Extracting the multiscale backbone of complex weighted networks

M. Ángeles Serrano[a,1], Marián Boguñá[b], and Alessandro Vespignani[c,d]

[a]Instituto de Física Interdisciplinar y Sistemas Complejos, Consejo Superior de Investigaciones Científicas-Universitat Illes Balears, E-07122 Palma de Mallorca, Spain; [b]Departament de Física Fonamental, Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain; [c]Center for Complex Networks and Systems Research, School of Informatics, Indiana University, 919 East 10th Street, Bloomington, IN 47406; and [d]Complex Networks Lagrange Laboratory, Institute for Scientific Interchange, 10133 Torino, Italy

A large number of complex systems find a natural abstraction in the form of weighted networks whose nodes represent the elements of the system and the weighted edges identify the presence of an interaction and its relative strength. In recent years, the study of an increasing number of large-scale networks has highlighted the statistical heterogeneity of their interaction pattern, with degree and weight distributions that vary over many orders of magnitude. These features, along with the large number of elements and links, make the extraction of the truly relevant connections forming the network's backbone a very challenging problem. More specifically, coarse-graining approaches and filtering techniques come into conflict with the multiscale nature of large-scale systems. Here, we define a filtering method that offers a practical procedure to extract the relevant connection backbone in complex multiscale networks, preserving the edges that represent statistically significant deviations with respect to a null model for the local assignment of weights to edges. An important aspect of the method is that it does not belittle small-scale interactions and operates at all scales defined by the weight distribution. We apply our method to real-world network instances and compare the obtained results with alternative backbone extraction techniques.

disordered systems | multiscale phenomena | filtering | visualization

In recent years, a huge amount of data on large-scale social, biological, and communication networks, meticulously collected and catalogued, has become available for scientific analysis and study. Examples can be found in all domains; from technological to social systems and transportation networks on a local and global scale, and down to the microscopic scale of biochemical networks (1–3). Common traits of these networks can be found in the statistical properties characterized by large-scale heterogeneity with statistical observables such as nodes' degree and traffic varying over a wide range of scales (4). The sheer size and multiscale nature of these networks make very difficult the extraction of the relevant information that would allow a reduced representation while preserving the key features we want to highlight. A typical example is seen in the visualization of networks. Although, in general, it is possible to create wonderful images of large-scale heterogeneous networks, the amount of valuable information gathered is in most cases very little because of the redundant intricacy generated by the overwhelming number of connections. Problems such as the extraction of the relevant backbone or the isolation of the statistically relevant structures/signal that would allow reduced but meaningful representations of the system are indeed major challenges in the analysis of large-scale networks.

In complex weighted networks, the discrimination of the right trade-off between the level of network reduction and the amount of relevant information preserved in the new representation faces us with additional problems. In many cases, the probability distribution $P(\omega)$ that any given link is carrying a weight ω is broadly distributed, spanning several orders of magnitude. This feature implies the lack of a characteristic scale and any method based

on thresholding would simply overlook the information present above or below the arbitrary cutoff scale. Although this issue would not be a major drawback in networks where the intensities of all the edges are independently and identically distributed, the cutoff of the $P(\omega)$ tail would destroy the multiscale nature of more realistic networks where weights are locally correlated on edges incident to the same node and nontrivially coupled to topology (5). Thus, the presence of multiscale fluctuations calls for reduction techniques that consistently highlight the relevant structures and hierarchies without favoring any particular resolution scale. Furthermore, it also demands a change in the focus toward a local perspective rather than a global one, where the relevance of the connections could be decided at the level of nodes in relative terms.

In this work, we concentrate on a particular technique that operates at all the scales defined by the weighted network structure. This method, based on the local identification of the statistically relevant weight heterogeneities, is able to filter out the backbone of dominant connections in weighted networks with strong disorder, preserving structural properties and hierarchies at all scales. We discuss our multiscale filter in relation to the appropriate null model that provides the basis for the statistical significance of the heterogeneity measurements. We apply the technique to two real-world networks, the U.S. airport network and the Florida Bay food web, and compare the results with those obtained by the application of thresholding methods.

Results and Discussion

In statistical mathematics, as in other areas, filtering techniques aimed at uncovering the relevant information in datasets are popular and successful. One could cite, for instance, the Principal Components Analysis to identify hidden patterns by reducing the effective dimension of multivariate data (6). In the following, we will refer to the *network reduction* as the construction of a network that contains far fewer data (in our case, links) and allows the discrimination and computational tractability of the relevant features of the original networks; for instance, the traffic backbone of a large-scale transportation infrastructure. Reduction schemes can be divided into two main categories: coarse-graining and filtering/pruning. In the first case, nodes sharing a common attribute could be gathered together in the same class—group, community, etc.—and then substituted by a single new unit that represents the whole class in a new network representation of the system (7–10). This coarse-graining is indeed zooming out the system so that it can be observed at different scales. Something completely different is done when a filter is applied. In this case, the observation scale is fixed and the representation that the network symbolizes is

APPLIED PHYSICAL SCIENCES

not changed. Instead, those elements, nodes and edges, that carry relevant information about the network structure are kept while the rest are discarded. An example of a well-known hierarchical topological filter, although usually not referred to as such, is the *k*-core decomposition of a network (11), with a filtering rule that acts on the connectivity of the nodes.

In the case of weighted networks (5), two basic reduction techniques refer to the extraction of the minimum spanning tree and the application of a global threshold on the weights of the links so that just those that beat the threshold are preserved. The minimum spanning tree of a graph $\mathcal{G}$, a classical concept of graph theory (12), is the shortest-length tree subgraph that contains all the nodes of $\mathcal{G}$. These definitions can be generalized for weighted graphs (13). A minimum spanning tree of a weighted graph $\mathcal{G}$ is the spanning tree of $\mathcal{G}$ whose edges sum to minimum weight. This idea has been exploited along with percolation criticality to define superhighways in weighted networks (14). By using opportune transformation rules for the weights, it is also possible to define maximum weighted spanning trees and other analogous definitions. One of the big limitations of this method is that spanning trees are by construction acyclic. This means that reduced networks obtained by this algorithm are overly structural simplifications that destroy local cycles, clustering coefficient, and the clustering hierarchies often present in real world networks.

These previous drawbacks are not present in the application of a threshold to the global weight distribution that removes all connections with a weight below a given value ω_c. This filter has been used, for instance, in the study of functional networks connecting correlated human brain sites (15) and food web resistance as a function of link magnitude (16). This approach, however, belittles nodes with a small strength s (defined as the sum of weights incident to the node $s_i = \sum_j w_{ij}$), since the introduction of ω_c induces a characteristic scale from the outset. As a consequence, strongly disordered networks with heavy-tailed statistical distributions $P(s)$ and $P(\omega)$ cause this simple thresholding algorithm to be very poorly performing since nodes with small s are systematically overlooked. This is an even more serious drawback when weights are correlated at the local level. In this type of network, interesting features and structures are present at all scales and the introduction of such an artificial cutoff drastically removes all information below the cutoff scale.

Local Fluctuations. To develop a multiscale reduction algorithm, we take advantage of the local fluctuations of weights on the links emanated by single nodes. In heterogeneous weighted networks with strong disorder, i.e., heavy-tailed $P(\omega)$ and $P(s)$ distributions, a few links carry the largest proportion of the node's total strength. Furthermore, most real networks have nodes surrounded by incident edges with associated weights that are heterogeneously distributed and correlated between them. The fingerprint of these correlations is observed in the nontrivial dependence between weights and topology (5). The better a node is connected to the rest of the network, the higher the weight of its edges so that the strength tends to grow superlinearly with the degree. However, the strength alone is not enough to capture the weighted structure of nodes even at the local level. We need to introduce some measure of the fluctuations of the weights attached to a given node, and we want to do it at the local level in relative terms so that each node could independently assess the importance of its connections. To this end, we first normalize the weights of edges linking node *i* with its neighbors as $p_{ij} = \omega_{ij}/s_i$, being s_i the strength of node *i* and w_{ij} the weight of its connections to its neighbor *j*. Then, by using the disparity function defined in *Materials and Methods*, it is possible to see that, even at the local level defined by the edges adjacent to a single node, a few of those edges carry a disproportionate fraction p_{ij} of the node's strength, with the remaining edges carrying just a small fraction of the node's strength (17, 18).

Being more specific, we are interested in all edges with weights representing a significant fraction of the local strength and weight magnitude of each given node. However, local heterogeneities could simply be produced by random fluctuations. It is then fundamental to introduce a null model that informs us about the random expectation for the distribution of weights associated to the connections of a particular node. Empirical values not statistically compatible with the null model define, on a node-by-node basis, whether the observed weight heterogeneity and intensity are statistically significant and define the relevant part of the signal due to specific and relevant organizing principles of the network structure. This procedure would determine without arbitrariness how many connections for every node belong to the backbone of connections that carry a statistically disproportionate weight—be they one, zero, or many—providing sparse subnetworks of connected links selected according to the total amount of weight we intend to characterize. This reduction scheme necessarily encodes a wealth of information because the reduced network not only contains the links carrying the largest weight in the network, but also all links that can be considered, according to a predefined statistical significance level, to define the relevant structure (signal) generated by the weight and strength assignment with respect to the simple randomness of the null hypothesis. An important aspect of this construction is that the ensuing reduction algorithm does not belittle small nodes in terms of strength and then offers a practical procedure to reduce the number of connections taking into account all of the scales present in the system.

The Disparity Filter. In the following, we discuss the disparity filter for undirected weighted networks, although it is also applicable to directed ones as reported in the supporting information (SI) *Appendix*. The null model that we use to define anomalous fluctuations provides the expectation for the disparity measure of a given node in a pure random case. It is based on the following null hypothesis: the normalized weights that correspond to the connections of a certain node of degree *k* are produced by a random assignment from a uniform distribution. To visualize this process, $k-1$ points are distributed with uniform probability in the interval [0, 1] so that it ends up divided into *k* subintervals. Their lengths would represent the expected values for the *k* normalized weights p_{ij} according to the null hypothesis. The probability density function for one of these variables taking a particular value *x* is

$$\rho(x)dx = (k-1)(1-x)^{k-2}dx, \qquad \textbf{[1]}$$

which depends on the degree *k* of the node under consideration. In *Materials and Methods* we provide a detailed analysis of the null model with respect to the actual weight distribution in two real-world networks.

The disparity filter proceeds by identifying which links for each node should be preserved in the network. The null model allows this discrimination by the calculation for each edge of a given node of the probability α_{ij} that its normalized weight p_{ij} is compatible with the null hypothesis. In statistical inference, this concept is known as the *p* value, the probability that, if the null hypothesis is true, one obtains a value for the variable under consideration larger than or equal to the observed one. By imposing a significance level α, the links that carry weights that can be considered not compatible with a random distribution can be filtered out with an certain statistical significance. All the links with $\alpha_{ij} < \alpha$ reject the null hypothesis and can be considered as significant heterogeneities due to the network-organizing principles. By changing the significance level we can filter out the links progressively focusing on more relevant edges. The statistically relevant edges will be those whose weights satisfy the relation

$$\alpha_{ij} = 1 - (k-1)\int_0^{p_{ij}}(1-x)^{k-2}dx < \alpha. \qquad \textbf{[2]}$$

Table 1. Sizes of the disparity backbones in terms of the percentage of total weight (%W_T), nodes (%N_T), and edges (%E_T) for different values of the significance level α

U.S. airport network				Florida Bay food web			
α	%W_T	%N_T	%E_T	α	%W_T	%N_T	%E_T
0.2	94	77	24	0.2	90	98	31
0.1	89	71	20	0.1	78	98	23
0.05(a)	83	66	17	0.05	72	97	16
0.01	65	59	12	0.01	55	87	9
0.005	58	56	10	0.0008(a)	49	64	5
0.003(b)	51	54	9	0.0002(b)	43	57	4

See points a and b in Fig. 3.

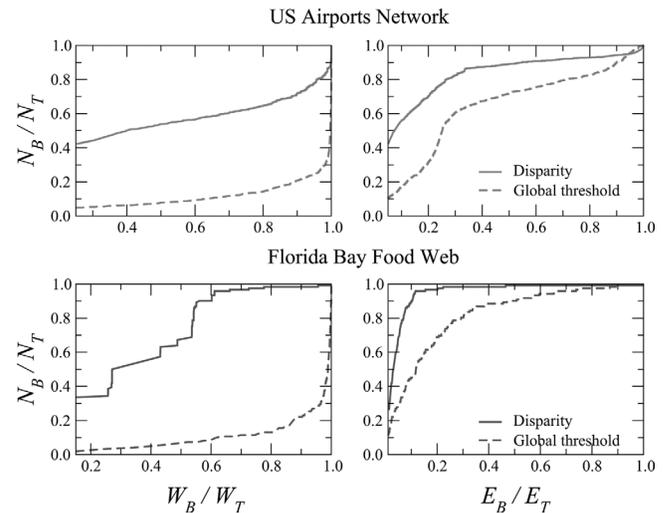**Fig. 1.** Fraction of nodes kept in the backbones as a function of the fraction of weight (*Left*) and edges (*Right*) retained by the filters.

Note that this expression depends on the number of connections k of the node to which the link under consideration is attached.

The multiscale backbone is then obtained by preserving all the links that satisfy the above criterion for at least one of the two nodes at the ends of the link while discounting the rest.* In this way, small nodes in terms of strength are not belittled so that the system remains in the percolated phase. In other words, we single out the relevant part of the network that carries the statistically relevant signal provided by the distribution with respect to local uniform randomness null hypotheses. By choosing a constant significance level α we obtain a homogeneous criterion that allows us to compare inhomogeneities in nodes with different magnitudes in degree and strength. By decreasing the statistical confidence, more restrictive subsets are obtained, giving place to a potential hierarchy of backbones. This strategy will be efficient whenever the level of heterogeneity is high and weights are locally correlated. Otherwise, the pruning could lose its hierarchical attribute producing results analogous to the global threshold algorithm (see section on networks with uncorrelated weights in *SI Appendix*).

The Multiscale Backbone of Real Networks. To test the performance of the disparity filter algorithm, we apply it to the extraction of the multiscale backbone of two real-world networks. We also compare the obtained results with the reduced networks obtained by applying a simple global threshold strategy that preserves connections above a given weight ω_c. As examples of strongly disordered networks, we consider the domestic nonstop segment of the U.S. airport transportation system for the year 2006 (http://www.transtats.bts.gov) and the Florida Bay ecosystem in the dry season (19). The U.S. airport transportation system for the year 2006 gathers the data reported by air carriers about flights between 1,078 U.S. airports connected by 11,890 links. Weights are given by the number of passengers traveling the corresponding route in the year symmetrized to produce an undirected representation. The resulting graph has a high density of connections, $\langle k \rangle = 22$, making difficult both its analysis and visualization. The Florida Bay food web comes from the ATLSS Project by the University of Maryland (http://www.cbl.umces.edu/atlss.html). Trophic interactions in food webs are symbolized by directed and weighted links representing carbon flows (mg C y^{-1}m^{-2}) between species. The network consists of a total of 122 separate components joined by 1,799 directed links.

In Table 1 and Fig. 1, we show statistics for the relative sizes—in terms of fractions of total weight W_T, nodes N_T, and edges E_T—preserved in the backbones when the network is filtered by the disparity filter and by the application of a global threshold, respectively. The disparity filter reduces the number of edges significantly even when the significance level α is close to 1, keeping at the same time almost all of the weight and a high fraction of nodes. Smaller values of α reduce even more the number of edges but, interestingly, the total weight and number of nodes remain nearly constant. Only for very low values of α—when the filter becomes very restrictive—do the total weight and number of nodes start decreasing significantly. In the case of the airports network, values around $\alpha \approx 0.05$ extract backbones with >80% of the total weight, 66% of nodes, and only 17% of edges. The global threshold filter, on the other hand, is not able to maintain the majority of the nodes in the backbone for similar values of retained weight or edges, as it is clearly seen in the first and second columns of Fig. 1, respectively.

It is particularly interesting to analyze the behavior of the topological properties of the filtered network at increasing levels of reduction. Fig. 2 shows the evolution of the cumulative degree distribution, i.e., $P_c(k) = \sum_{k' \geq k} P(k')$, for different values of α (*Left Top*) and ω_c (*Right Top*), respectively. The original airports network is heavy tailed although it cannot be fitted by a pure power-law function. Interestingly, the disparity filter reveals a clear power-law behavior as α decreases, with an exponent $\gamma \approx 2.3$. On the other hand, the global threshold filter produces subgraphs with a degree of distribution similar to the original one, but with a sharp cutoff that becomes smaller as the filter gets more restrictive. However, the weight distribution $P(\omega)$ for the disparity filter (*Left Middle*) shows that almost all scales are kept during the filtering process and only the region of very small weights is affected, in contrast to the global threshold filter that, by definition, cuts $P(\omega)$ off below ω_c (*Right Middle*).

In Fig. 2 (*Bottom*), we show the clustering coefficient C measured as the average over nodes of degree >1. It remains nearly constant in both filters until they become too restrictive, in which case clustering goes to zero.[†] In the case of the disparity filter, clustering remains constant up to values of $\alpha \approx 0.01$. This is precisely the value below which both the number of nodes and the weight in the backbone start decreasing significantly. Therefore, we can conclude that values of α in the range $[0.01, 0.5]$ are optimal, in the sense that backbones in this region have a large proportion of nodes and weight, the same clustering of the original network, and a stable stationary degree distribution, all

---

* In the case of a node i of degree $k_i = 1$ connected to a node j of degree $k_j > 1$, we keep the connection only if it beats the threshold for node j.

[†] The sudden increase of clustering for $E_B/E_T = 0.2$ is due to the reduction of the number of nodes in the network, increasing then the chances of having a random contribution.

Fig. 2. Topology of the filtered subgraphs for the U.S. airports network. (*Top*) Cumulative degree distribution, $P_c(k)$, for the disparity (*Left*) and global threshold (*Right*) backbones. The values of ω_c on the right plot are chosen to generate subgraphs with the same weight as the ones shown on the left plot. (*Middle*) Distribution of links' weights of the different subgraphs generated by the two filters. Symbols are the same as in the top plots. (*Bottom*) Clustering coefficient averaged over nodes of degrees >1 for the two methods as a function of the fraction of edges in the backbones. Dashed lines show the fraction of nodes and weight for a given fraction of edges.

Fig. 3. Fraction of edges in different global threshold backbones (GTB) included in the disparity backbone (DB) as a function of the significance level. As shown, points **a** and **b** in the U.S. airport network mark disparity backbones including a 100% of the 40 W and 10 W global threshold backbones, respectively; points **a** and **b** in the Florida Bay food web mark disparity backbones including a 100% of the 40 W and 13 W global threshold backbones, respectively. See also Table 1.

with a very small number of connections compared with the original network. It is important to stress that the disparity filtering also includes the connections with the largest weight present in the system. This is because the heavy tail of the $P(\omega)$ distribution is mainly determined by relevant large-scale weight. This is clearly illustrated in Fig. 3, where we show that for statistical significance levels up to $\alpha \simeq 10^{-3}$, all of the edges included in the 10–20% of the $P(\omega)$ tail are included in the extracted multiscale backbone.

As an illustration of the efficacy of the disparity filter, we visualize the obtained multiscale backbone in Fig. 4. In the case of the U.S. airport network we use the significance value $\alpha = 0.003$ [see entry (b) in Table 1 and Fig. 3]. Interestingly, the disparity filter offers a perspective of the network that reveals its geographic constraints (notice that each node is placed in the plane according to its actual coordinates on the earth). It is possible to identify local hubs with very well defined basins of attraction made of small airports connected to them (21), a star-like pattern that is particularly clear in Alaska airports or midwestern cities. In addition, the hierarchy of the transportation system is fully highlighted, including not just the most high flux connections but also small weight edges that are statistically significant because they represent relevant signal at the small scales. In this way, all important connection on the local and global level are considered at once. This would not be possible with a global threshold algorithm, which would simply eliminate all connections below the scale introduced by the cutoff threshold.

The Florida Bay food web is a directed network (see *SI Appendix* for an explanation of the methodology in the case of weighted directed neworks). We draw its multiscale backbone for $\alpha = 0.0008$, which contains the top 40% of heaviest links (see entry (a) in Table 1 and Fig. 3). Notice that, in this case, the concentration of weight in a few links is so important that the represented disparity backbone contains approximately half of the total weight in the network. Again, star motifs are uncovered, formed by mainly incoming connections, as for the pelican, or mainly outgoing ones, bivalves. More in general, specific subsystems dominated by significant fluxes can be easily identified, which might be evidence of an historical evolution of the network from smaller modular and disconnected structures to the complete ecosystem we observe today. Another interesting remark refers the presence in the backbone of species with relatively few trophic links. Species with few connections are usually assumed to have a low impact on the ecosystems. However, counterexamples can be found and such species may act as the structural equivalent of keystone species, whereas species with many trophic linkages may be more conceptually similar to dominant species (22). Because of its local approach, our filter mixes both types in the backbones, where simultaneously big hubs coexist—like the Predatory Shrimp, which in the complete network approximately has an average number of incoming connections and the maximum number of outgoing ones, 13 and 61, respectively—with more modest species in terms of connections—like Benthic Flagellates, with in-degree 1 and out-degree 10, both below the average.

Fig. 4. Pajek representations (20) of disparity backbones. (*Left*) The $\alpha = 0.003$ multiscale backbone of the 2006 domestic segment of the U.S. airport transportation system. This disparity backbone includes entirely the top 10% of the heaviest edges. (*Right*) The $\alpha = 0.0008$ multiscale backbone of the Florida Bay ecosystem in the dry season. This disparity backbone includes entirely the top 40% of the heaviest edges. These disparity backbones correspond to points (b) for the U.S. airport network and (a) for the Florida Bay food web in Table 1 and Fig. 3. The connection with maximum weight for the U.S. airport network is Atlanta-Orlando, with value $\omega_{max} = 1,290,488$ passengers/year and for the Florida Bay Food Web Free Bacteria to Water Flagellates with value $\omega_{max} = 12.90$ mg C y^{-1}m^{-2}.

Conclusions

The disparity filter exploits local heterogeneity and local correlations among weights to extract the network backbone by considering the relevant edges at all the scales present in the system. The methodology preserves an edge whenever its intensity is statistically not compatible with respect to a null hypothesis of uniform randomness for at least one of the two nodes the edge is incident to, which ensures that small nodes in terms of strength are not neglected. As a result, the disparity filter reduces the number of edges in the original network significantly, keeping, at the same time, almost all of the weight and a large fraction of nodes. As well, this filter preserves the cutoff of the degree distribution, the form of the weight distribution, and the clustering coefficient.

As a criticism, one could say that it only works in the case of systems with strong disorder, where the weights are heterogeneously distributed both at the global and local level. Nevertheless, all filters present limitations; one has to take them into account in relation to the problem under analysis. Which strategy is the most appropriate for a particular problem should be carefully judged and we cannot exclude the possibility that a combination of different techniques turns out to be the most appropriate. Yet, the ubiquitous presence of fluctuations and disorder spanning many length scales uncovered in many real networks provides a wide range of potential applications for the present methodology in biology (metabolic networks, brain, periodically regulated genes), information technology (Internet, World Wide Web), economics (World Trade Web), and finance (stock markets).

Materials and Methods

Local Heterogeneity of Edges' Weight. To assess the effect of inhomogeneities in the weights at the local level, for each node i with k neighbors one can calculate the function (17, 18)

$$\Upsilon_i(k) \equiv kY_i(k) = k\sum_j p_{ij}^2. \qquad [3]$$

The function $Y_i(k)$ has been extensively used in several fields as a standard indicator of concentration for more than half a century: in ecology (23), economics (24, 25), physics (26), and recently in the complex networks literature where it is known as the disparity measure (17). In all cases, $Y_i(k)$ characterizes the level of local heterogeneity. Under perfect homogeneity, when all the links share the same amount of the strength of the node, $\Upsilon_i(k)$ equals 1 independently of k, while in the case of perfect heterogeneity, when just one of the links carries the whole strength of the node, this function is $\Upsilon_i(k) = k$. An intermediate behavior is usually observed in real systems with $\Upsilon_i(k) \propto k^\alpha$ and the exponent close to 1/2. In this case, the weights associated with a node are then peaked on a small number of links with the remaining connections carrying just a small fraction of the node's strength. This is the situation

where our filter will be more useful, highlighting structures impossible to detect using the global threshold filter. In this way, the disparity function can be used as a preliminary indicator of the presence of local heterogeneities.

The Null Model. The probability density function of Eq. **1**, along with the joint probability distribution for two intervals given by

$$\rho(x,y)dxdy = (k-1)(k-2)(1-x-y)^{k-3}\Theta(1-x-y)dxdy, \qquad [4]$$
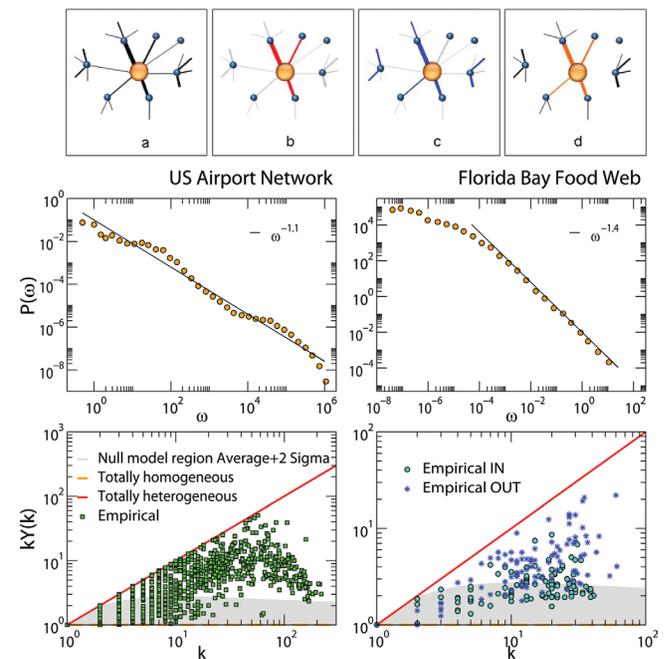
Fig. 5. Heterogeneity of weights at the local and global scales. (*Top*) Sequential diagram illustrating the disparity filtering technique at the local level. We focus on the central node in orange and its first neighborhood. (*a*) Original network; (*b*) edges of the central node with weights that are statistically significant heterogeneity; (*c*) the same for the neighbors; (*d*) intersection of the colored edges in B and C that are finally selected in the backbone. (*Middle*) Distribution of link's weights spanning for six decades. Even though this distribution does not have a clear functional form, a direct power-law fit of the form $\omega^{-\beta}$ yields an exponent $\beta = 1.1$, so with a diverging first moment. (*Bottom*) Scattered plot of the disparity measure for individuals airports of the U.S. airport network. The gray area corresponds to the average plus 2 standard deviations given by the null model.

Serrano et al.

where $\Theta(\cdot)$ is the Heaviside step function, can be used to calculate the statistics of $\Upsilon_{null}(k)$ for the null model. The average $\mu(\Upsilon_{null}(k)) = k\mu(Y_{null}(k))$ and the variance $\sigma^2(\Upsilon_{null}(k)) = k^2\sigma^2(Y_{null}(k))$ are found to be:

$$\mu(\Upsilon_{null}(k)) = \frac{2k}{k+1} \qquad [5]$$

$$\sigma^2(\Upsilon_{null}(k)) = k^2 \left(\frac{20+4k}{(k+1)(k+2)(k+3)} - \frac{4}{(k+1)^2} \right). \qquad [6]$$

Notice that the two moments depend on the degree k so that each node in the network with a certain degree k should be compared with the corresponding null model.

The observed values $\Upsilon_{ob}(k)$ compatible with the null hypothesis could be defined as those in the region between $\langle\Upsilon_{null}(k)\rangle + a\cdot\sigma(\Upsilon_{null}(k))$ and perfect homogeneity, so that local heterogeneity will be recognized only if the observed values lie outside this area,

$$\Upsilon_{ob}(k) > \mu(\Upsilon_{null}(k)) + a\cdot\sigma(\Upsilon_{null}(k)). \qquad [7]$$

The variable a is a constant determining the confidence interval for the evaluation of the null hypothesis. The larger it is the more restrictive becomes the null model and the more disordered weights should be for local heterogeneity to be detected. A typical value in analogy to gaussian statistics could be for instance $a = 2$.

As shown in Fig. 5, the overall distributions of weights for both networks considered here are very broad, with tails approaching power-law behaviors spanning six decades for the U.S. airport network and more than four for the Florida Bay food web. At the local level, $\Upsilon(k)$ measurements cannot be explained by the null model for most nodes.

1. Newman MEJ (2003) The structure and function of complex networks. *SIAM Rev* 45:167–256.
2. Dorogovtsev SN, Goltsev AV, Mendes JFF (2007) Critical phenomena in complex networks. arXiv:0705.0010v2 [cond-mat.stat-mech].
3. Caldarelli G (2007) *Scale-Free Networks* (Oxford Univ Press, Oxford).
4. Barabási A-L, Albert R (1999) Emergence of scaling in random networks. *Science* 286:509–512.
5. Barrat A, Barthélemy M, Pastor-Satorras R, Vespignani A (2004) The architecture of complex weighted networks. *Proc Natl Acad Sci USA* 101:3747–3752.
6. Jolliffe I (2002) *Principal Component Analysis* (Springer, New York), 2nd Ed.
7. Kim BJ (2004) Geographical coarse graining of complex networks. *Phys Rev Lett* 93:168701.
8. Song C, Havlin S, Makse HA (2005) Self-similarity of complex networks. *Nature* 433:392–395.
9. Itzkovitz S, et al. (2005) Coarse-graining and self-dissimilarity of complex networks. *Phys Rev E* 71:016127.
10. Gfeller D, los Rios PD (2007) Spectral coarse-graining of complex networks. *Phys Rev Lett* 99:038701.
11. Chalupa J, Leath PL, Reich GR (1979) Bootstrap percolation on a Bethe lattice. *J Phys C* 12:L31.
12. Kruskal JB (1956) On the shortest spanning subtree of a graph and the traveling salesman problem. *Proc Am Math Soc* 7:48–50.
13. Macdonald PJ, Almas E, Barabási A-L (2005) Minimum spanning trees of weighted scale-free networks. *Europhys Lett* 72:308.
14. Wu Z, Braunstein LA, Havlin S, Stanley HE (2006) Transport in weighted networks: Partition into superhighways and roads. *Phys Rev Lett* 96:148702.
15. Eguíluz VM, Chialvo DR, Cecchi GA, Baliki M, Apkarian AV (2005) Scale-free brain functional networks. *Phys Rev Lett* 92:028102.
16. Allesina S, Bodinia A, Bondavalli C (2006) Secondary extinctions in ecological networks: Bottlenecks unveiled. *Ecol Model* 194:150–161.
17. Barthélemy M, Gondran B, Guichard E (2003) Spatial structure Internet traffic. *Physica A* 319:633–642.
18. Almaas E, Kovács B, Vicsek T, Oltvai ZN, Barabási A-L (2004) Global Organization of metabolic fluxes in the bacterium *Escherichia coli*. *Nature* 427:839–843.
19. Ulanowicz RE, Bondavalli C, Egnotovich MS (1998) *Network Analysis of Trophic Dynamics in South Florida Ecosystem, FY 97: The Florida Bay Ecosystem*, Ref. No. [UMCES]CBL 98-123 (Chesapeake Biological Laboratory, Solomons, MD).
20. Batagelj V, Mrvar A (2003) *Visualization of Large Networks*, eds Jünger M, Mutzel P (Springer, Berlin), pp 77–103.
21. Barthélemy M, Flammini A (2006) Optimal traffic networks. *J Stat Mech*, L07002.
22. Dunne JA, Williams RJ, Martinez ND (2002) Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecol Lett* 5:558–567.
23. Simpson EH (1949) Measurement of diversity. *Nature* 163:688.
24. Herfindahl OC (1959) *Copper Costs and Prices: 1870-1957* (John Hopkins Univ Press, Baltimore, MD), pp 1-260.
25. Hirschman AO (1964) The paternity of an index. *Am Econ Rev* 54:761–762.
26. Derrida B, Flyvbjerg H (1987) Statistical properties of randomly broken objects and of multivalley structures in disordered systems. *J Phys A* 20:5273–5288.

Serrano et al.