Networks, Dynamics, and Modularity

Evan A. Variano
Department of Civil and Environmental Engineering, Cornell University, Ithaca, New York 14853, USA

Jonathan H. McCoy*
Department of Physics, Cornell University, Ithaca, New York 14853, USA

Hod Lipson
Department of Mechanical and Aerospace Engineering, Cornell University, Ithaca, New York 14853, USA
(Received 27 May 2003; revised manuscript received 13 November 2003; published 7 May 2004)

The identification of general principles relating structure to dynamics has been a major goal in the study of complex networks. We propose that the special case of linear network dynamics provides a natural framework within which a number of interesting yet tractable problems can be defined. We report the emergence of modularity and hierarchical organization in evolved networks supporting asymptotically stable linear dynamics. Numerical experiments demonstrate that linear stability benefits from the presence of a hierarchy of modules and that this architecture improves the robustness of network stability to random perturbations in network structure. This work illustrates an approach to network science which is simultaneously structural and dynamical in nature.

DOI: 10.1103/PhysRevLett.92.188701 PACS numbers: 89.75.Fb, 45.30.+s, 87.23.Kg, 89.75.Hc

In the science of complex networks, there is a longstanding problem concerning the identification of general principles relating network structure to dynamics [1]. However, as several members of the physics community have pointed out recently [1,2], research efforts have for the most part been focused on structural properties of networks and the evolution of these properties. The structural principles identified are often fascinating and suggestive. Prominent examples include small-world [3] and scale-free architectures [4] and community and modular structure [5] arising in social, biological, and technological networks. Some interesting work has been done connecting these structural properties to network function. It has been shown, for instance, that "shortcut" links generate phase transitions in the communication of information and disease in a small-world network [3,6] and that scale-free networks possess a surprising robustness to random failures [7]. Unfortunately, it is not clear how to address general questions about dynamics on networks; this may account for the apparent shortage of tractable questions about the general principles relating structure and dynamics.

We propose that linear dynamics on a network provides a natural starting problem. A network of n nodes, labeled 1 through n, with directed, weighted links can be represented by an n-by-n matrix M in a straightforward manner: the matrix element M_{ij} defines the edge weight of the directed edge connecting node i to node j, with a weight of 0 indicating the absence of that particular edge. M also, however, defines a linear dynamical system, $\dot{\mathbf{x}} = M\mathbf{x}$. If we use x_i to represent a state variable associated with the ith node, this correspondence allows us to think of the matrix M as simultaneously representing both a directed, edge-weighted network and a linear dynamical system; in this way, the entries of M naturally acquire both structural and dynamical significance. This viewpoint carries with it many interesting and important questions about how dynamical features of the network are related to its topology. For instance, which topologies support stable linear dynamics? Since stability is extremely rare in large, random matrices [8], we must expect that special structural features accompany linear stability in a network dynamical system. What are these features? What does a large, linearly stable network look like? This question is of broad importance, even for real networks exhibiting nonlinear dynamics, since the theory of local dynamics around equilibria in a nonlinear system relies on a solid understanding of linear dynamics.

In this Letter, we report the emergence of *modularity* and *hierarchical organization* in evolved networks supporting asymptotically stable linear dynamics. Unlike other forms of modularity and hierarchical organization reported recently [5], the architecture we present depends crucially on the network having directed edges (directional links provide an important source of complexity in real networks [1,9]). Furthermore, since the elements of M have simultaneous structural and dynamical significance, this modular architecture is mathematically manifest in the dynamics in a precise manner. A number of authors have conjectured about the intertwining of modularity and stability in various contexts [8,10]—our work provides straightforward, concrete support for these conjectures. Numerical experiments demonstrate that linear stability benefits from the presence of a hierarchy of modules and that this architecture improves the robustness of network stability to random perturbations in network structure.

We cannot get a good sample of linearly stable networks by random search, but a standard genetic algorithm (GA) allows us to evolve a sample in a straightforward manner. We begin with a population of 1000 25-node random networks. These networks are generated by drawing a directed edge from node i to node j (where j can equal i) with probability p and randomly choosing edge weights from a uniform distribution on $[-1, 1]$. We choose $p = 0.1$ so the networks are sparse but well above the giant-component transition [1]. The algorithm consists of the following steps: First, it assigns a fitness value (defined below) to each network in the population. Next, it randomly selects pairs of parents, with an individual selected a number of times proportional on average to that individual's fitness relative to the rest of the population. Then, it recombines the attributes of each pair of parents to produce a pair of offspring. Recombination is achieved by swapping the values in a rectangular region of one parent matrix with the corresponding values in its "spouse" matrix. The rectangular region is chosen by randomly selecting corners such that all matrix entries are equally likely to be included. Finally, the offspring replace the parents, producing a new population of networks and the above steps repeat. We sample and study the stable individuals at various generations.

By basing our fitness function on linear stability, we interweave the dynamics on the network with evolution of the network structure. Our fitness function rewards a network for each eigenvalue with a negative real part. Instead of adjusting the magnitude of its leading eigenvalue, each network evolves to become more stable by reducing the dimension of, and eventually eliminating, the unstable manifold of its equilibrium point. Figure 1 shows the evolution of stability by plotting the average fitness of the population for successive generations. Individuals of fitness 1 are asymptotically stable and, thus, the evolution of populations with average fitness near 1 indicates that our GA has indeed found many large, linearly stable networks. The average Hamming distance between two networks in a population provides a measure of diversity. Because this measure decreases slowly while fitness increases rapidly, the population remains diverse; our results are not due to the structure of a single network infecting the population.

At first glance, the evolved networks look no different than the random networks comprising the initial population—they display similar degree distributions and similar densities (where density is defined to be the fraction of nonzero entries in M). However, these similarities belie important differences. As evolution progresses and the population becomes more fit, individuals show increased modular structure, in two senses. First, the network breaks into *components*. Nodes i and j are in two separate components if and only if there exists no directed path connecting i to j in either direction. Second, within each component, a number of *strongly connected components* (SCCs) emerge. Nodes i and j are in the same SCC if and only if there exists a directed path from i to j *and* from j back to i (thus, i and j are part of the same directed cycle). Two nodes are in different SCCs but the same component if there exists a directed path from one node to the other but not vice versa; such paths include edges that link nodes from different SCCs. These cross-SCC connections by definition cannot form directed cycles. Thus, they define a hierarchy of modules, represented in Fig. 2.

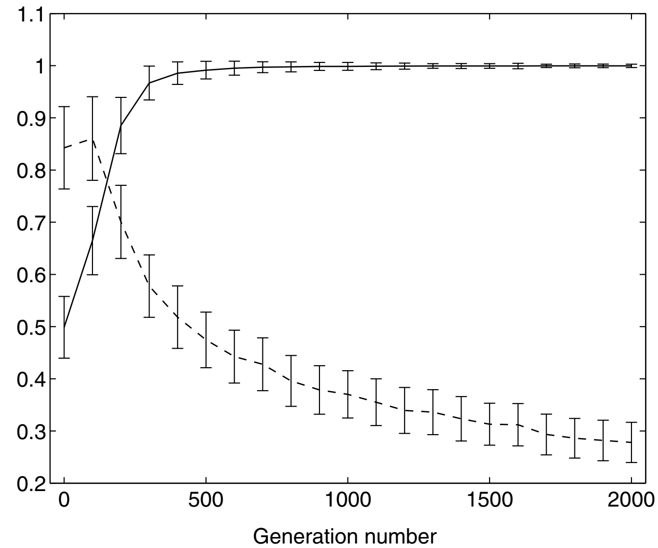

FIG. 1. Average fitness of population (solid line) and average Hamming distance between members of the population (dashed line, scaled to fit axes) as evolution progresses. These curves demonstrate that our algorithm successfully evolves a large, diverse population of asymptotically stable networks.

We observe in Fig. 3(a) that the number of components increases and the average component size decreases as evolution progresses. Most of the change occurs in the earlier generations, where the average fitness rises steeply from one generation to the next. The earlier generations exhibit a bimodal distribution in component sizes, dominated by large components but with a significant presence of tiny components as well. As the average fitness of the population rises, this distribution remains bimodal but the tiny component peak emerges as dominant. At no time

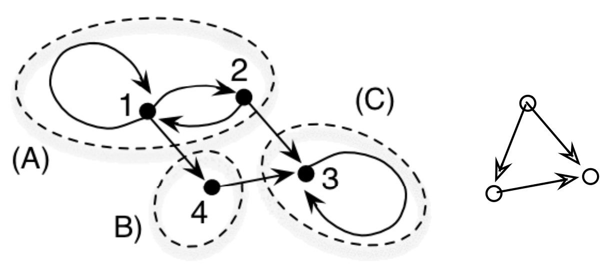

FIG. 2. Decomposition of a network into a hierarchy of strongly connected components (SCCs): The network on the left is partitioned into three SCCs by dotted curves. The connectivity of these SCCs is represented by the network on the right; note the absence of directed loops in this network.

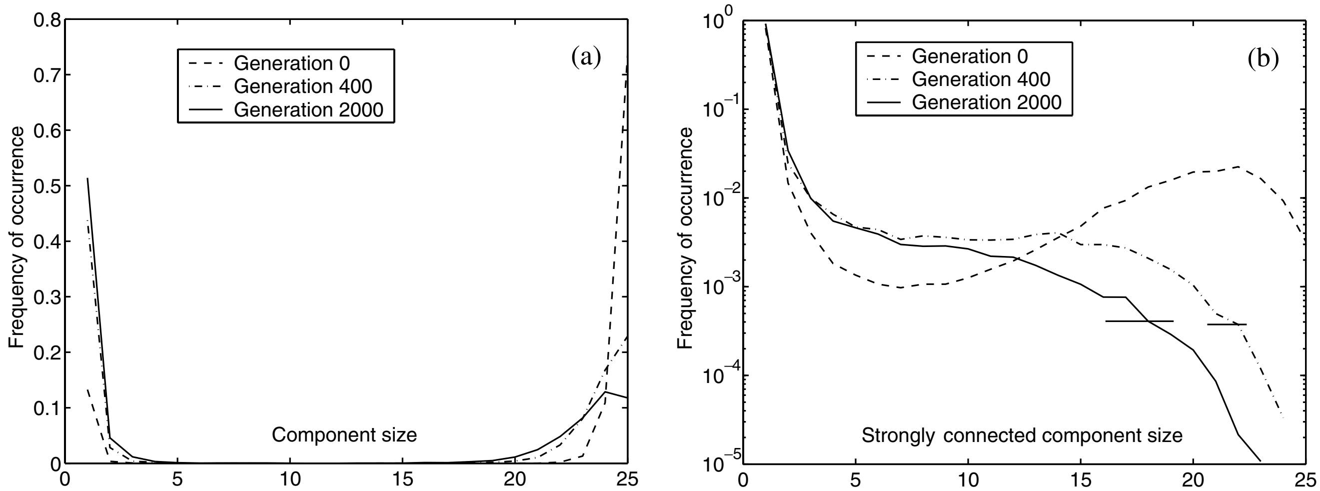

FIG. 3. Distribution of component sizes (a) and SCC sizes (b) in evolved populations. Leftward motion of the probability indicates increased modularity in the stable networks. Results are averaged over five independent runs; error bars indicate typical horizontal spread. Note that the evolved networks retain a giant component.

do we observe any medium-size components. Similarly, in Fig. 3(b), we observe that the number of SCCs increases and the average SCC size decreases, even if we restrict our analysis to those networks consisting of only a single component. Again, the earlier generations exhibit a bimodal distribution in SCC sizes, except here the small SCCs already dominate and medium-size SCCs are plentiful. As the average fitness of the population rises, the weight of this distribution moves to the left until the large SCC mode has vanished. All of the above trends indicate a motion towards increased modularity in the evolved networks. Repeating the experiment with $N = 50$ and also with a completely different (nonevolutionary) algorithm, we see qualitatively similar results for networks of similar density.

Two control experiments use alternative fitness functions to ensure that the emergent structure is not an artifact of the GA. In the first, we eliminate our fitness function entirely by assigning equal fitness to each individual in every generation. In the second, we observe the consequences of assigning random fitnesses from a uniform distribution. In neither case do we observe *any* evolution away from the characteristics of the initially random networks. These control experiments demonstrate that our results cannot be explained as an artifact of recombination alone or recombination plus random drift.

We observe a decrease in network density as evolution proceeds towards stability. As the following control experiment demonstrates, this "sparsifying" or "pruning" effect is clearly a by-product and not the mechanism by which evolution achieves stability. Starting from a population of random networks, we remove edges at random until the population has a density distribution identical to that of stable networks evolved from the same initial population. These randomly pruned networks are no more fit than random networks; i.e., they are not stable. Furthermore, they are far less modular than evolved networks: typical evolved networks have 18.72 ($\pm$ 1.40) SCCs, while typical randomly pruned networks have only 7.63 ($\pm$ 0.35). Because pruning alone yields neither the stability nor the modularity we observe in evolved networks, our GA must be *reorganizing* and not just pruning. In other words, network stability and modularity are affected by both the presence and the *placement* of zeros in M. It is interesting and important to recognize that, as the networks evolve, they reorganize to become more hierarchical without disconnecting.

We have demonstrated that evolution uses hierarchical modularity as a strategy for building linearly stable networks. Since it is so much easier to make small stable networks, larger stable networks are often best built by combining smaller ones in a modular hierarchy. This works because the eigenvalues of a matrix are precisely the eigenvalues of its SCCs—the linear stability of the network as a whole is determined by that of its modules and, hence, connections between distinct modules do not impact the linear stability of those modules. In other words, a hierarchy protects the modules; when small networks are combined at random, we find that the modules tend to coalesce and the result is unlikely to be stable.

Considering how sensitive stability is to the details of network structure, it might be surprising that two stable matrices can produce stable offspring by recombination. That is, there is no reason to expect that a stable matrix would remain stable after random replacement of some of its entries. On the other hand, this is precisely what we observe—the evolved networks are *structurally stable*, i.e., their dynamics are robust to structural mutations [8]. This property emerges naturally as evolution proceeds, since the networks which are not structurally stable produce unstable offspring which are less likely to reproduce. Thus, the genetic material of structurally unstable networks tends to be weeded out of the population. To quantify this trend, each stable network is subjected to a perturbation in which a randomly selected entry in M is replaced with a new value randomly selected from a uniform distribution on $[-1, 1]$ (a rude intrusion which could seriously alter the eigenvalue spectrum of M). At generation 200, a typical stable network is only 7.7% ($\pm$ 2.4%) resistant to mutation, on average (that is, random perturbations lead to instability 92.3% of the time). As evolution progresses and modularity increases, this percentage rises sharply until, by generation 2000, a typical network is 97.2% ($\pm$ 3.6%) resistant to mutation. We test the generality of this apparent connection between robustness and modularity by repeating the above experiment on stable hierarchies [11] and stable indecomposable networks, both created using an alternative (nonevolutionary) algorithm. Preliminary results confirm that hierarchical modularity improves the robustness of network stability to random structural perturbations.

By virtue of being hierarchical, our networks suffer a natural vulnerability to targeted attack which complements their robustness to random attack. This is reminiscent of a property of scale-free networks, with the important difference that our results are concerned with whether random perturbations destabilize network dynamics, rather than the purely structural question of whether random node deletion severs the giant component. Our framework also bears a qualitative resemblance to that of highly optimized tolerance [12], in that an evolutionary algorithm leads to highly structured systems characterized by robust performance and whose dynamical and structural properties are shown to be closely interrelated.

In conclusion, our evolved networks achieve stability and robustness by means of a hierarchical organization of stable modules, rather than a subdivision into a number of disconnected networks which are trivially easier to stabilize. It is interesting to note that the emergent distribution of SCC sizes (Fig. 3) has a heavy tail—smaller modules predominate, but larger ones occur more frequently than if the distribution were exponential. Furthermore, since their eigenvalues determine the dynamics of the entire network, the modules themselves are *functional* units and our results demonstrate how functional modularity can be closely linked to structural modularity. The dynamical function of a real network may require that some of its modules be large and complex, even if its need to resist mutation calls for many small modules. A heavy-tail distribution of module sizes therefore can provide a compromise between the conflicting demands of function, flexibility, and robustness.

Although we are not implying that a need for linear stability has governed the structure of real systems, it is highly suggestive nevertheless that so many significant properties emerge from a simple GA which selects only for stability. Given the directed structure of many real networks, e.g., genetic, metabolic, social, and business networks, our results may be relevant in a number of applications. Real networks are assembled according to system-specific rules, which may act at the level of individual nodes or on the network as a whole. Our work demonstrates that whenever these rules guide the system towards a modular architecture, the resulting networks are more likely to be linearly stable and robust. Many exciting questions remain: What are other features of large linearly stable networks, e.g., degree correlations [1], diagonal values, and scaling properties? How does the relationship between structure and dynamics change in discrete-time systems, where $\mathbf{x}_{n+1} = M\mathbf{x}_n$ replaces $\dot{\mathbf{x}} = M\mathbf{x}$? How might evolutionary rules acting at the level of individual nodes lead to the self-organization of stable networks [10]? We hope that our work will stimulate further investigation of general principles connecting structure and dynamics and that researchers can use the linear special case as a tool for exploring the richer behavior of nonlinear dynamics on networks.

This research is supported by the National Science Foundation, IGERT Grant No. DGE-9870631. We wish to thank Steve Strogatz, Mary Lou Zeeman, and Richard Wiener for their support and interest in this work.

*Email address: jhm28@cornell.edu

[1] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002); M. E. J. Newman, SIAM Rev. **45**, 167 (2003); S. H. Strogatz, Nature (London) **410**, 268 (2001).
[2] V. M. Eguíluz and K. Klemm, Phys. Rev. Lett. **89**, 108701 (2002); M. Timme, F. Wolf, and T. Geisel, Phys. Rev. Lett. **89**, 258701 (2002); C. Furusawa and K. Kaneko, Phys. Rev. Lett. **90**, 088102 (2003).
[3] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[4] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **401**, 130 (1999); A.-L. Barabási and R. Albert, Science **286**, 509 (1999); H. Jeong *et al.*, Nature (London) **407**, 651 (2000); S. Redner, Eur. Phys. J. B **4**, 131 (1998).
[5] M. Girvan and M. E. J. Newman, Proc. Natl. Acad. Sci. U.S.A. **99**, 7821 (2002); W. Rives and T. Galitsky, Proc. Natl. Acad. Sci. U.S.A. **100**, 1128 (2003); E. Ravasz *et al.*, Science **297**, 1551 (2002).
[6] C. Moore and M. E. J. Newman, Phys. Rev. E **61**, 5678 (2000); M. Kuperman and G. Abramson, Phys. Rev. Lett. **86**, 2909 (2001); D. H. Zanette, Phys. Rev. E **64**, 050901(R) (2001).
[7] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **406**, 378 (2000).
[8] R. M. May, *Stability and Complexity in Model Ecosystems* (Princeton University Press, Princeton, 1974).
[9] A. D. Sánchez, J. M. López, and M. A. Rodríguez, Phys. Rev. Lett. **88**, 048701 (2002).
[10] S. Levin, *Fragile Dominion* (Perseus, Reading, 1999).
[11] We select random modules, with sizes chosen from a $1/x$ distribution, and connect these hierarchically into a single $N = 25$ component.
[12] J. M. Carlson and J. Doyle, Phys. Rev. E **60**, 1412 (1999).