# Dynamic structure of networks updated according to simple, local rules

Kate L. Morrow,[1] Todd Rowland,[2] and Christopher M. Danforth[1,3]
[1]*Department of Mathematics and Statistics, University of Vermont, Burlington, Vermont 05401, USA*
[2]*Wolfram Research, Champaign, Illinois 61820, USA*
[3]*Complex Systems Center and Vermont Advanced Computing Center, Burlington, Vermont 05401, USA*
(Received 12 August 2008; revised manuscript received 27 February 2009; published 8 July 2009)

While most studies of deterministic network growth have been of one- or two-case models, here a more diverse and comprehensive method of deterministic network evolution is presented. The range of observed behavior is classified and the underlying causes of the various types of growth are investigated. The potential for prediction of the different types of network growth is also examined. It is discovered that a wide variety of behavior can be produced by a simple evolutionary setup and that the networks resulting from this method of evolution warrant further study.

DOI: 10.1103/PhysRevE.80.016103 PACS number(s): 89.75.Hc, 89.75.Fb

## I. INTRODUCTION

Increasingly, the study of networks has been attracting attention in the last few years [1,2], in part due to the interconnected nature of modern society and the increasingly available data on social, technological, and biological networks. The seminal work of Watts and Strogatz [3,4] demonstrated that networks are capable of exhibiting remarkable, nonintuitive characteristics and catalyzed what is now a very active area of research. Due to the popularity of their work, much attention has since been paid to designing models of network evolution which produce scale-free degree distributions and other "small-world" characteristics [5–10], most notably the Barabási-Albert preferential attachment model [11]. The overwhelming majority of such models have been stochastic, although a handful of deterministic models have been presented [12–16]. While many of these models (both stochastic and deterministic) often successfully capture the observed small-world properties of real-world networks, it has been shown that some of the most popular models are unable to produce all of the various types of topological community structure seen in real-world networks [17]. It is evident that a more flexible and comprehensive approach to studying structural network evolution is needed.

Here the focus is on a deterministic method of evolving network structure, a special case of the class of rules proposed by Wolfram [18]. The benefits of this approach are evident in both implementation and result. Examining deterministic models of network growth enables potential identification of the cause of behavior of interest, whereas stochastic models inherently lack such well-defined causality. When complex behavior is seen in a stochastic model, it can be a subtle matter to determine whether the complexity arises from the architecture of the model itself or from the persistent introduction of randomness into the system. Furthermore, the model presented here is much more flexible and comprehensive than previously proposed deterministic methods that have focused mostly on one- or two-case models directly giving rise to networks with scale-free degree distributions. In contrast, the model presented in this paper was not designed with any specific end goal in mind and possesses a vast number of cases. This width of scope and versatility gives rise to a rich spectrum of observed network behavior.

In this study, directed networks were evolved according to simple deterministic rules. Each rule had a number of possible cases, from which a particular node would select the case to be applied depending on the local structure of the network (see Appendix for details). Two different classes of rules were examined, depending on how far along the network a node was to search—in terms of following successive forward links—in order to determine a case. While nothing but simple behavior was found in the "distance-one" class of rules, the "distance-two" class presented a breadth of significantly more complex behavior.

Given the size of the distance-two rule space (detailed in the following section), this paper does not attempt to draw conclusions about the statistics of the rule space as a whole. Instead, the motivation for this study was to provide an introduction to this method of network structure evolution and begin investigation into the potential utility of these networks in both applied and theoretical research. The variety of observed behavior in the distance-two class was sorted into categories with common qualitative growth characteristics. Quantitative tools were developed in an effort to better understand the underlying mechanisms driving certain types of network growth behavior, as well as to explore the potential for structural behavior prediction. The conclusion discusses the effectiveness of the presented analysis methods as tools for network prediction, as well as some thoughts on potential applications of this method of network evolution.

## II. DYNAMIC STRUCTURE RULES

The networks in this study are composed of directed links, with each node having exactly two outgoing links, distinguished as the "up" and "down" links for ease of reference. (This outgoing link restriction is simply for the purpose of establishing an initial class of rules to examine. As stated later in Sec. V, future studies could certainly concern themselves with rules that allow for a different number of outgoing links.) No restrictions are placed on the number of incoming links a node may obtain. All networks examined are evolved from an initial condition of three nodes, with every

 

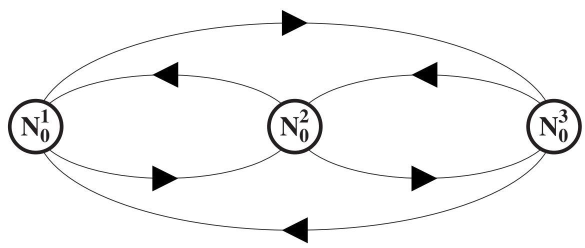


FIG. 1. Initial condition for all experiments: cyclic net of three nodes, where every node is forward connected to each of the other two.

node forward connected to each of the other nodes (Fig. 1).

Analogous to ordinary cellular automata [18], the up and down links of a particular node can be altered according to one of several cases within a given rule. The case of the rule to be applied depends on the local structure, namely, on the total number of neighbors a node has when following its forward connections over a given distance.

Depending on the rule and case, links can be rerouted or entirely new nodes (and their accompanying links) can be placed in their path. Each node is updated in sequence according to its index. Nodes are denoted by $N$ and indexed with a subscript $s$ and superscript 1 through $n_s$, where $n_s$ is the total number of nodes in the network at step $s$ (e.g., $n_0 = 3$). One "step" of the evolution is complete when each node has been updated once. At the end of each step, nodes are deleted if they can no longer be reached from $N_s^1$, the node indexed one at step $s$. Next, nodes are reindexed so that there are no gaps in the index. (The reindexing at each step requires that when a specific node in the network is referenced, both the index and step must be specified, since the index of a node can change between steps if lower-indexed nodes were deleted from the network. Note that $N_s^1$ is a special case and is always the same for all steps.) This type of network evolution can then in effect be thought of as tracing the history/evolution of the initial cluster, the set of nodes that remains reachable from $N_0^1$. If so desired, detached clusters can be examined by evolving a network under the same rule with different initial conditions.

Two classes of these dynamic network structure rules were examined: distance one and distance two. The distance-one rules each have two cases, which are applied according to the number of neighbors a node has when looking forward to a distance of one. At distance one, a node can either see one neighbor (its outgoing links are pointing to the same neighbor) or two (its outgoing links have distinct destinations). The distance two rules each have four cases. The case is chosen depending on the sum of the number of neighbors a node has at distance one and at distance two. This sum can then have values from 2 to 6, with sum two and three being combined into one case in order to make the size of the rule space computationally tractable. There are 1296 possible distance-one rules and approximately 9.6 trillion distance two rules (reduced from over $10^{16}$ rules by the combination of sums two and three into a single case).

As an example, distance-one rule 219 states that

| If the sum of the immediate neighbors is: | Then: |
|---|---|
| 1: Both outgoing links point to the same neighbor | Redirect the "up" link to the current destination of the "down" link and redirect the down link to the current destination of the up link. |
| 2: Each outgoing link points to a distinct neighbor | Leave the up link pointing to its current destination and place a new node as the new destination of the down link—whose links will point to the same starting destinations as the up and down links of the original node, respectively. |

See the Appendix for a description of rule enumeration and for more detailed explanation of the structure and application of rules and their cases.

## III. SIMULATION RESULTS

An exhaustive search of the distance-one rules was performed. Only two types of behavior were observed: namely, monotonic growth or periodic behavior. Rules which result in periodicity typically have a brief transient period of no more than five time steps. Of the 1296 distance-one rules, 116 result in fixed points, while 52 result in period two behavior. Behavior of higher periods is not observed.

A random search of over 50 000 distance-two rules was conducted. Most of the observed behaviors fit into one of a handful of categories. However, six of the searched rules exhibit potentially more complex behavior. The frequency of periodic behavior was somewhat higher than expected and rules were found with periods as high as several hundred steps. The distance-two rules with periodic behavior took much longer to settle into their periodic cycles than the distance-one rules, with observed transient phases generally on the order of 40–50 steps.

Classification of network behavior is a nontrivial problem, since direct visualization of network evolution is infeasible for large numbers of rules. In addition to the obvious computational difficulty of plotting large, distributed networks, there is also the risk of artificial creation of meaning for graphs of any size. Depending on the graphing algorithm used, nodes which appear close together in a graph may actually be separated by several steps. In the case of a directed network, one node may not even be reachable from another at all despite visual proximity. Structures could appear in the graph as a by-product of the graphing algorithm and not have physical meaning when the network is analyzed mathematically. The best solution is to take purely quantitative approaches to analyzing networks to ensure that the emergent properties observed are a direct result of the network evolution rules and not an accidental by-product of the analysis method.

### A. Classification of observed behavior

Behavior was classified according to the style of population growth seen. The full range of observed behavior could then be divided into four categories:

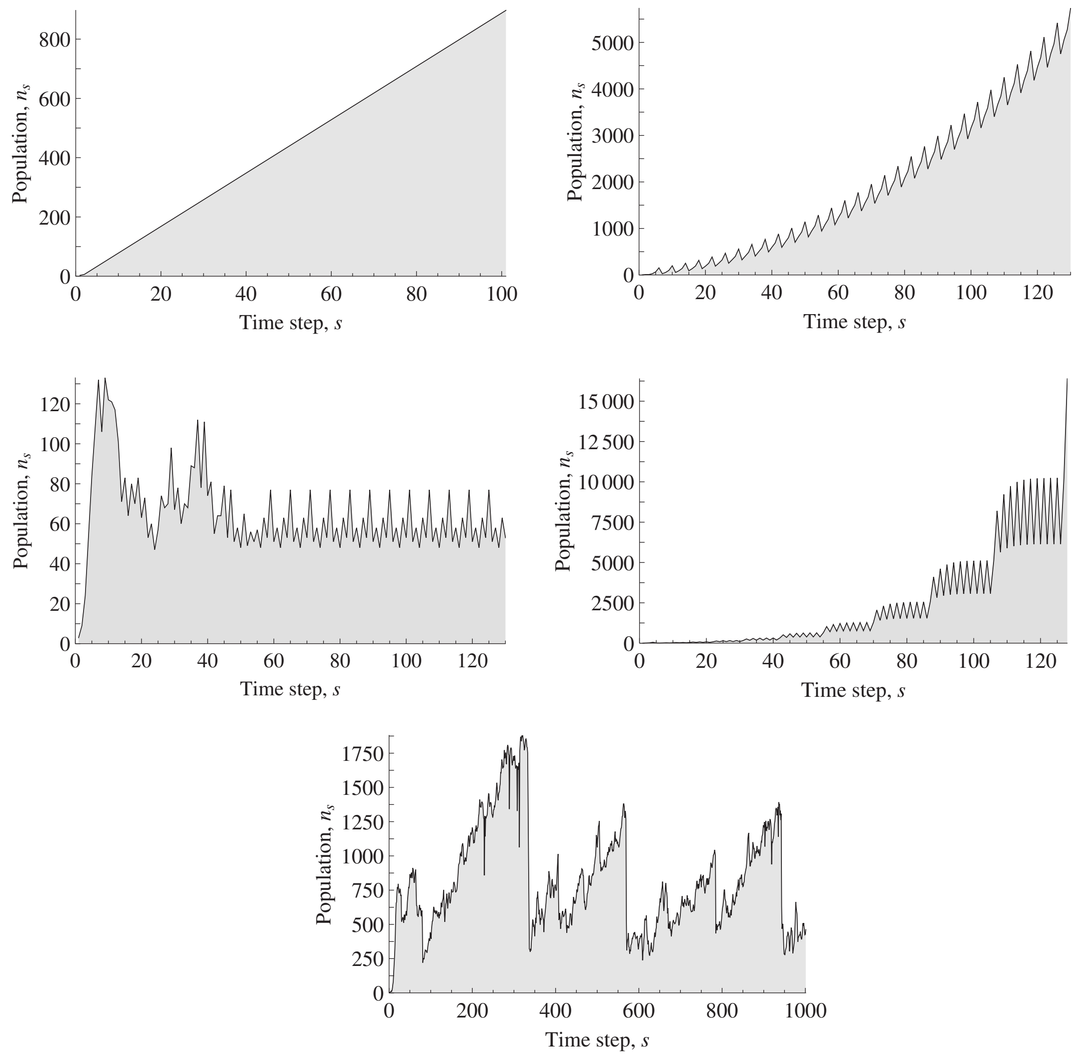


FIG. 2. Population, $n_s$, vs time step, $s$, for different categories of behavior: (a) monotonic growth, (b) monotonic trend with superimposed regular behavior, (c) periodic behavior, (d) nested behavior, and (e) complex behavior.

*1. Monotonic growth (both linear and exponential).* These rules exhibited purely monotonic growth. Both linear and exponential growths were observed, but no sublinear growth was found in the sample [see Fig. 2(a)].

*2. Monotonic growth trend with superimposed regular behavior.* These rules exhibited a positive linear or exponential growth trend, but with periodic behavior superimposed [see Fig. 2(b) for an example]. A few instances of superimposed behavior that appeared to be dampened out over time were found.

*3. Periodic or nested/resonant behavior.* These rules exhibited purely periodic or nested behavior. (For "purely periodic," the exact link structure is compared from cycle to cycle. Since these evolution rules are deterministic, once period $t$ behavior is found, the network is guaranteed to repeat its exact structure every $t$ steps.) "Nested" behavior was characterized by repeated network growth patterns that seemed to grow in amplitude each repetition. Successions of sublinear trends were seen within a few of these nested rules. For an example of periodic behavior, see Fig. 2(c); for nested behavior, see Fig. 2(d).

*4. Complex behavior.* These rules exhibited growth patterns without any apparent regular growth behavior (includes rules 1 374 996 482 325, 1 885 294 065 141, 5 969 384 073 490, 6 669 033 908 439, 8 068 664 081 521, and 8 414 212 167 895, see the Appendix for numbering scheme) [see Fig. 2(e)].

Many rules exhibited a seemingly random transient phase before settling down into one of the regular behavior categories. Consequently, one goal was to determine if there were any indicators during the transient phase which could be used to predict the category of growth a network would eventually exhibit in the long term. Indeed, a valid concern is that the rules categorized as "complex behavior" could simply be rules which eventually demonstrate behavior belonging to one of categories 1–3, yet experience extraordinarily long transient phases.

The distributions of both incoming degree and clustering coefficient were examined for the entire network as a function of time step and the discrete Fourier transform [Eq. (1)] of the population time series was also examined. Although all three tools provided useful insight into the mechanisms that underlie specific types of behavior, their utility as prediction tools was varied. Certain observed properties of net-

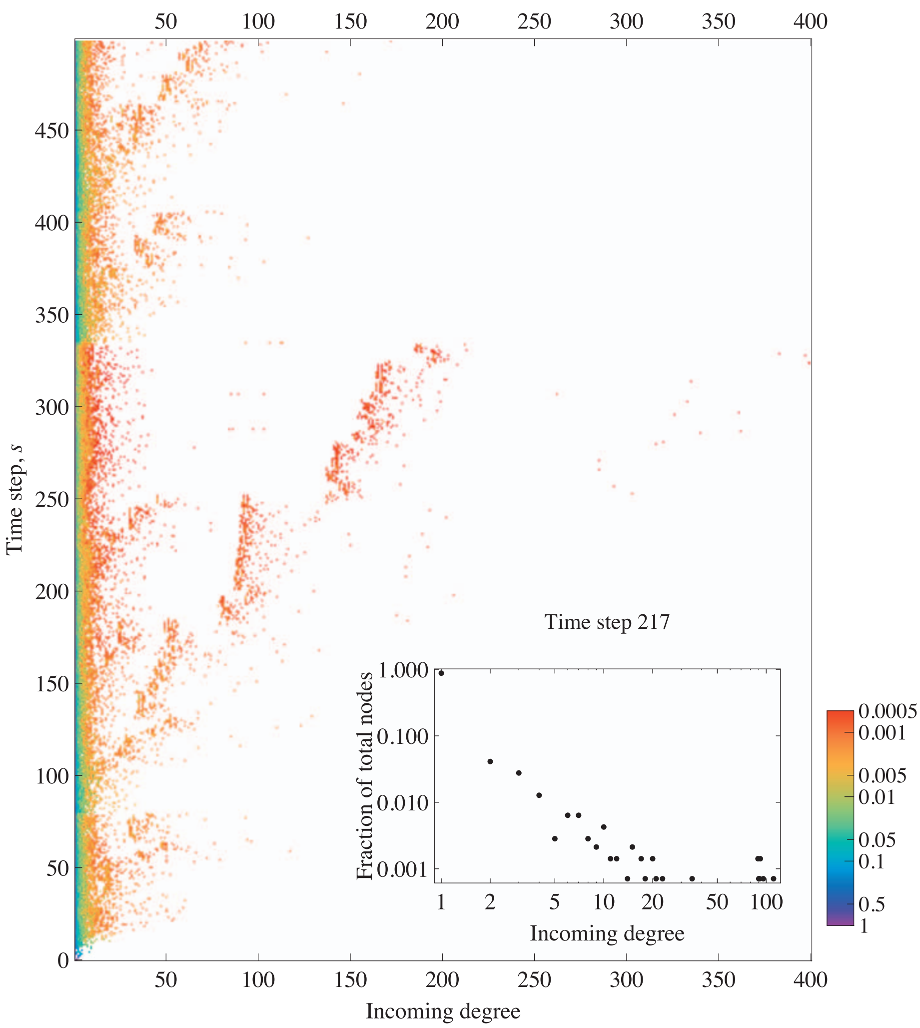


FIG. 3. (Color) Incoming degree distribution plot for rule 5 969 384 073 490 [Fig. 2(e)]. Colors represent the fraction of nodes in the network at each step possessing a given degree. Inset shows a "cross section" of the color plot, giving a log-log comparison of the network fraction possessing a given incoming degree at step 217.

works often meant the network was more prone to one type of behavior than another. However, none of these tools were able to routinely forecast the future evolution category of a given rule.

A preliminary observation of the underlying community structure of these networks was also performed, using the thirteen three-node motifs as presented by Milo *et al.* [19]. The time series of all thirteen motifs were examined for representative samples of rules from each category. The initial results suggest certain types of motifs or trends in their network total time series may be correlated with overall population dynamics, but a more rigorous study would be necessary in order to make a proper assertion.

### B. Analytical tools

#### *1. Incoming degree distribution*

Perhaps the most frequently examined property of networks is the degree distribution. In a directed network, a distinction is made between the incoming and outgoing degrees of a node. Here, the outgoing degree distribution was dismissed as trivial, since one of the basic properties of these networks is that each node is restricted to exactly two outgoing links. The degree distribution over an entire evolution was of interest rather than at a selection of individual steps. The distributions for individual steps were plotted in successive horizontal lines using color to denote the fraction of nodes which had a particular degree (see Fig. 3). Unfortunately, no correlations were seen between the type of incoming degree distribution and the eventual category of behavior.

Examination of the incoming degree distribution provided useful insight into the level of heterogeneity of a network. Observed degree distributions were diverse and ranged from very compact and homogenous [Fig. 13(b)] to apparently scalefree (inset, Fig. 3), sometimes even within the same evolution. This tool also proved to be useful in determining when a network was prone to sharp population drops. A greater number of nodes of high incoming degree implies fewer links remain for other nodes. Therefore, the existence of nodes of higher degree implies the existence of nodes at increased risk of being unreachable from $N_1^s$. Furthermore, nodes of greater incoming degree are somewhat more likely to increase in degree than nodes of lesser incoming degree, since the rewiring instructions in a rule are given in terms of reachable destinations in the network. This unbalanced dynamic may cause links contributing to nodes of lesser degrees to rewire to nodes of greater degree, which could cause

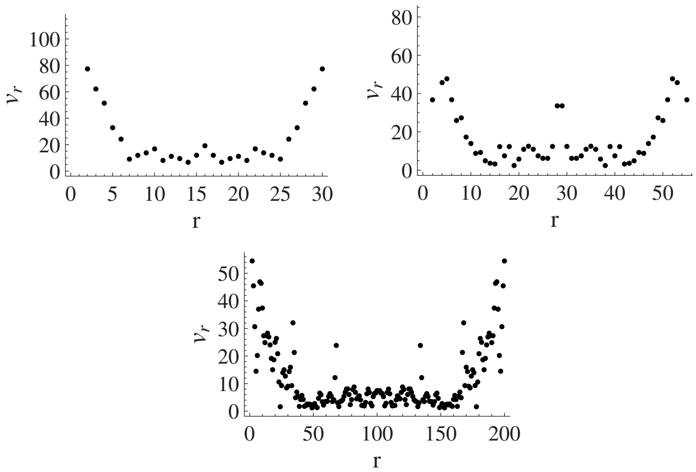


FIG. 4. Discrete Fourier transform spectra for three intervals in the transient phase of periodic rule 2 195 950 540 592 [Fig. 2(c)]. Figures compare steps (a) 1 through 30, (b) 1 through 55, and (c) 1 through 200.

deletion of fractions of the network when nodes of decreasing degree become unreachable from $N_s^1$.

#### 2. Fourier transform of population time series

The discrete Fourier transform (DFT) of the population time series was used in an effort to find a method of reliable identification of impending periodic behavior. The formula used to compute the discrete Fourier transform vector was given by

$$v_r = \frac{1}{\sqrt{t}} \sum_{s=1}^{t} n_s e^{2\pi i (s-1)(r-1)/t}, \tag{1}$$

where $v_r$ is the $r$th element of the resulting transform vector, $n_s$ is the $s$th element of the vector being analyzed (here, the population time-series), and $t$ is the length of $n$.

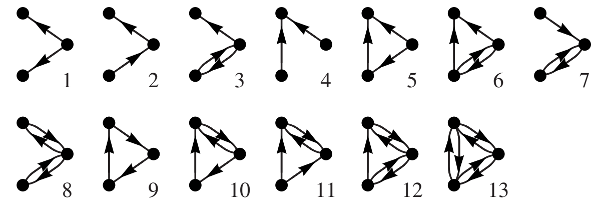


FIG. 6. The 13 three-node motifs as presented by Milo *et al.* [19]. Note that motif 13 is the initial condition used throughout this study.

Although the discrete Fourier transform performed well at recognizing periodic behavior after it had already become apparent in the population versus time plot, it displayed mixed success at predicting such behavior during the transient phase (see Fig. 4). This would seem to suggest that periodic behavior arises spontaneously rather than gradually when a rule by chance comes across a network configuration for which it is periodic. However, results from examining the clustering coefficients of evolutions [Eq. (2)] would suggest that dynamics between community structures within the network play a nontrivial role in overall network behavior and these community structures seem to appear gradually rather than spontaneously (see Fig. 5) . It would be worthwhile to evolve a catalog of periodic rules from varying initial conditions to see if some rules guarantee periodicity or if periodicity is dependent on initial condition.

The discrete Fourier transform did prove to be useful for evolutions where there was suspected complex behavior in conjunction with a monotonic growth trend. The DFT was able to detect periodic trends in these rules, where it is difficult to directly test for quasiperiodic behavior superimposed upon linear or exponential growth. There was one seemingly complex rule discovered in this study which may not be complex upon closer examination (Fig. 15); the DFT

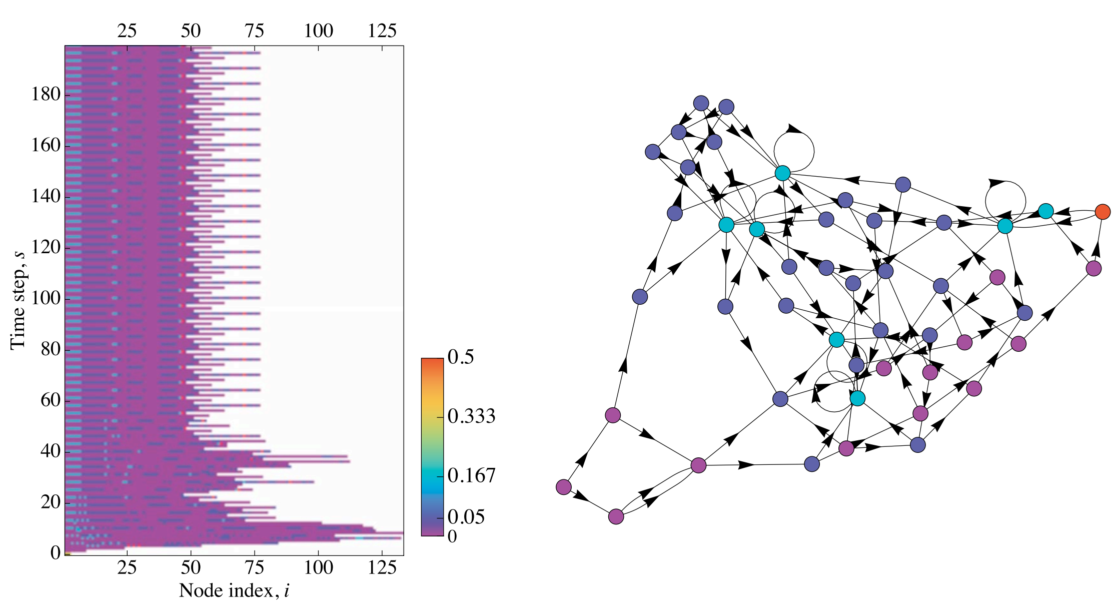


FIG. 5. (Color online) (a) Clustering coefficient per node and (b) diagram of the network at step 121 for periodic rule 2 195 950 540 592 [Fig. 2(c)]. Nodes in the network diagram are colored according to the value of the clustering coefficient at each location.

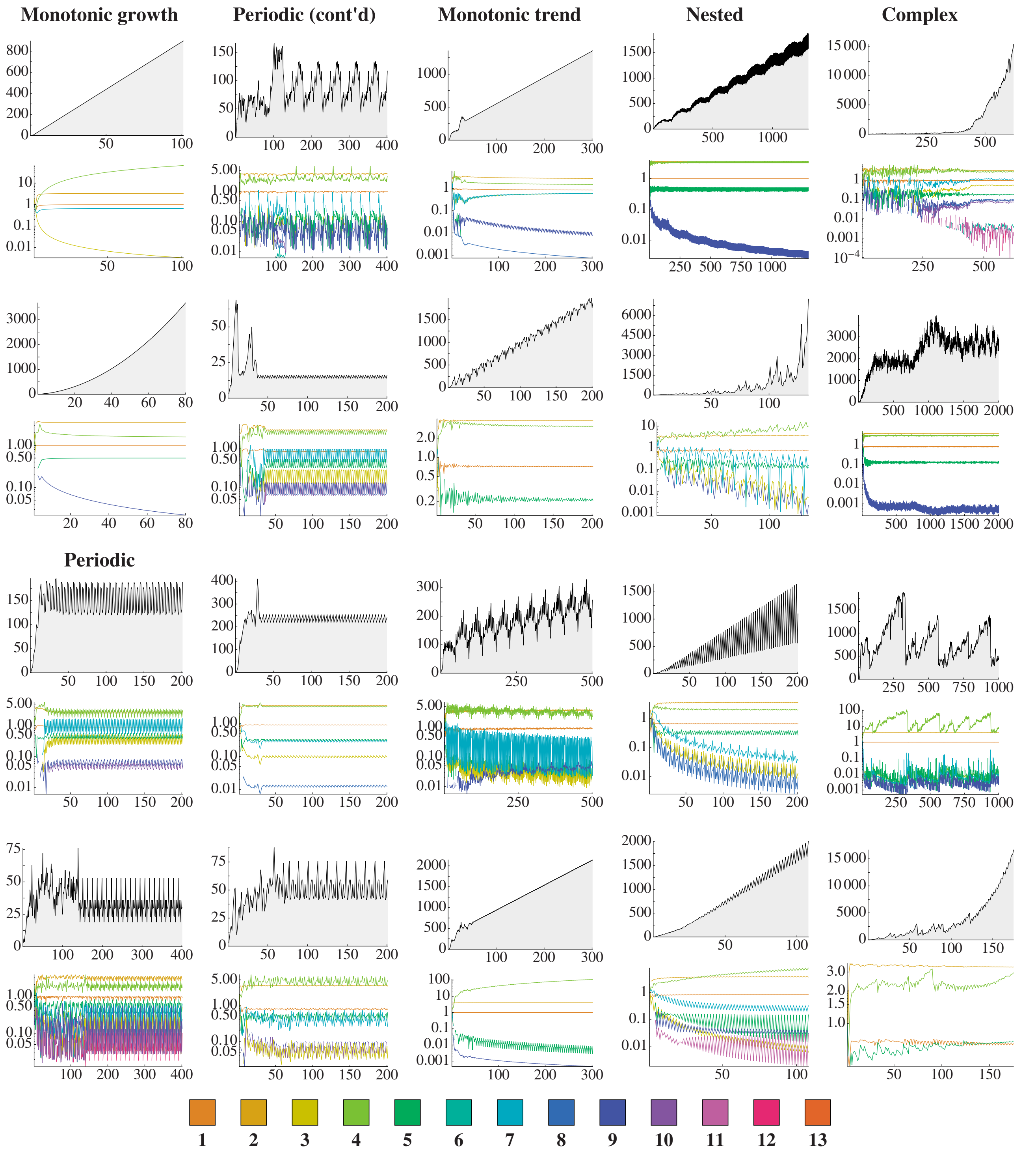


FIG. 7. (Color online) Examples of network motifs per capita. The 13 three-node motifs examined by Milo *et al.* [19] were here tallied at each step of a network evolution. These totals were all divided by the size of the entire network at the particular step to remove trends in motif totals due solely to an increase in the size of the network. Note that all motif plots are logarithmic along their vertical axes.

of the evolution suggests that regular behavior may not be readily seen upon simple visual inspection (Fig. 16).

### *3. Clustering coefficient*

The formula used to find the clustering coefficient of an individual node $i$ was a special case of the form proposed by Hansen *et al.* [20],

$$C_i(m,n) = \frac{p_i(m,n)}{\prod_{j=0}^{n-1}(T_i^n - j)}, \tag{2}$$

where $T_i^n$ is the total number of unique nodes reachable from node $i$ within a distance of $n$ and $p_i(m,n)$ is the total number of paths of length $n$ which travel from node $i$ to any node in

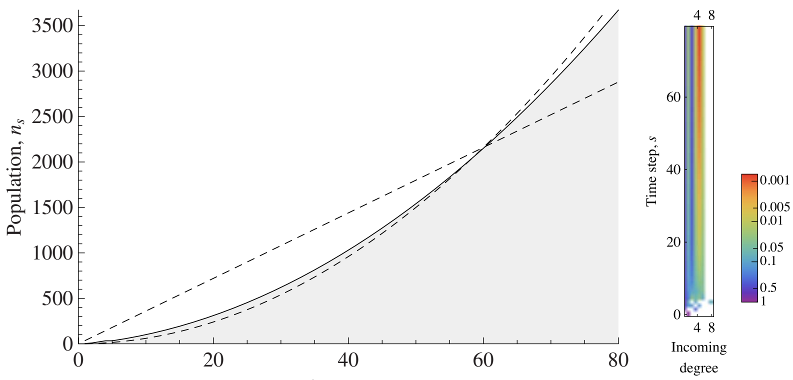


FIG. 8. (Color online) Monotonic growth example, rule 4 8364 37 700 860: (a) Population vs time graph, dashed linear and quadratic lines plotted for comparison. (b) Incoming degree distribution.

the set which node $i$ sees at exactly distance $m$ (i.e., if $m=2$, nodes seen at distance$=1$ are not included in the set). This is identical to Eq. 4 from [20], except $r$ is always equal to $n$ and $m \leq n$. To avoid dividing by zero, $C_i=0$ by default for any node where $T_i^n \leq (n-1)$. Although the clustering coefficient is usually examined in aggregate form, as the average over all individual nodes in a network [1,20,21], an alternative method was explored in this study. By examining the time series of solely the network average of the clustering coefficient, detailed information regarding the structure of the network is lost. In order to better understand the underlying structure and potentially uncover mechanisms signaling or even driving certain types of network growth behavior, the clustering coefficient [Eq. (2)] was examined as a distribution at each time step, similar to how the degree of a node is conventionally examined as a distribution instead of a network average [1]. This technique proved useful in gaining understanding of underlying network dynamics and even in predicting the frequency of successive levels of nesting in network evolutions corresponding to nested behavior (category 3).

Another factor contributing to the need for a more general approach was the restriction that each node in the present networks is limited to having at most two nearest neighbors (when looking *forward* through the network, following the directions of links). A significant advantage to using Eq. (2) for a measure of local network structure as opposed to the "classical" clustering coefficient [21] is the former's versatility. Whereas the classical clustering coefficient is only concerned with how many of node $i$'s nearest neighbors are also nearest neighbors themselves, the clustering coefficient employed here can be used to examine network structure over arbitrarily small or large distances [It is easily shown [20] that the classical clustering coefficient is a special case of Eq. (2)]. Furthermore, Eq. (2) is readily extendable to networks with directed links, whereas nontrivial specifications must be made before applying the definition of the classical clustering coefficient to directed networks [21], which may or may not take into account the actual reachability of particular nodes from the one at which the coefficient is being calculated.

#### *4. Network motifs*

A brief examination of the community structure in the presented networks was also conducted. The time series of the three-node motifs presented by Milo *et al.* [19] were examined over the course of a given network evolution (see Fig. 6). These series of values were divided by the size of the network at each corresponding step to remove trends in motif totals due solely to the trends of the entire network size.

Even this preliminary investigation into network community behavior displayed an interesting breadth of behavior (see Fig. 7 for examples). Motifs 1 and 2 generally scaled along with the network size as expected (note that these are represented by straight horizontal lines in the diagrams). Motif 4 frequently scaled with the size of the network, with some interesting exceptions, but was usually more volatile than 1 and 2 even when it did scale with the network size. Motif 6 was somewhat frequently spotted scaling with the size of the network as well. Motifs 12 and 13 (note that 13 is identical to the initial condition used throughout this study) were almost never observed.

The periodic evolutions displayed a wider spectrum of motif activity more frequently than evolutions from other categories. The monotonic trend rules displayed very smooth motif time series, whereas the time-series values from the other categories were more volatile. However, any observed regular behavior in the network population was consistently reflected in the behavior of the motif values. This seems to strongly suggest that regular population behavior of these networks is not an accident, but instead implies organized structure.

The motifs per capita time series displayed a variety of trends; sometimes the values would remain constant, some-

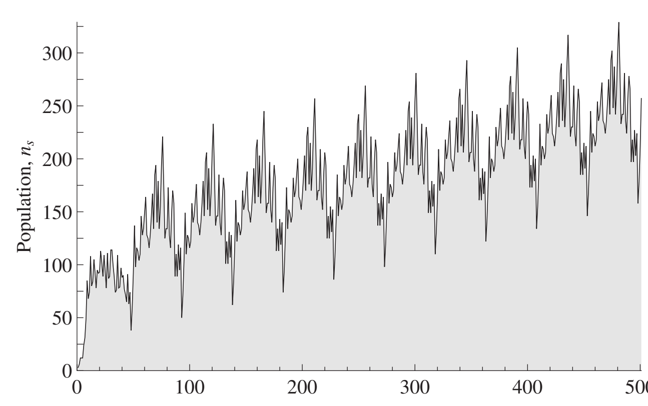


FIG. 9. Population vs time graph for monotonic growth trend example, rule 1 727 495 112 417.

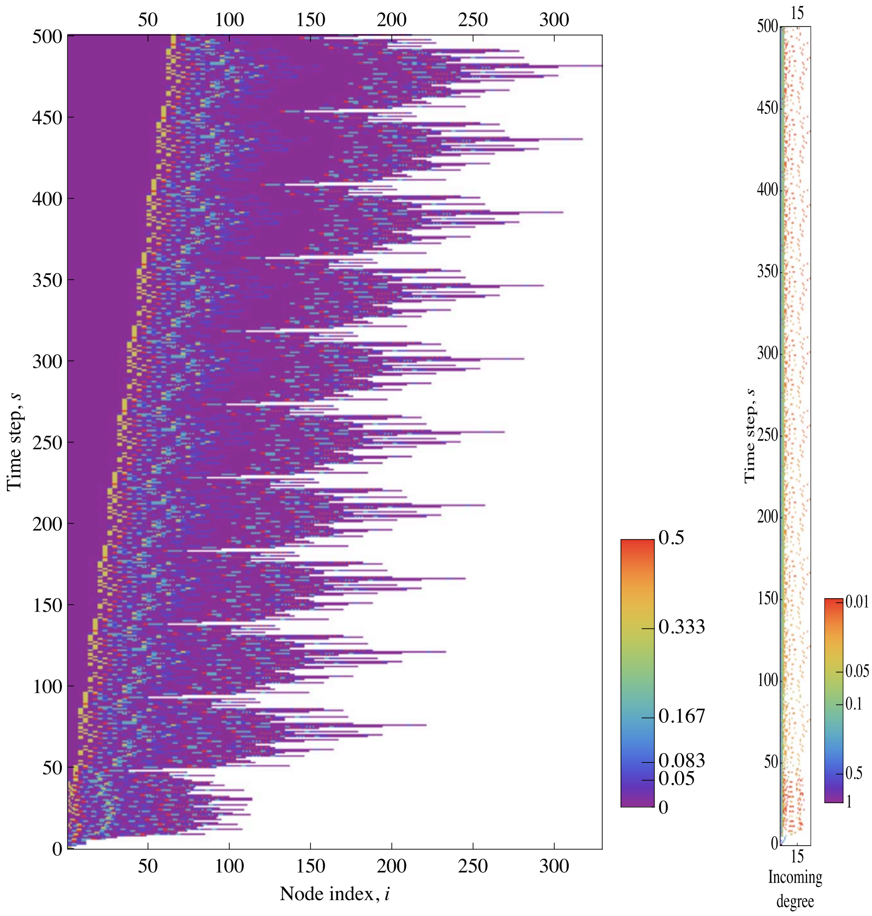


FIG. 10. (Color online) Monotonic growth trend example, rule 1 727 495 112 417. (a) Plot of clustering coefficient per node. (b) Incoming degree distribution

times they would decrease as the network grew (possibly implying a constant absolute number of motifs in the network), and in a few observed cases, the values would even increase. It could be argued that perhaps a motif which appears to correlate with growth in a given network evolution only appears as such because that motif is somehow indicative of one or more cases in the given rule that add nodes. But the cases of a rule that add nodes may not leave the acted-upon node seeing the same structure as it did before the new node was added or reproduce the previous structure for the new node. Put more simply, there are no guarantees that a given case of a rule will create more of itself or even preserve the already existing instances. So a case-occurrence count, if performed analogously to the presented motif count, might fluctuate rather unpredictably. Furthermore, these motif diagrams seem to suggest something more emergent in the behavior of these networks and it is probably more fruitful to think of them as indicators of the way information is shared across the network, rather than strictly related to specific rule cases.

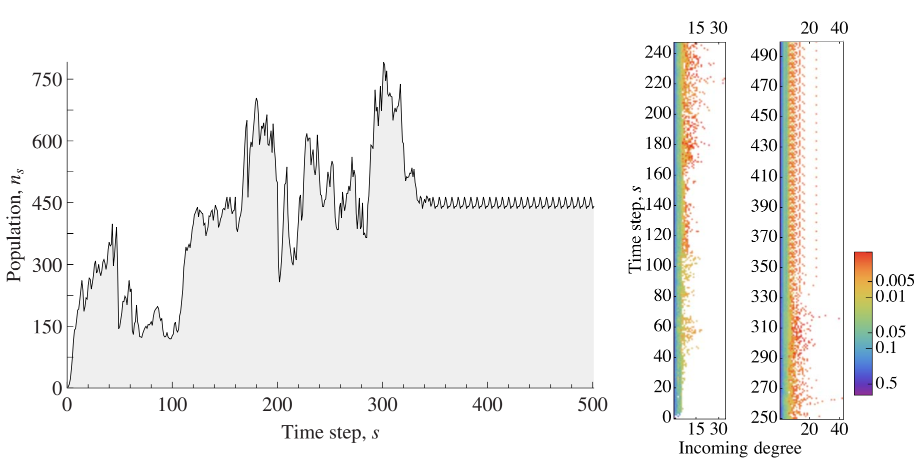


FIG. 11. (Color online) Periodic example, rule 7 427 685 516 255. At time step 346, the rule begins to exhibit period-six behavior. (a) Population vs time graph. (b) Incoming degree distribution.

## IV. OBSERVATION: PROPERTIES OF THE FOUR CATEGORIES OF BEHAVIOR

### A. Monotonic growth

Like many of the evolutions in the monotonic growth category, the example rule illustrated in Fig. 8 exhibits a markedly homogenous clustering coefficient, taking on only three values during the entire course of the above evolution: 0, $\frac{1}{20}$, and $\frac{1}{3}$. Also characteristic of this category of network growth, this example rule displays an incoming degree distribution that maintains a fixed width over time. Monotonic growth rates ranging from linear to exponential were observed. Although no examples of sublinear growth were found, there does not appear to be any property of the evolution method that would explicitly prevent such growth. The rule whose behavior is described by Fig. 8(a) was selected as an example for its interesting growth rate, which falls somewhere between linear and quadratic.

### B. Monotonic growth trend with superimposed regular behavior

The example rule represented in Fig. 9 displays clustering behavior typical of the monotonic trend rules encountered in this study. For rules in this category, the pattern of the clustering coefficient across individual nodes always exhibits a rather regular structure. Often these networks appear to be made up of several distinct communities, which either grow in diameter or maintain a steady size over the course of an evolution [see Fig. 10(a)]. It would certainly be worth investigating the potential causes for the stability of these subnetwork structures and the dynamics of their interactions.

Another commonly observed characteristic of this category, also seen in the monotonic growth rules, is the steady width of the incoming degree distribution [Fig. 10(b)]. As can be seen by the changing colors of the distribution—especially on the right-hand side—the fraction of the network which exhibits a certain incoming degree is changing, but the range of the distribution is not. It appears that the nodes being added to this particular network are not connecting to the pre-existing hub, but rather contributing to a more sparse, homogenous lattice—an observation supported by the growing diameter of partitions with clustering coefficient equal to zero [Fig. 10(a)].

### C. Periodic or nested/resonant behavior

#### *1. Periodic example: Rule* 7 427 685 516 255

The example rule illustrated in Fig. 11 displays the longest transient phase of the periodic rules covered in this study. At time step 346, the rule begins to display period-six behavior. Various attempts were made to find a metric which could predict this eventual periodic behavior. The discrete Fourier transform showed no significant peaks when examined over different time intervals within the transient phase, including steps 1–100, 1–200, 1–300, and 1–340.

The time series of the network average clustering coefficient [$\overline{C_i(1,2)}$, from Eq. (2)] was examined for all of the periodic rules in this study by comparing to a distribution of the same average for a random network of the same number of nodes and links. The averages during periodicity were split rather evenly between exhibiting a greater or lesser amount of clustering than a comparable random network; however, a trend was seen between the size of a network during periodicity and whether the average clustering was above or below that of a random network. Networks of larger average size (on the order of a few hundred nodes or more) were more likely to have a clustering coefficient consistently above that of a random network, while the clustering of networks of smaller average size (no more than 100 or 200 nodes) was more likely to be below that of a random network (Fig. 12).

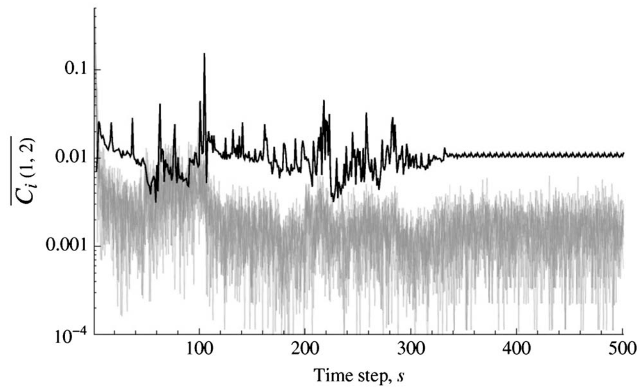


FIG. 12. Comparison of network average clustering coefficient for periodic rule 7 427 685 516 255 (black) and distribution of random networks (gray) with the same number of nodes and links.

#### *2. Nested example: Rule* 5 060 443 886 396

As compared to other categories of behavior, the nested rules display either very short transient phases or virtually none at all. Therefore, prediction of nested behavior was not a significant focus and instead, the underlying mechanisms causing such behavior were of interest. The rule illustrated in Fig. 13(a) was chosen as an interesting example for this subcategory due to its unexpected jumps, seeming phase transitions in growth trend rate.

As with the other rules in this category of network growth, this example rule displays a globally regular struc-

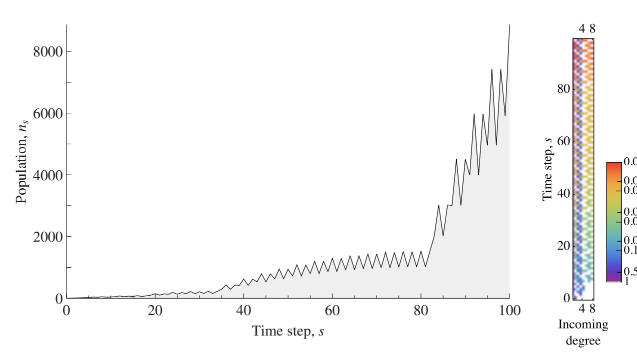


FIG. 13. (Color online) Nested example, rule 5 060 443 886 396. (a) Population vs time graph. (b) Incoming degree distribution.

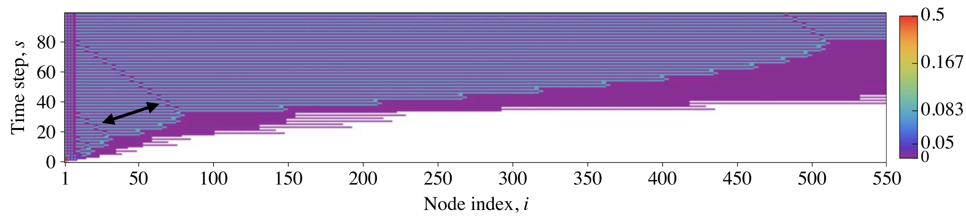


FIG. 14. (Color online) Section of the clustering coefficient diagram for nested rule 5 060 443 886 396. An apparent motif in the network signals the jumps in growth rates seen in the population time series. After six steps, the network displays only two values of clustering coefficient: zero and one-sixth.

ture in its clustering coefficient plot. This is similar to what is seen for the monotonic trend rules [Fig. 10(a)], but whereas the monotonic trend rules display regular structure within sections of the network, the nested rules display an overall coherent pattern with little apparent partitioning. However, unlike many of the other rules in this subcategory, the evolution of this example rule displays a single, small network motif which appears to work itself through the network and cause—or at least signal—the observed growth trend transitions (Fig. 14).

### D. Complex behavior

The main property of the complex rules is that they do not exhibit the strictly regular behavior of the other categories. Therefore, Fig. 15 shows the evolution of a rule which at first looks complex, but upon closer inspection may in fact belong to the monotonic trend category (shown in Fig. 9). It appears that this example rule exhibits complex behavior superimposed on a linear growth trend, but examination of this rule through the various tools here presented suggests that the superimposed behavior may actually achieve a regular nature, as found in the monotonic trend rules, sometime after step 300.

The discrete Fourier transform of rule 2 854 763 462 287 uncovers a regular dynamic in the growth behavior, which is apparent after 300 steps (Fig. 16). A review of the clustering coefficient diagram for the evolution of this rule also displays signs of regular behavior, showing some homogenous patterns similar to those found in other linear monotonic trend rules. There also appears to be a section within the diagram that displays an irregular pattern, much like the other complex rules, yet judging from the DFT, it is likely this may be a regular pattern as well that is simply not obvious to the eye.

## V. DISCUSSION

The flexibility in the rule specification potentially enables modeling of real-world networks, particularly those concerned with directional traffic of goods or information across a network, such as the natural distribution systems presented in Banavar *et al.* [22]. Since nodes in these networks are allowed to have more than one link pointing to the same destination, this method of network evolution may not be of use in modeling some types of social networks, where it would be nonsensical to imply that a person can have more than one link or, say, friendship, with another person. (Of course, it may be possible to create another method of network evolution in a similar spirit—with a constant set of rules concerned with local structure—that places the necessary restriction on the destinations of links.) Where the presented model can be of use is in modeling real-world networks free of this link restriction, including networks that have conventionally been modeled using the single link restriction, but whose underlying real-world existences do not necessarily require such a restriction. One such example is the network of connections among airports, which has often been presented assuming the single link restriction. However, a single directed link in an airport-to-airport network does not represent a single connection or flight. In fact, it represents multiple flights, likely at different departure times and by different airlines. The single link restriction results in an overly compressed representation of such real-world networks, since the dozens of daily flights from one major airport to another will be represented identically to the perhaps half-dozen daily flights from a small regional airport to a major airport. Such networks could be viewed equally (if not more) fruitfully by allowing for multiple links with the same origin and destination nodes.

The presented network evolutions display the capacity for scale-free degree distributions (Figs. 3 and 17), a rich variety of local community structure (Fig. 7), and in some instances, the small-world property (Fig. 18). Searching the above-presented rule space, or ones like it, could provide closer

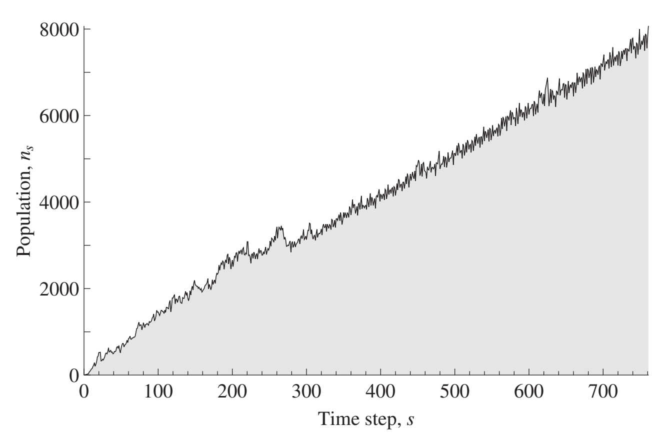


FIG. 15. Population vs time graph for unclassified rule 2 854 763 462 287. Visually, the rule appears to exhibit a linear growth trend with superimposed irregular behavior. However, examination of the discrete Fourier transform of the population time series suggests the superimposed behavior may be periodic in nature.

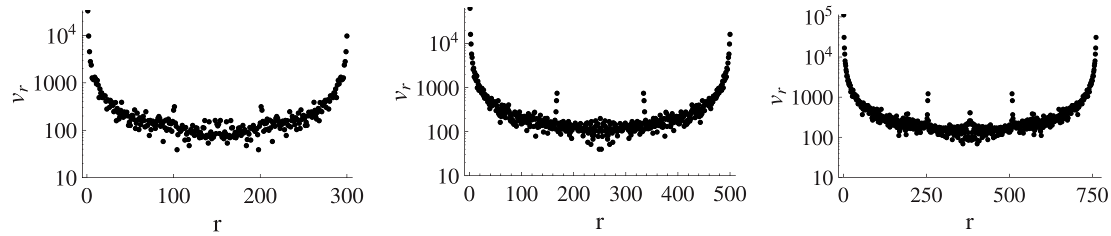


FIG. 16. Logarithmic plots of DFT for rule 2 854 763 462 287 for steps 1 through 300, 500, and 761, respectively. Examination of the peaks implies the influence of some period-three behavior in this rule evolution.

models for many real-world networks—along with better understanding of the forces that drive these networks when the specific cases of the underlying rule are more closely examined and translated into the particular real-world context. In addition to the presented class of rules, model searches could be conducted in rule spaces where the nodes have a different number of outgoing links or even where the number of outgoing links is variable. (The style of rule cases would have to be redefined, but there is nothing explicitly prohibiting the exploration of these alternate spaces of possible rules.) Even the order in which nodes are updated could be altered, if so desired. It is worth noting that the flexibility found in this method of network growth results in the drawback that exhaustive searches of rule spaces will often be insurmountably cumbersome. But when seeking to model or even simply better understand a real-world, dynamic network, it is perhaps better to have too many options rather than too few.

Apart from direct applications of this style of network evolution, the tools employed in their analysis can support better understanding of dynamic networks in general, particularly the use of the clustering coefficient as a distribution instead of a network average. By examining spaces of networks updated according to simple rules, one can gain intuition about network dynamics, especially those which operate locally. Also, there is certainly potential for theoretical studies of this method of network evolution, perhaps in search of an analog for Langton's lambda parameter [23] for ordinary cellular automata or other interpretations of the rules and the impact they have upon the evolution of a network.

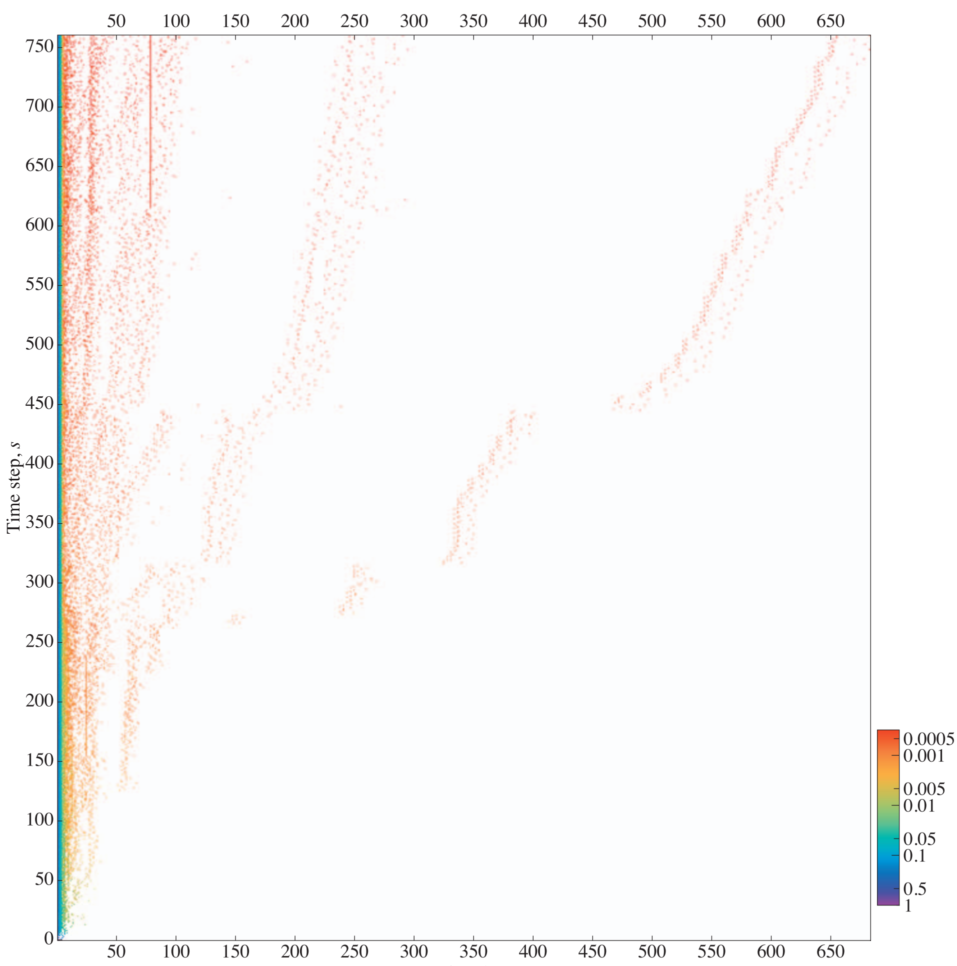


FIG. 17. (Color) Incoming degree distribution for rule 2 854 763 462 287.

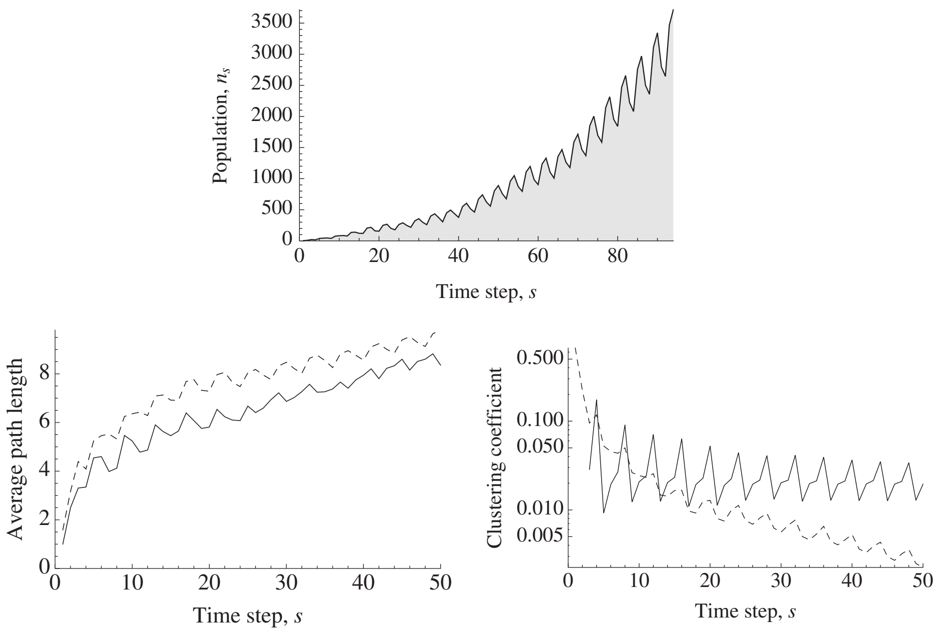


FIG. 18. Network displaying "small-world"-like properties over the course of evolution of rule 2 005 240 550 289. (a) Population vs. time. (b) Comparison of average path length between reachable nodes in evolved rule (solid) to expected value of average path length in a random, undirected network with equal numbers of nodes and links (dashed). (c) Comparison of average clustering coefficient of evolved rule (solid) to expected value of clustering coefficient for a corresponding random network (dashed).

Finally, it is clear that a more detailed study of the underlying community structure [24] of these networks is needed, both for a better understanding of the behavior found in this model and for further insight into the effects of dynamical structural processes in real-world networks. For example, when people talk about the "growth" of the internet, they are really referring to its *aggregate* growth. The internet is made up of many types of communities and groups, differing not only in their subject matter, but also in their purpose and utility. It would also not be surprising if these different subsets of the internet had different internal link structures as a product of their differing function [19]. However, all of these unique communities are interconnected and collectively produce the overall growth or decay of the entire network. Examining how intercommunity dynamics give rise to the aggregate growth of a network will likely provide greater insight into the underlying causes of observed real-world network behavior. A deterministic setting, such as the network evolution model here presented, would be ideal grounds for studying such intercommunity dynamics.

## ACKNOWLEDGMENTS

The authors would like to acknowledge support from NASA Phase II Grant No. NNG 06GE87G to the Vermont Advanced Computing Center and a NSF-EPSCoR Complex Systems grant. The basis for the research discussed was performed by the lead author as part of an undergraduate New Kind of Science summer school program. The authors would also like to thank Peter Dodds for helpful conversation as well as three anonymous reviewers for valuable criticisms.

## APPENDIX: RULE ENUMERATION AND APPLICATION OF CASES

### 1. Distance one rules

Rules are enumerated separately for distances one and two. For distance one, there are 36 possible sets of instructions and two cases which can have any set of instructions regardless of what is assigned to the other. Therefore, there exist $36^2 = 1296$ possible rules and each rule is assigned to an index ranging from 0 to 1295. The specific instructions described by a particular rule are then found as follows:

(1) The index is expressed in its base-six representation, with trivial zeros added to the left, if needed, to make a total string of four digits.

(2) The digits in this representation are then replaced with expressions according to the following pattern:

$$0 = \{1\} \; 2 = \{\{1\},\{1\}\} \; 4 = \{\{2\},\{1\}\},$$

$$1 = \{2\} \; 3 = \{\{1\},\{2\}\} \; 5 = \{\{2\},\{2\}\},$$

where a "{1}" represents following the current up link and "{2}" represents the down link. The nested expressions designate the addition of a new node, whose two outgoing link destinations correspond to the directions given in the two elements of the list.

(3) The four resulting expressions, one for each digit in the base-six representation, are then organized into the actual instructions for the rule. The first and second expressions define case 1 of the rule, corresponding to instructions for a node's up and down links, respectively. The second two expressions define case 2 of the rule.

For example, a distance-one rule may take the form [rule 219, Fig. 19(a)]

$$\{1\} \to \{\{2\},\{1\}\}, \; \{2\} \to \{\{1\},\{\{1\},\{2\}\}\},$$

where the elements before the arrow denote which case of the rule they represent. The first position in the lists after the arrows gives the instructions for the up link and the second position gives instructions for the down link. This example rule specifies that if the node being updated has only one neighbor (corresponding to case 1), its up link is to be redirected to where the down link currently goes and its down link is to be rerouted to where the up link currently goes. For

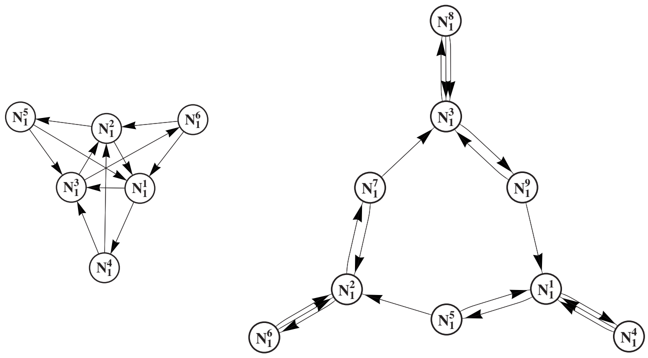

FIG. 19. One step in the evolution of two different example networks, each evolved from an initial condition of three nodes (see Fig. 1).

case 2, the up link remains at its current location and the down link is instructed by the nested expression to create a new node and link to it. The elements in the nested expression then provide the destinations for the new node's links, which are interpreted in the same manner as already seen—{1} means follow the up link and {2} means follow the down link. Also as before, the first position in the nested list provides the instructions for the new node's up link, while the second position provides the instructions for its down link. Here, the up link of the new node will be wired to the same destination as the up link of the original node and the down link will be wired to the same destination as the original node's down link.

### 2. Distance two rules

For distance two, rules are assigned an index ranging from 0 to 9 682 651 996 415. Here, there are 1764 possible sets of instructions and four cases to a rule, resulting in $1764^4$=9 682 651 996 416 possible rules. The extraction of specific instructions from the index number is similar to that of distance one, although different integer bases are required. This process is outlined below and accompanied by the example expansion of rule 5 060 443 886 396 [Fig. 19(b)].

(1) The index is expressed in its base-1764 representation, adding trivial zeros to the left, if necessary, to make a total string of four numbers.

*Example:* {921, 1621, 901, 1592}.

(2) Each number in the string is then broken down into its respective representation in base 42 (again adding zeros to make a list of length two for each of the four numbers).

*Example:* {21,39}, {38,25}, {21,19}, {37,38}.

(3) Since there are four cases in a distance two rule (see Sec. II), each of these four pairs of numbers is then assigned to a particular case. The first pair is assigned to case 2|3, the second to case 4, and so on.

$$\text{Example: } \{2|3\} \to \{21,39\},$$
$$\{4\} \to \{38,25\},$$
$$\{\{5\} \to \{21,19\}\},$$
$$\{\{6\} \to \{37,38\}.$$

(4) Each argument in each pair is then replaced by its base-six representation (adding zeros if necessary to make a string of three digits). The resulting three-digit strings give the directions for a specific link in a specific case. Depending on whether it begins with a 1 or 0, one or two of the digits are selected to expand into explicit directions:

$\{1,0,x\}=x$ (will become instructions which preserve the number of nodes)

$\{0,y,z\}=\{y,z\}$ (will become instructions which add nodes).

| | | | |
|---|---|---|---|
| *Example:* | {2\|3}→{{0,3,3},{1,0,3}} | ⇒ | {2\|3}→{{3,3},3} |
| | {4}→{1,0,2},{0,4,1}, | | {4}→{2,{4,1}}, |
| | {5}→{0,3,3},{0,3,1}, | | {5}→{{3,3},{3,1}}, |
| | {6}→{{1,0,1},{1,0,2}} | | {6}→{1,2} |

(5) The digits in these remaining strings are then expanded as follows:

$$0=\{1\} \quad 2=\{1,1\} \quad 4=\{2,1\},$$

$$1=\{2\} \quad 3=\{1,2\} \quad 5=\{2,2\}.$$

The final form of this example rule is then

$$\{2|3\} \to \{\{\{1,2\},\{1,2\},1,2\}\},$$

$$\{4\} \to \{\{1,1\},\{\{2,1\},\{2\}\},$$

$$\{5\} \to \{\{1,2\},\{1,2\},\{1,2\},\{2\}\},$$

$$\{6\} \to \{\{2\},\{1,1\}\},$$

where strings of length two are now seen, since in distance two rules the new link destinations can be indicated by any location reachable at distance one *or* two from the "active" node. The instruction strings of length two are to be understood in terms of *successive* destinations in the network, so an instruction of {1, 2} denotes the destination reached by following the up link from the original node, then following the down link from that subsequent node.

In this example rule, nodes will be added in cases 2|3 and 4 and two nodes will be added if case 5 is applied, one in the place of each link. The only instances of rewiring without adding a new node occur for the up links of cases 4 and 6 and for the down link of case 2|3. Since nodes are apparently added so frequently in the application of this rule, one might expect to see nothing but monotonic growth arise from this rule. But in fact what is seen is more complex than simple monotonic growth [see Fig. 13(a)].

[1] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[2] M. Newman, SIAM Rev. **45**, 167 (2003).
[3] D. Watts and S. Strogatz, Nature (London) **393**, 440 (1998).
[4] D. J. Watts, *Six Degrees: The Science of a Connected Age* (W.W. Norton & Company, New York, 2004).
[5] S. Bernhardsson and P. Minnhagen, Phys. Rev. E **74**, 026104 (2006).
[6] C. Moore, G. Ghoshal, and M. E. J. Newman, Phys. Rev. E **74**, 036121 (2006).
[7] L. G. Morelli, Phys. Rev. E **67**, 066107 (2003).
[8] H. Štefančić and V. Zlatić, Phys. Rev. E **72**, 036105 (2005).
[9] C. Bedogne' and G. J. Rodgers, Phys. Rev. E **74**, 046115 (2006).
[10] J. Ozik, B. R. Hunt, and E. Ott, Phys. Rev. E **69**, 026108 (2004).
[11] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[12] S. Jung, S. Kim, and B. Kahng, Phys. Rev. E **65**, 056101 (2002).
[13] S. Dorogovtsev, A. Goltsev, and J. Mendes, Phys. Rev. E **65**, 066122 (2002).
[14] A.-L. Barabási, E. Ravasz, and T. Vicsek, Physica A **299**, 559 (2001).
[15] E. Ravasz and A.-L. Barabási, Phys. Rev. E **67**, 026112 (2003).
[16] Z. Zhang and L. Rong, Physica A **363**, 567 (2006).
[17] E. Estrada, Phys. Rev. E **75**, 016103 (2007).
[18] S. Wolfram, *A New Kind of Science* (Wolfram Media, Champaign, IL, 2002).
[19] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, Science **298**, 824 (2002).
[20] H. Hansen, C. Andresen, and A. Hansen, Europhys. Lett. **78**, 48005 (2007).
[21] G. Fagiolo, Phys. Rev. E **76**, 026107 (2007).
[22] J. R. Banavar, A. Maritan, and A. Rinaldo, Nature (London) **399**, 130 (1999).
[23] C. G. Langton, Physica D **42**, 12 (1990).
[24] M. E. J. Newman and M. Girvan, Phys. Rev. E **69**, 026113 (2004).