Ulanowicz, R.E. 2013. Circumscribed complexity in ecological networks. Pp. 249-258 In: M. Dehmer, A. Mowshowitz and F. Emmert-Streib (Eds.) *Advances in Network Complexity*, Volume 4. Weinheim : Wiley-Blackwell.

11
Circumscribed Complexity in Ecological Networks

Robert E. Ulanowicz

11.1
A New Metaphor

As the world reveals itself to be ever more complex, trust in science as a reductionistic exercise has begun to wane. The Enlightenment assumption of universal atomism appears to be contradicted in all too many instances [1], and it is becoming increasingly evident that relationships can track better the behavior of complex systems than can the component elements by themselves. Nowhere has this been more obvious than in the study of ecosystems, where networks of relationships have been employed to describe ecosystems for more than seven decades [2,3]. It is hardly surprising, then, that scientists in other fields have turned toward networks as relational metaphors to represent the broader dynamics of nature and society [4]. The overwhelming influence that computational and communications networks exert on contemporary life only underscores the importance of describing and understanding networks.

It should be remarked, however, that the recent surge in network research has not paralleled the historical treatment of ecosystem networks in one important respect. Within physics and sociology, work has focused on common graphs, or at most, digraphs (ibid.). Ecologists, by contrast, elected from the very beginning to tackle ecosystems in terms of *weighted* digraphs [2]. The opinion prevalent among physicists has been that ecologists are concerned with only a special subclass of networks that can be conveniently ignored in order to concentrate on "more general" topological forms. It is worth noting that this popular attitude could be mistaken and ultimately counter-productive. For example, the strategy known as engineering science [5] is one that begins with the most general available descriptions of dynamics (such as the full equations of motion) and simplifies as necessary to describe the immediate problem (e.g., the Navier–Stokes equations to describe flow in a pipe).

That a description of the simple and specific can be abstracted from that of the most complex and general is the stratagem that will be pursued here. The focus of what follows will be the quantification of weighted digraphs. Any results from this

Advances in Network Complexity, First Edition. Edited by M. Dehmer, A. Mowshowitz, and F. Emmert-Streib.

more detailed representation will apply *a fortiori* to common graphs and digraphs as degenerate corollaries (i.e., by equating all weights and/or removing arrows).

11.2
Entropy as a Descriptor of Structure

Various forms of the statistical entropy have proven to be useful quantitative descriptors of the complexity of networks [6]. Such metrics are especially relevant to the description of ecological networks, because the statistical entropy is a form of the logarithmic average (which in turn reveals the geometric mean). Measures in ecosystems commonly range over 9 to 12 orders of magnitude, making the logarithmic mean far more meaningful than a simple average.

A second (and generally unappreciated) advantage of statistical entropies is that they allow one to extrapolate discrete topological measures into the continuous realm of weighted structures [7]. For example, the node in Figure 11.1a has three topological edges emanating from it. If these edges are equiponderant, then each flow accounts for one-third of the total weight. Therefore, the logarithmic mean becomes $H = -\sum_i p_i \ln(p_i)$, where $p_1 = p_2 = p_3 = 300/900 = 1/3$ so that $H = \ln(3)$. One may define an "effective" number of flows as $F \equiv e^H = e^{\ln 3} = 3$, which in this particular instance corresponds exactly to the discrete topological count.

In general, however, the weights of the edges can be far from equal, as in Figure 11.1b. Here the logarithmic mean fraction, H, is calculated as $-(.001)\ln(.001) - (.005)\ln(.005) - (.995)\ln(.995) = .0384$, and the effective number of flows thereby becomes $F = e^H = 1.039$. Clearly, simple topological counts can misrepresent what is going on in a system. In particular, a discrete count seriously overestimates complexity whenever certain edges become dominant, which is most of the time in natural systems.

A third advantage of using entropy-like measures is perhaps the least appreciated, but turns out to be far more important than the first two. This is because networks are rarely recognized for what they are – namely, a metaphor for the entanglement of constraint and freedom [8]. Rather, networks are commonly seen through the

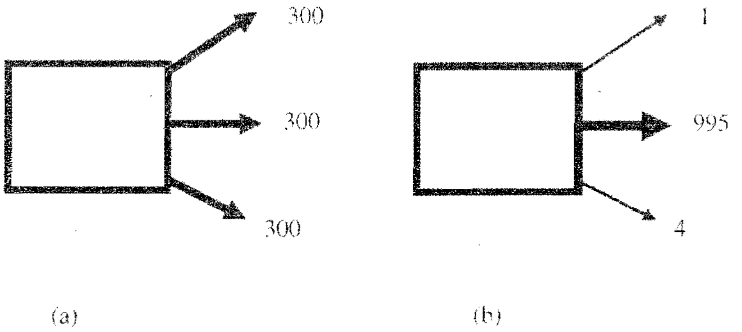

Figure 11.1 Two nodes, each with three efferent edges (a) of equal magnitudes and (b) of highly disparate magnitudes.

prevailing mechanical lens as determinate structures. Perhaps this is a historical consequence of the fact that the earliest networks consisted of electrical circuits, which could be analyzed in mechanistic fashion. Whence, the aim in so many papers on networks is to uncover a "mechanism" to explain a particular structure, so imbued is science in its mechanistic origins.

While electrical circuits are often complicated, they rarely are complex. Such is not the case with ecological, economic, or social networks, all of which embody demonstrably aleatoric behaviors. The normal way of dealing with such structures is to identify a set of rules that operate in abstraction from random events, that is, the conventional dichotomy between chance and necessity. But behaviors in networks do not readily lend themselves to such dichotomy. In most networks each node is connected to only a small subset of other nodes. That is, behaviors are *constrained* to affect only particular other nodes. Furthermore, contingent behaviors in networks cannot easily be interpreted as "blind chance," given the constrained, anisotropic directions in which they may act. Nor can the efferent effects from any node be described as strictly mechanical because of the ambiguity in the destinations they might affect at any one time. Networks represent a *complex* amalgam of contingency and rules.

11.3
Addressing Both Topology and Magnitude

Serendipitously, the statistical entropy is well-suited to quantify such complexity [9]. To see this, one notes that the magnitude of any edge is *jointly* associated with two nodes – its origin and its terminus. Accordingly, one may define a joint frequency, $p_{ij} \equiv (T_{ij}/T_{..})$, where T_{ij} is the magnitude assigned to the edge connecting i with j, and $T_{..}$ is the sum of all the edge magnitudes. The complexity of the network can then be represented by

$$H = -\sum_{i,j} p_{ij} \ln\left(p_{ij}\right) = -\sum_{i,j} \frac{T_{ij}}{T_{..}} \ln\left(\frac{T_{ij}}{T_{..}}\right) \tag{11.1}$$

As has just been argued, this complexity consists of amalgamated constraint and freedom. Can the degrees to which each attribute is manifest in any network be separately quantified? Toward this end, it is necessary only to calculate the marginal sums of the magnitudes as they apply to afferent and efferent edges, respectively [10]. That is, let $T_{.j} = \sum_i T_{ij}$ and $T_{i.} = \sum_j T_{ij}$. H can then be decomposed into two non-negative terms,

$$H = \sum_{i,j} \frac{T_{ij}}{T_{..}} \ln\left(\frac{T_{ij} T_{..}}{T_{i.} T_{.j}}\right) - \sum_{i,j} \frac{T_{ij}}{T_{..}} \ln\left(\frac{T_{ij}^2}{T_{i.} T_{.j}}\right)$$

where the first term

$$A = \sum_{i,j} \frac{T_{ij}}{T_{..}} \ln\left(\frac{T_{ij} T_{..}}{T_{i.} T_{.j}}\right) \tag{11.2}$$

quantifies the degree of constraint that structures the system, and the second,

$$\Phi = -\sum_{i,j} \frac{T_{ij}}{T_{..}} \ln\left(\frac{T_{ij}^2}{T_{i.}T_{.j}}\right) \tag{11.3}$$

captures the extent of freedom manifest in the system [3].

To recapitulate, various forms of the statistical entropy can be invoked to characterize the complexity of a network [6]. Usually, these measures have been applied to static structures, such as molecules [11]. Weighted digraphs, however, may also pertain to dynamical situations where, for example, the edges might represent physical flows of various magnitudes. In such dynamical situations it is now possible to parse the complexity to assess how much of it pertains to internal constraints extant within the system and how much can be attributed to residual incoherencies or external factors.

11.4
Amalgamating Topology with Magnitudes

There is a topological aspect to the decomposition just described. In particular, it is rather easy to demonstrate that the term Φ can be used to quantify the effective number of edges that impinge upon a typical node of a weighted digraph. In the introduction, it was shown how the effective number of edges in any collection can be calculated as e^H. This result generalizes to a network of interacting nodes as the measure $C \equiv e^{\Phi/2}$, where C is the effective number of edges *either* efferent *or* afferent to a typical node [12].[1] The measure C corresponds to the link density [13], defined in conventional foodweb analysis as the number of edges divided by the number of nodes (F/N). As with the examples in Figure 11.1, this continuous measure yields the appropriate result when applied to discrete, degenerate digraphs ([12], Appendix A).

What then of A, the measure of constraint? Zorach and Ulanowicz [14] demonstrated how this term describes the number of distinct "roles," R, evident in the network structure. They defined $R \equiv e^A$. Just as A and Φ are complementary in the algebraic sense that they sum to yield H, one may consider R and C to be "orthogonal." For example, if C is considered to measure the "breadth" of freedom at each node, then one may conceive of R as the "depth" (or "length") of the network. In ecosystem trophic networks, R can be shown to equal one greater than the number of trophic levels that are functioning in the ecosystem.

In summary, any weighted digraph can be characterized by two numbers, a network "breadth," C, and a corresponding "depth," R. The conventional graph theoretic property "network diameter" corresponds to the sum $C + R$ in this scheme of reckoning effective attributes.

---

[1] The factor $^1/_2$ in the exponent assures that each edge is not counted twice.

11.5
Effective Network Attributes

As shown in the introductory example, the effective number of overall flows is gauged by $F = e^H$ [14]. Because one now has measures both of effective link density, C, and of the effective number of flows, F, it now becomes possible to calculate a homologous "effective" number of nodes as the quotient, $N = F/C$. Here N represents not the actual nodes as defined by the observer but rather the effective number of nodes as functionally grouped by that particular flow structure. N is usually less than the actual number of nodes, because constraints tend to bind nodes that might otherwise act independently of each other.

As Bersier [7] indicated, conventional web measures, such as the number of nodes, number of edges, link density, and the number of roles, as counted in conventional fashion for digraphs, all have their weighted homologs in N, F, C and R, respectively. Bersier argued further that logarithmic counterparts converge more rapidly to the properties they are intended to represent than do conventional indexes. It can even be argued that the effective measures indicate limits that otherwise might remain obscure under conventional definitions – limits such as how complex networks of natural system can become.

11.6
Limits to Complexity

The common experience is that natural systems tend to increase in complexity up to a point, after which they either fall apart due to lack of coherence or simplify at a larger scale under the aegis of some synchronous dynamic. That is, the complexity of natural systems appears to be bounded, but the question remains, "To what degree?" As regards network breadth, anecdotal evidence points toward a limit on link density in the neighborhood of three edges per node: Pimm [15], for example, noted that his collection of food webs averaged 3.1 links per node, while Kauffman [16] reported that networks of genetic mechanisms tended to become unstable above about three links per node. Wagensberg *et al.* [17] wrote about the "magic number 3" as the watershed beyond which networks of ecosystem dynamics do not persist.

In an effort to quantify the upper bound on link density, May [18] pointed to the Wigner [19] semicircle law in reference to network properties. May suggested that whenever $\alpha \leq (n\Gamma)^{-\frac{1}{2}}$, the system would likely be stable. (α is the effective strength of interaction, n is the number of nodes in the system, and Γ is the fraction of possible connections that are realized in the given system.) Ulanowicz [20] reinterpreted May's criterion in logarithmic terms as

$$\ln(C) \leq \frac{3\ln(C^*)}{2\sqrt{C^*}}$$

where C^* is the value of C after all edge magnitudes have been set equal. One may demonstrate that C is maximal when $C^* = e^2$ so that the greatest C possible, C_ω, becomes

$$C_\omega = e^{\left(\frac{3}{e}\right)} \cong 3.01$$

which agrees well with the anecdotal consensus.

As for food chain length, there is less agreement as to where the limit might lie. Pimm and Lawton [21] suggested an upper bound on chain length of five transfers ($R = 6$). Although individual feeding pathways of length greater than five can be identified within feeding networks, the average trophic level at which any compartment feeds is generally four or fewer (e.g., [22]), indicating that the maximum effective trophic level for the entire system should be even lower.

Because the metaphor of the network implies that contingency and necessity may be deeply entwined in a natural system, R and C are thus likely to have some close relationship with each other. A clue to such coupling can be found in the distribution of the quotient $a = A/H$, which is purported to measure the relative degree of organization [23]. Whereas a was expected to be liberally distributed over the interval $1 \geq a \geq 0$, ecosystem networks with greater than 13 nodes as estimated from a wide distribution of habitats happened to cluster closely around the value $a \approx 0.40$ [24]. If a is nearly constant (calling said constant K), then R and C are close to being functionally related. In particular, it is straightforward to show that for constant K (≈ 0.40),

$$\ln(C) = \frac{1-K}{2K}\ln(R)$$

Whence, the value of R corresponding to C_ω turns out to be $R_\omega\,(3.613 \exp[6/\{e^2 - e\}] \approx 4.383$. The bottomline is that the large majority of ecological systems have an effective trophic length less than 3.4, which may account for why so many ecosystems can be reasonably modeled as a three-step trophic cascade [25].

Knowing C_ω and R_ω makes it possible to calculate N_ω and F_ω, the upper limits on the effective numbers of nodes and flows, respectively. It works out that $N_\omega = R_\omega \times C_\omega = 4.383 \times 3.01 \approx 13.2$ nodes and $F_\omega = R_\omega \times {C_\omega}^2 \approx 39.9$ flows. These are remarkable results. Combined they imply that a given ecological network can possess hundreds of actual nodes and thousands of real flows, but from a dynamical viewpoint the system should be adequately represented by a virtual network having at most 13 nodes and 40 flows. As surprising as this conclusion might seem, it makes intuitive sense. The world economy, for example, consists of some 170 national sectors and well over 10,000 bilateral international trading partnerships. When it comes to managing the global economy, however, most decisions are made by the international G8 group. Presumably, the remaining 160 or so economies can all be folded into the five remaining virtual nodes in power-law fashion.[2]

---

[2] This assumes that the value $K = 0.40$ determined from ecological data applies as well to economic systems, which may not be the case. It is unlikely, however, that the limit N_ω for economic systems will differ radically from the 13 estimated for ecosystems.

A corollary result is that the effective overall complexity of natural ecosystem networks remains circumscribed. That is, one does not expect to encounter ecosystem networks having an overall complexity much greater than $H_\omega \approx 3.67$ nats.

The value of $C_\omega \approx 3.01$ is basically a theoretical result [26], whereas the value $K = 0.4$ remains a phenomenological observation. For now, one can only conjecture as to what sets the balance between A and Φ that causes systems to cluster around $a \approx 0.4$. It would appear that when a is lower than 0.4, manifold opportunities exist for autocatalytic cycles to arise spontaneously, and the selection generated by such autocatalysis increases A, because participating autocatalytic elements tend to grow at the expense of nonparticipating nodes [27]. Conversely, when a is significantly greater than 0.4, some autocatalytic cycles will have become too specialized and fall vulnerable to being short-circuited by smaller, shorter cycles as the latter continue to appear [23].

11.7
An Example Ecosystem Network

Example 11.1

The reader is referred to the trophic network depicted in Figure 11.2, in which carbon flows (in mg m^{-2} d^{-1}) among the 17 components of a tidal marsh gut adjoining the Crystal River in Florida [3,28]. Substituting the

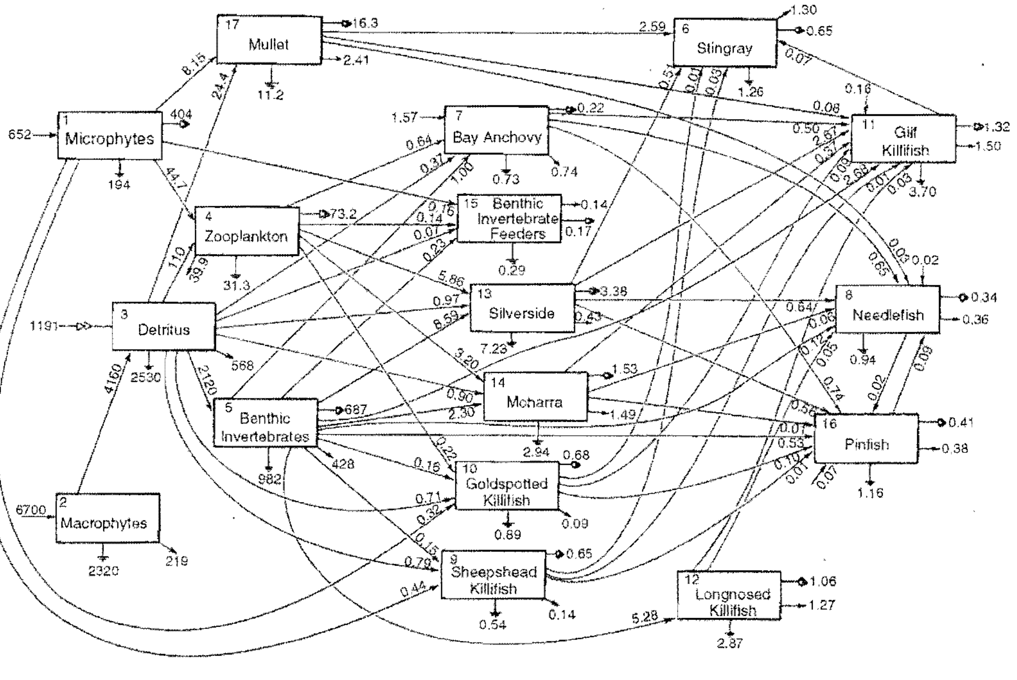

Figure 11.2 A weighted digraph of carbon flows (mg m^{-2} d^{-1}) among the principal taxa of a tidal marsh gut ecosystem, Crystal River, FL. The linked arrows indicate returns to the detritus (compartment #3). After Homer *et al.* [28], with kind permission from Elsevier.

magnitudes of the exchanges from the figure as the T_{ij}'s in Eqs. (Eqs. (11.1)–(11.3) yields,

$$H = 2.160 \text{ nats}$$

$$A = 0.8761 \text{ nats}$$

and $\Phi = 1.3099$ nats, respectively.

. The degree of order ($a = A/H$) works out to 0.401, which is very close to the mean for the entire collection of networks. The remaining parameters of the effective virtual network thereby become

$$R = e^{A} = 2.402 \text{ roles}$$

$$C = e^{\Phi/2} = 1.925 \text{ flows/node}$$

$$F = e^{H} = 8.671 \text{ effective flows}$$

and $N = R \times C = 4.624$ effective nodes.

The 2.402 roles translate into an effective trophic length for the network of 1.402.

The low values for F and N indicate that the Crystal River marsh ecosystem network is relatively simple, as is readily apparent from the many parallel functional pathways evident in Figure 11.1.

Example 11.2

The second example is one of the most highly articulated ecological networks estimated to date. The network is of the shallow water marine ecosystem of Florida Bay during the dry season. It consists of 125 compartments with 2135 exchanges of carbon (gC m^{-2} y^{-1}) among them. The community is too complicated to depict as a flow diagram, but a system description and the raw data used here can be found in ATLSS [29].

Evaluation of H, A, and Φ according to Eqs. (11.1)–(11.3), respectively, yield the following magnitudes:

$$H = 3.651 \text{ nats}$$

$$A = 1.390 \text{ nats}$$

$$\Phi = 2.261 \text{ nats}$$

$$R = e^{A} = 4.012 \text{ roles}$$

$$C = e^{\Phi/2} = 3.098 \text{ flows/node}$$

$$F = e^{H} = 41.05 \text{ effective flows}$$

and $N = R \times C = 12.43$ effective nodes.

Unlike the network in Example 11.1, the Florida Bay weighted digraph pushes the envelope on complexity. In fact, two parameters slightly exceed the upper bounds indicated above. ($C = 3.098 > C_{\omega} = 3.01$, and $F = 41.05 >$

F_o, (39.9.) One needs bear in mind, however, that the values used for the T_{ij}'s are estimates that can easily differ from actual values by factors of 2 or 3. Furthermore, similar schemes were used to estimate many of the flows. Such uniformity tends to even out the distribution of flow magnitudes, resulting in slight inflation of H and Φ (and consequently, C). Given the looseness in the flow estimates, the approximation of the calculated parameters to their theoretical limits is quite acceptable.

11.8
A New Window on Complex Dynamics

To summarize, the statistical entropy has been invoked heretofore in a myriad of ways, usually to characterize static structures, such as molecules, which can be aptly characterized in terms of common graphs or digraphs. In contrast, ecosystems are dynamical entities and the connections therein span a great range of magnitudes. Anything less than a description of these systems as weighted digraphs will ignore important relative features of their behaviors. Fortunately, the statistical entropy can be adapted to address the dynamical features of these more complex systems. The conditional nature of dynamics in ecosystems prompts the introduction of joint and conditional probabilities into the calculus, which allows the decomposition of the network entropy into separate terms that reflect the complementary attributes of constraint and flexibility.

The importance of this separation cannot be overemphasized. For too long now, science has focused solely on the laws and constraints that guide (but not determine) how phenomena transpire to the exclusion of the manifold noise, inefficiencies, and opportunities that actually determine true change in evolutionary systems. By separating H into A and Φ, one is able to follow the progress of these antagonistic tendencies within the framework of a unified calculus [20,24].

In addition, the introduction of a logarithmic calculus has facilitated the identification of the limits to complexity in natural systems. Arbitrarily complex systems fall apart spontaneously, whereas excessively efficient and streamlined systems perish, because they cannot adequately adapt to novel conditions. The decomposition of statistical entropy into its Bayesian components allows one to pinpoint the propitious balance between constraint and contingency that enables systems to persist [20].

Finally, adapting entropy measures to fully detailed, weighted digraphs empowers the investigator to apply those same indexes in corollary fashion to more degenerate networks, that is, digraphs and common graphs. For example, it remains to be seen what, if anything, the separate indexes for constraint and redundancy will reveal about the structures of large, complicated organic molecules.

Treating the statistical entropy in Bayesian fashion leads to a totally different description of living reality – one that resembles less the mechanistic clockworks of decades past and more a transactional image that provides a more complete picture of the complex living world [1,30].

References

1 Ulanowicz, R.E. (2009a) *A Third Window: Natural Life Beyond Newton*, Templeton Foundation Press, West Conshohocken, Pennsylvania, PA, USA, 196 p.

2 Lindeman, R.L. (1942) The trophic-dynamic aspect of ecology. *Ecology*, **23**, 399–418.

3 Ulanowicz, R.E. (1986) *Growth and Development: Ecosystems Phenomenology*, Springer, NY, USA, 203 p.

4 Barabási, A.-L. (2002) *Linked: The New Science of Networks*, Perseus, Cambridge, MA, USA, 280 p.

5 UToronto (2012) What is Engineering Science? http://engsci.utoronto.ca/explore_our_program/about_engsci.htm (last accessed April 8, 2013).

6 Dehmer, M. and Mowshowitz, A. (2011) A history of graph entropy measures. *Inform. Sciences*, **220** (1), 57–78.

7 Bersier, L. (2002) Quantitative descriptors of food web matrices. *Ecology*, **83** (9), 2394–2407.

8 Pahl-Wostl, C. (1995) *The Dynamic Nature of Ecosystems: Chaos and Order Entwined*, John Wiley & Sons, Inc., New York, NY, USA, 267 p.

9 MacArthur, R.H. (1955) Fluctuations of animal populations and a measure of community stability. *Ecology*, **36**, 533–536.

10 Rutledge, R.W., Basorre, B.L., and Mulholland, R.J. (1976) Ecological stability: an information theory viewpoint. *J. Theor. Biol.*, **57**, 355–371.

11 Bonchev, D. (1983) *Information Theoretic Indices for Characterization of Chemical Structures*, Research Studies Press, Chichester, UK.

12 Ulanowicz, R.E. and (1991) W.F., Wolff, Ecosystem flow networks: loaded dice? *Math. Biosci.*, **103**, 45–68.

13 Cohen, J.E., Briand, F., Newman, C.M., and Palka, Z.J. (1990) *Community Food Webs: Data and Theory*, Springer, Berlin, 308 p.

14 Zorach, A.C. and Ulanowicz, R.E. (2003) Quantifying the complexity of flow networks: how many roles are there? *Complexity*, **8** (3), 68–76.

15 Pimm, S.L. (1982) *Food Webs*, Chapman and Hall, London, 219 p.

16 Kauffman, S.A. (1991) Antichaos and adaptation. *Sci. Am.*, **265**, 78–84.

17 Wagensberg, J., Garcia, A., and Solé, R.V. (1990) Connectivity and information transfer in flow networks: two magic numbers in ecology. *Bull. Math. Biol.*, **52**, 733–740.

18 May, R.M. (1972) Will a large complex system be stable. *Nature*, **238**, 413–414.

19 Wigner, E. (1958) On the distribution of the roots of certain symmetric matrices. *Ann. Math.*, **67**, 325–328.

20 Ulanowicz, R.E. (2002) The balance between adaptability and adaptation. *BioSystems*, **64**, 13–22.

21 Pimm, S.L. and Lawton, J.H. (1977) Number of trophic levels in ecological communities. *Nature*, **268**, 329–331.

22 Baird, D. and Ulanowicz, R.E. (1989) The seasonal dynamics of the Chesapeake Bay ecosystem. *Ecol. Monogr.*, **59**, 329–364.

23 Ulanowicz, R.E., Holt, R.D., and Barfield, M. Limits on ecosystem trophic complexity. *Ecol. Lett.*, in review.

24 Ulanowicz, R.E. (2009b) The dual nature of ecosystem dynamics. *Ecol. Model.*, **220**, 1886–1892.

25 Holt, R.D., Holdo, R.M., and van Veen, F.J. (2010) Theoretical perspectives on trophic cascades: current trends and future directions, in *Trophic Cascades: Predators, Prey and the Changing Dynamics of Nature* (eds J. Terborgh and J.A. Estes), Island Press, Washington, DC, USA, pp. 301–318.

26 Tao, T., Vu, V., and Krishnapur, M. (2010) Random matrices: Universality of ESDs and the circular law. *Ann. Probab.*, 38 (5), 2023–2065.

27 Ulanowicz, R.E. (1997) Ecology, the Ascendent Perspective, Columbia University Press, New York, NY, USA, 201p.

28 Homer, M., Kemp, W.M., and McKellar, H. (1976) Trophic analysis of an estuarine ecosystem: salt marsh-tidal creek system near Crystal River, Florida. Unpublished manuscript, Department of Environmental Engineering, University of Florida, Gainesville.

29 ATLSS (1998) Network Analysisof the trophic dynamics of South Florida Ecosystems: Florida Bay. http://www.cbl.umces.edu/~atlss/FBay001.html (last accessed April 8, 2013).

30 Ulanowicz, R.E. (2011) Towards quantifying a wider reality: Shannon exonerata. *Information*, 2, 624–634.