[7] C. Huepe and M. Aldana-González, *Dynamical phase transition in a neural network model with noise: an exact solution*, Journal of Statistical Physics **108(3-4**) (2002), 527-540, DOI 10.1023/A:1015777824097.

[8] R. Albert and H. Othmer, *The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in* Drosophila melanogaster, Journal of Theoretical Biology **223** (2003), 1-18, DOI 10.1016/s0022-5193(03)00035-3.

[9] T. Pentzien, B.L. Puniya, T. Helikar, and M.T. Matache, *Identification of biologically essential nodes via determinative power in logical models of cellular processes*, Frontiers in Physiology **9** (2018), DOI 10.3389/fphys.2018.01185.

[10] V. Kaufman, T. Mihaljev, and B. Drossel, *Scaling in critical random Boolean networks*, Physical Review E **72(4)** (2005), 046124, DOI 10.1103/PhysRevLett.90.068702.

[11] V. Kaufman and B. Drossel, *Relevant components in critical random Boolean networks*, New Journal of Physics **8(10)** (2006), DOI 10.1088/1367-2630/8/10/228.

[12] J.E. Socolar and S.A. Kauffman, *Scaling in ordered and critical random Boolean networks*, Physical Review Letters **90(6)** (2003), 068702, DOI 10.1103/PhysRevLett.90.068702.

[13] S. Bilke and F. Sjunnesson, *Stability of the Kauffman model*, Physical Review E **65** (2001), 016129, DOI 10.1103/PhysRevE.65.016129.

[14] K.A. Richardson, *Simplifying Boolean networks*, Advances in Complex Systems **8(4)** (2004), 365-381, DOI 10.1142/S0219525905000518.

[15] A. Saadatpour, R. Albert, and T.C. Reluga, *A reduction method for Boolean network models proven to conserve attractors*, SIAM Journal on Applied Dynamical Systems **12** (2013), 1997-2011, DOI 10.1137/13090537X.

[16] I. Ivanov, R. Pal, and E. Dougherty, *Dynamics preserving size reduction mappings for probabilistic Boolean networks*, IEEE Transactions on Signal Processing **55(5)** (2007), DOI 10.1109/TSP.2006.890929.

[17] J.G. Klotz, D. Kracht, M. Bossert, and S. Schober, *Canalizing Boolean functions maximize the mutual information*, IEEE Transactions on Information Theory **60(4)** (2014), 2139-2147, DOI 10.1109/TIT.2014.2304952.

[18] A. Jiménez, *A complex network model for seismicity based on mutual information*, Physica A **392** (2013), 2498-2506. DOE 10.1016/j.physa.2013.01.062.

[19] R. Heckel, S. Schober, and M. Bossert, *Harmonic analysis of Boolean networks: determinative power and perturbations*, EURASIP Journal on Bioinformatics and Systems Biology (2013), 1-13, DOI 10.1186/1687-4153-2013-6.

[20] A.S. Ribeiro, S.A. Kauffman, J. Lloyd-Price, B. Samuelsson, and J.E.S. Socolar, *Mutual information in random Boolean models of regulatory networks*, Physical Review E **77** (2008), 011901, DOI 10.1103/PhysRevE.77.011901.

[21] M.T. Matache and V. Matache, *Logical reduction of biological networks to their most determinative components*, Bulletin of Mathematical Biology **78** (2016), 1520-1545, DOI 10.1007/s11538-016-0193-x.

[22] P. Erdrich, R. Steuer, and S. Klamt, *An algorithm for the reduction of genome-scale metabolic models to meaningful core models*, BMC Systems Biology **9** (2015), DOI 10.1186/s12918-015-0191-x.

[23] L. Gorjao, A. Saha, G. Ansmann, U. Feudel, and K. Kehnertz, *Complexity and irreducibility of dynamics on networks of networks*, Chaos **28** 106306, (2018), DOI 10.1063/1.5039483.

[24] P. Krawitz and I. Shmulevich, *Basin entropy in Boolean network ensembles*, Physical Review Letters **98** (2007), 158701, DOI 10.1103/PhysRevLett.98.158701.

[25] P. Krawitz and I. Shmulevich, *Entropy of complex relevant components of Boolean networks*, Physical Review E **76** (2007), 036115, DOI 10.1103/PhysRevE.76.036115.

[26] A. Shreim, A. Berdahl, F. Greil, J. Davidsen, M. Paczuski, *Attractor and basin entropies of random Boolean networks under asynchronous stochastic update*, Physical Review E **82(3)** (2010), 035102, DOI: 10.1103/PhysRevE.82.035102.

[27] T. Helikar, B. Kowal, S. McClenathan, M. Bruckner, T. Rowley, B. Wicks, M. Shrestha, K. Limbu, and J.A. Rogers, *The cell collective: toward an open and collaborative approach to systems biology*, BMC Systems Biology **6:96** (2012) DOI 10.1186/1752-0509-6-96.

[28] T. Helikar, B. Kowal, and J.A. Rogers, *A cell simulator platform: the cell collective*, Clinical Pharmacology and Therapeutics **93** (2013), 393-395, DOI 10.1038/clpt.2013.41.

[29] S.A. Kauffman, *Antichaos and adaptation*, Scientific American **265(2)** (1991), 78-85, DOI 10.1038/scientificamerican0891-78.

[30] D.J. Watts and S.H. Strogatz, *Collective dynamics of 'small-world' networks*, Nature **393** (1998) 440-442.

[31] A. Barabási, R. Albert, and H. Jeong, *Mean-field theory for scale-free random networks*, Physica A **272** (1999), 173-187, DOI 10.1016/S0378-4371(99)00291-5.

[32] A. Barabási and R. Albert, *Emergence of scaling in random networks*, Science **286** (1999), 509-512, DOI: 10.1126/science.286.5439.509.

[33] P. Holme and B.J. Kim, *Growing scale-free networks with tunable clustering*, Physical Review E **65(2 Pt 2)** (2002) 026107, DOI 10.1103/PhysRevE.65.026107.

[34] B.W. Wacker, M.T. Matache, and J.A. Rogers, *Boolean Network topologies and the determinative power of nodes*, submitted for publication, 2019.

[35] K. Jansen and M.T. Matache, *Phase transition of Boolean networks with partially nested canalizing functions*, European Physical Journal B **86(7**) (2013), DOI 10.1140/epjb/e2013-40009-4.

[36] L. Layne, E. Dimitrova, and M. Macauley, *Nested canalyzing depth and network stability*, Bulletin of Mathematical Biology **74(2)** (2011), 422-433, DOI 10.1007/s11538-011-9692-y.

[37] D.M. Wittman, C. Marr, and F.J. Theis, *Biologically meaningful updates rules increase the critical connectivity of generalized kauffman networks*, Journal of Theoretical Biology **266(3)** (2010), 436, DOI 10.1016/j.jtbi.2010.07.007.

[38] T.M. Cover and J.A. Thomas, *Elements of information theory*, Wiley-Interscience, 2006.

[39] D. Epstein and A.L.C. Bazzan, *The value of less connected agents om Boolean networks*, Physica A **392** (2013) 5387-5398. DOI 10.1016/j.physa.2013.07.004.