Spectral analysis of protein-protein interactions in *Drosophila melanogaster*

Christel Kamp*
Blackett Laboratory, Imperial College, Prince Consort Road, London SW7 2AZ, United Kingdom

Kim Christensen†
Blackett Laboratory, Imperial College, Prince Consort Road, London SW7 2AZ, United Kingdom and Physics of Geological Processes, University of Oslo, P. O. Box 1048, Blindern, N-0316 Oslo, Norway

(Received 26 May 2004; revised manuscript received 7 February 2005; published 21 April 2005)

Within a case study on the protein-protein interaction network (PIN) of *Drosophila melanogaster* we investigate the relation between the network's spectral properties and its structural features. The frequencies of loops of any size within the network can be derived from the spectrum; also the prevalence of specific subgraphs as a result of the network's evolutionary history affects its spectrum. The discrete part of the spectral density shows fingerprints of the PIN's topological features including a preference for loop structures. Duplicate nodes are also characteristic for PINs and we discuss their representation in the PIN's spectrum as well as their biological implications.

DOI: 10.1103/PhysRevE.71.041911 PACS number(s): 87.14.Ee, 89.20.−a, 89.75.Hc, 89.75.Fb

I. INTRODUCTION

Network structures can be observed in most diverse domains ranging from biological and technological systems to social and economical systems [1]. Genetic regulatory networks, protein-protein interaction networks, and metabolic networks support the functions of life in any living organism. Technological networks such as the internet or the World Wide Web have a huge impact on our lives and societies. Networks of acquaintances and the exchange of information within these networks shape social and economical systems. The investigation of networks has a long tradition in random graph theory [2,3]. However, during the last few years high quality data on real-world networks have revealed that they cannot be adequately described by standard models from random graph theory. Much attention has been devoted to the derivation of rather specific quantities like degree distributions or clustering coefficients [4] that do not allow for a classification and understanding of network topologies within a broader and self-consistent framework.

Spectral graph theory is a promising avenue for obtaining a more comprehensive description of networks [5–8]. A network of N nodes can be described by its adjacency matrix $\mathbf{A}=(a_{ij})$ with entries

$$a_{ij} = \begin{cases} 1 & \text{if there is a link between node } i \text{ and } j, \\ 0 & \text{otherwise}. \end{cases} \tag{1}$$

The adjacency matrix is a symmetric, non-negative matrix in the case of undirected networks and accordingly has real eigenvalues $\{\lambda_j\}$, $j=1,\ldots,N$, being solutions of $\det(\mathbf{A}-\lambda\mathbf{I})=0$ [9]. The relation between structural features of a network and properties of its spectral density

$$\rho(\lambda) = \frac{1}{N}\sum_{j=1}^{N} \delta(\lambda - \lambda_j) \tag{2}$$

with respect to its adjacency matrix is a topic of the present research. While dense classical random networks exhibit a semicircular spectral density of the adjacency matrix [10], networks with a broad or scale-free degree distribution give rise to a broader spectrum [11–16]. The spectral density of sparse random networks (i.e., networks with a small average degree) has peaks at eigenvalues corresponding to the strong prevalence of finite trees in such networks [10,17].[1] We address the general question whether peaks in the spectral density correspond to a strong prevalence of specific subgraphs. The search for subgraphs that are statistically overrepresented relative to a null model, so-called motifs, recently gained much attention [20,21]. As a case study on the relation between these two approaches, we investigate the spectral properties of the protein-protein interaction network (PIN) of the fruit fly *Drosophila melanogaster* [22]. We emphasize that spectral analysis is not hampered by an *a priori* bias toward predefined quantities like motifs of a given size and accordingly provides an unbiased characterization of a network's topological features. We find that although there is no simple correspondence between network motifs and the peaks in the spectral density, the discrete part of the spectral density does contain fingerprints of the prevalence of specific subgraphs. Moreover, the spectral density is affected by duplicate nodes and contains exact information on the loop size frequencies.

II. THE SPECTRUM OF THE PIN OF *DROSOPHILA MELANOGASTER*

For our study we used the PIN of *Drosophila melanogaster* as given in [22] and available via the Database of

*Electronic address: mail@christelkamp.de; c.kamp@imperial.ac.uk
†Electronic address: k.christensen@imperial.ac.uk

[1] A similar phenomenon has earlier been discussed in the context of a model of quantum percolation [18,19].

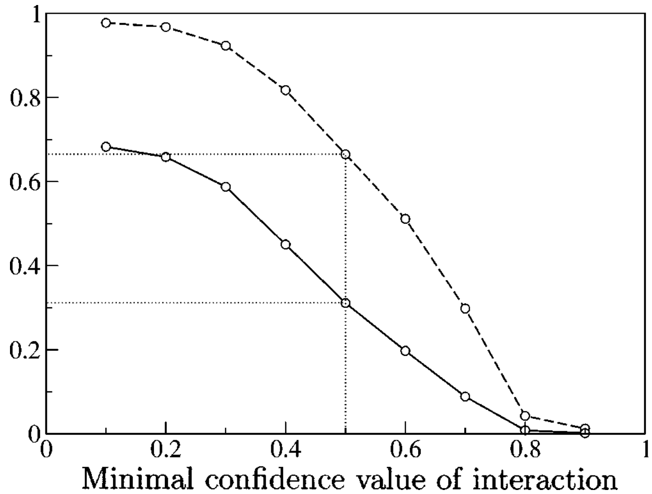

FIG. 1. The number of nodes divided by 10 000 (solid line) and the fraction of nodes (dashed line) in the largest connected component in the PIN as a function of the minimal confidence value of protein-protein interaction. We focus on the PIN defined by a minimal confidence value of 0.5, see dotted line.

Interacting Proteins [23]. The protein-protein interactions have been derived using the two-hybrid method which, however, is known to generate many false positives. Therefore each interaction in the network is classified by a confidence value between zero and 1 as described and provided in [22]. This defines a hierarchy of networks with increasing minimal confidence value for the protein-protein interactions. In Fig. 1 the size of the largest connected component in a network with a given minimal confidence value of interactions is shown.

For our further analysis we choose a network with a minimal confidence value of 0.5 which contains a total of 4681 proteins and 4794 interactions corresponding to an average degree $\langle k \rangle = 2.05$. The network is enriched with biologically meaningful interactions while it still shows a strong largest connected component (i.e., a giant component) containing about 2/3 of its nodes.

We determined the eigenvalues of the adjacency matrix corresponding to this PIN. The cumulative spectral density in Fig. 2(a) exhibits jumps at various eigenvalues which are represented by the discrete spectrum in Fig. 2(b).[2] Since about 2/3 of the network's nodes belong to its giant component and 49% of the eigenvalues in the network's spectrum are in the discrete spectrum, the emergence of spectral peaks cannot be explained by small isolated clusters alone.

A. The discrete spectrum and network motifs

To get a better understanding of the emergence of spectral peaks we compare the discrete spectrum with the corresponding spectra of two reference networks. First, we consider a randomized PIN of the same size and degree sequence but with randomized links following the procedure of [24]. Second, we consider a (classical) random network, that is, a network of the same size and average degree $\langle k \rangle = 2.05$, that is, a network with a probability $p = 0.000\,438$ for a link between any two nodes. In Fig. 3 the discrete spectrum of the adjacency matrix of the protein-protein interaction network of *Drosophila melanogaster* as well as of the two reference networks are shown, the latter being averages over ten reference networks.

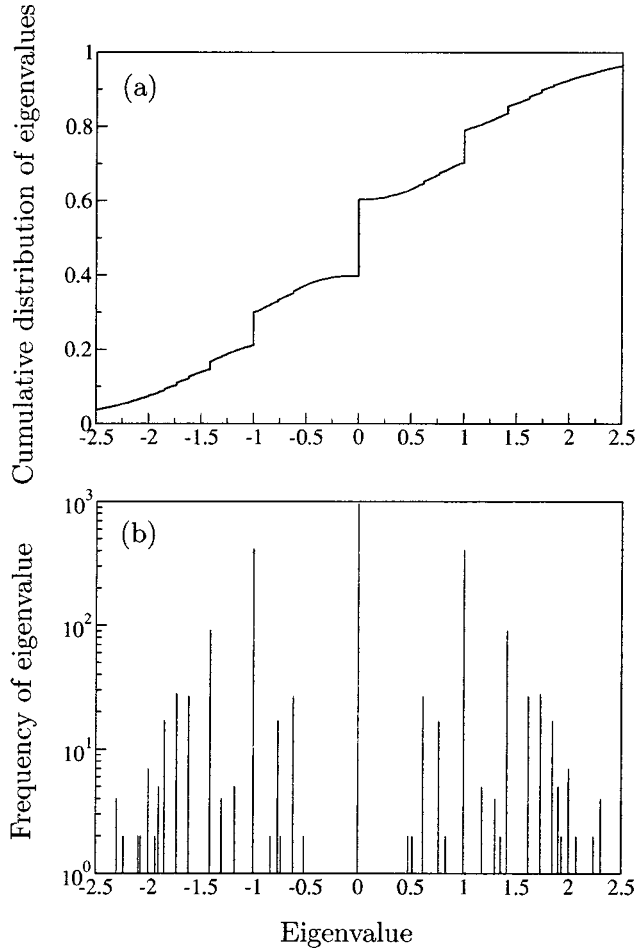

FIG. 2. Spectral analysis of the protein-protein interaction network of *Drosophila melanogaster*. (a) The cumulative spectral density. (b) The discrete frequency spectrum containing 49% of all eigenvalues.

To get more reliable results, we concentrate our further analysis only on eigenvalues that can be found more than twice in the spectrum of the original network. Qualitatively, we see that while the classical random network shows only a few peaks of that size corresponding to the eigenvalues of simple tree graphs (2.1, 3.1, 4.1, 4.2 in Fig. 4), additional eigenvalues appear in the discrete spectrum of the randomized PIN with the same degree sequence as the original network and eventually the original network (see Fig. 3). This change in the spectral properties indicates some differences in the structural organization of the underlying networks. In the following paragraphs, we will discuss how the observed hierarchy of spectral peaks reflects the networks' topologies and relates to other concepts like the search for motifs.

[2] An eigenvalue is considered to belong to the discrete spectrum if there exists at least one other eigenvalue that does not deviate by more than 10^{-12}.

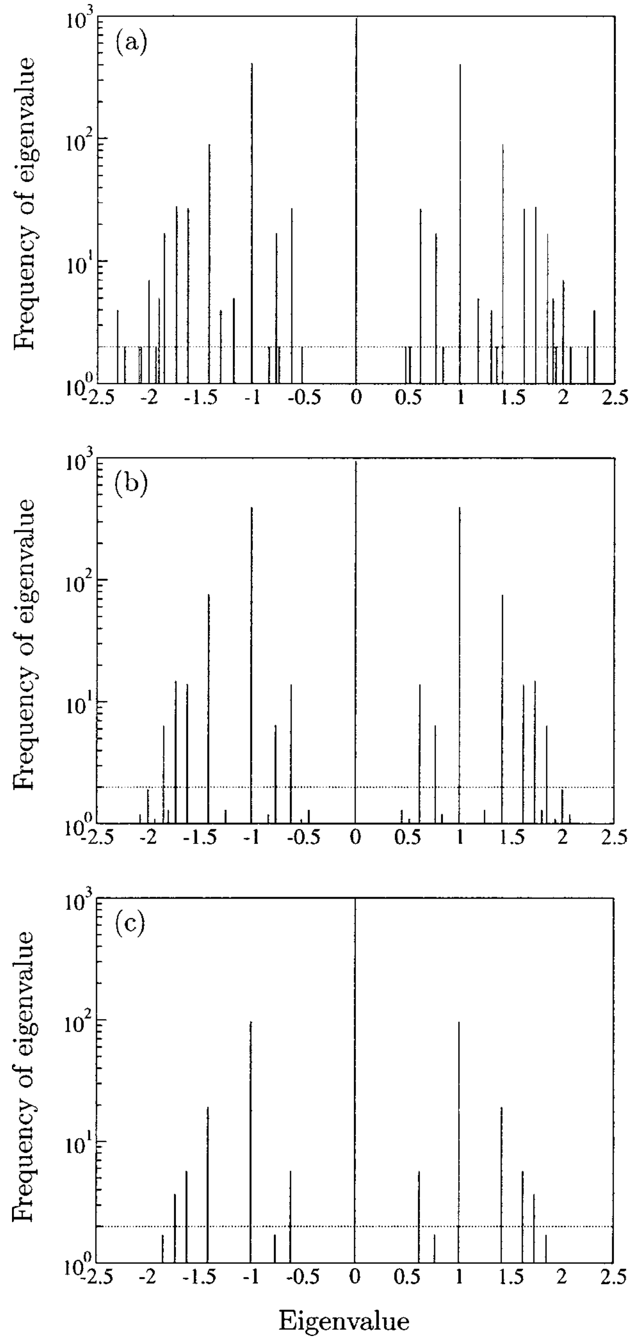

FIG. 3. The discrete frequency spectrum of (a) the PIN of *Drosophila melanogaster* containing 49% of all eigenvalues, (b) a randomized PIN with identical degree sequence containing 43% of all eigenvalues, and (c) a classical random network of identical size and average degree $\langle k \rangle = 2.05$ containing 27% of all eigenvalues.

Following the arguments of [17], we suggest that the prevalence of specific peaks in the discrete spectrum of a network corresponds to a strong representation of certain subgraphs. It has recently been shown that networks from different contexts show characteristic over-representation of specific subnetworks which are usually referred to as motifs [20,21]. Although motifs can be expected to leave marks in a network's spectrum, there is seemingly no simple correspondence between the eigenvalues of small subgraphs and spectral peaks. First, subgraphs are not generally represented by their eigenvalues in the spectrum of the whole network. Second, isospectral graphs are not necessarily isomorphic [6]. Nevertheless, a thorough comparative study of the discrete spectrum provides some insight into the networks' structure. In Figs. 4 and 5 we show the connected subgraphs up to size 5 with the full set of eigenvalues present in the discrete spectrum of the whole network. It shows that the spectrum of the PIN is more consistent with loop structures (cf. graphs 3.2, 4.3, 5.4, 5.5, 5.6, 5.8) than any randomized version which might hint to regulatory functionality supplied by this network. The eigenvalues behind these structures might correspond to eigenvalues of trees, e.g., the eigenvalues of a triangle (graph 3.2) or the box (graph 4.3, often also referred to as a bifan structure) might well be explained by graphs 2.1 and 5.3. However, to represent the eigenvalues of graphs 5.5 and 5.6 one has to consider trees of minimum size 8 and 7, respectively. The eigenvalues of graph 5.8 cannot be found among trees of size up to 10. Considering that the frequency of a given tree of size n in a sparse network decreases exponentially with n [10] and relating the findings in the PIN to those in the randomized reference networks we hypothesize that the spectral peculiarities reflect the loop structure in the original network.

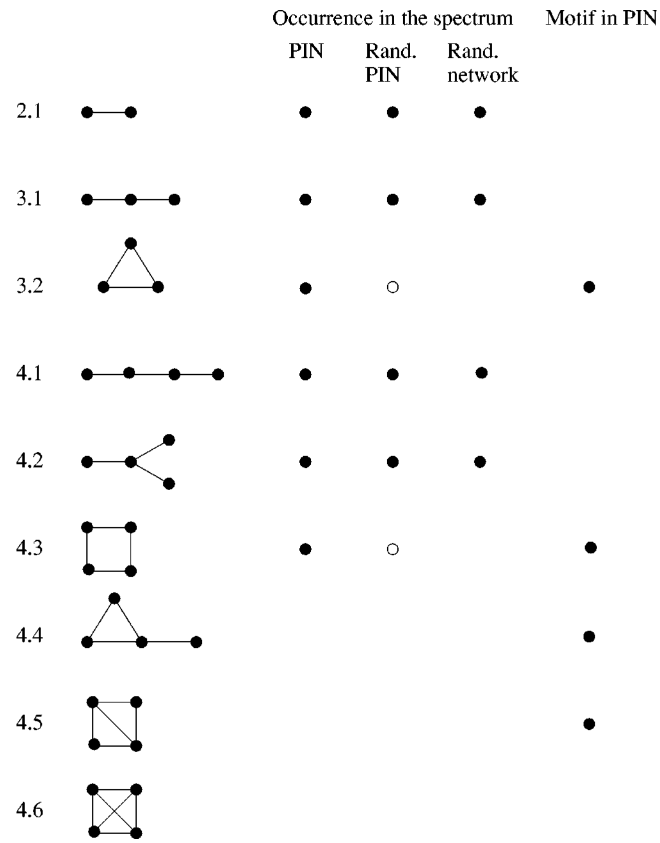

FIG. 4. Connected subgraphs with up to four nodes. A solid bullet (●) denotes that the eigenvalues of this graph can be found in the spectrum of the original network (PIN), the randomized network (Rand. PIN), or the classical random network (Rand. network), respectively. The rightmost column shows whether the subgraph is a motif according to the MFINDER software (default settings [20,25]). An open bullet (○) corresponds to single eigenvalue occurrences.

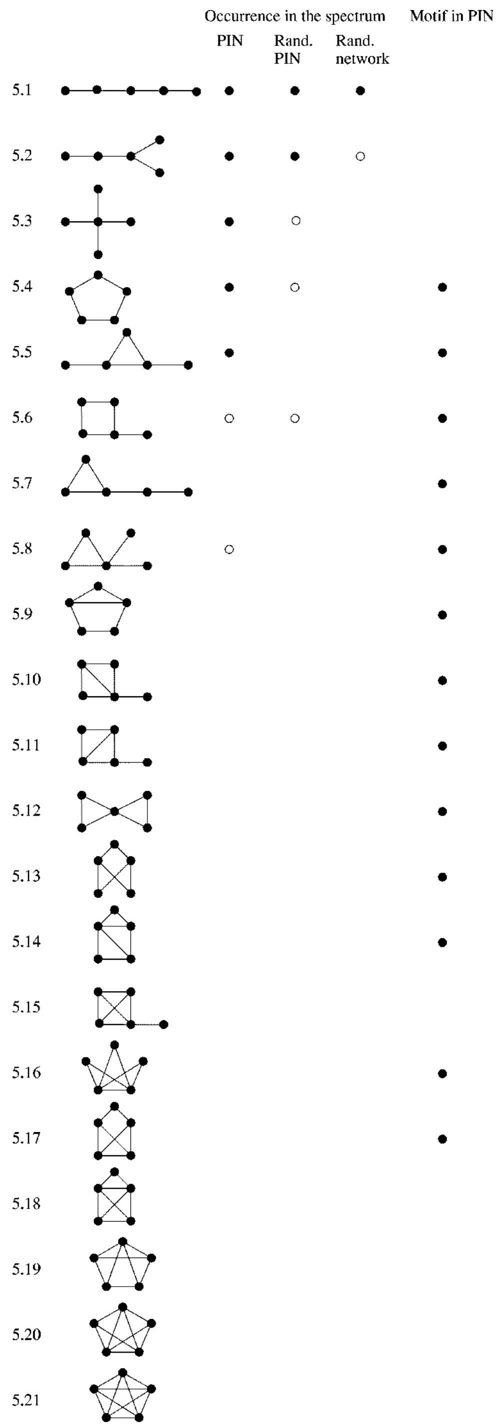

FIG. 5. Connected subgraphs with five nodes. A solid bullet (●) denotes that the eigenvalues of this graph can be found in the spectrum of the original network (PIN), the randomized network (Rand. PIN), or the classical random network (Rand. network), respectively. The rightmost column shows whether the subgraph is a motif according to the MFINDER software (default settings [20,25]). An open bullet (◯) corresponds to single eigenvalue occurrences.

To quantify the correspondence between the number of specific subgraphs in the PIN and the PIN's discrete spectrum we tried to decompose the spectrum into the contributions of connected subgraphs up to size 5. This, however, was not feasible indicating that higher order contributions, though being individually small, cannot be neglected as a whole.

Although we have to ascertain that there is no simple correspondence between subgraphs of a network and the prevalence of their eigenvalues in the discrete spectrum of the whole network we want to discuss the relation of spectral properties to the notion of motifs. According to the definition introduced in Ref. [20], a motif is a subnetwork that shows strong prevalence within the network relative to a randomized network. For our analysis we refer to the default requirements implemented in the MFINDER software [20,25], that is, a motif is a subgraph that occurs at least by two standard deviations more than in 100 randomized networks with the same degree sequence. In Figs. 4 and 5 the rightmost columns show which connected subgraphs up to size 5 are motifs in the PIN according to these criteria. There exist a lot of highly connected motifs while the spectrum reflects more the tree-structures in the network. However, the fingerprints of trees in the spectrum of both the original and the randomized PIN are consistent with the fact that they do not show up as motifs according to the above definition. One might further ask whether some motifs are hidden for spectral analysis because they are in fact building blocks of larger units. For example graph 5.5 as well as its subgraph 4.4 is a motif according to [20,25]. But only the eigenvalues of 5.5 can be found in the spectrum of the PIN. Moreover highly connected motifs do not occur in high (absolute) numbers and might accordingly be drowned in spectral analysis.

B. The circuitry of the PIN

In Sec. II A we have shown that the discrete spectrum of the PIN of *Drosophila melanogaster* favors the eigenvalues of loopy subgraphs. This observation derived from the investigation of distinct local structures and their eigenvalue representations can be confirmed by an assessment of the whole set of eigenvalues. Evaluating the trace of the matrix $\mathbf{A}^k$

$$\mathrm{Tr}(\mathbf{A}^k) = \sum_{i=1}^{N} \lambda_i^k \tag{3}$$

yields the number of directed loops of length k in the underlying network [6,11] as shown in Fig. 6, though neglecting details of the graphs underlying the loops. Note that loops of even length might be trivial going back and forth in a tree while odd loops are nontrivial. The difference in growth rates of the numbers of loops of increasing size between the original network and the classical random network is likely due to the strong fragmentation of the latter (many isolated nodes). However, the strong relative prevalence of loops of odd length in the original network is more remarkable with respect to the networks' topologies. This becomes obvious from Fig. 6(b) showing the number of loops of a given size in the original PIN normalized to the numbers in the two reference networks. While tree graphs have only trivial loops of even length, loops of odd length indicate nontrivial loops which confirms the results derived from the evaluation of the discrete spectrum on the basis of eigenvalue representations of small subgraphs.

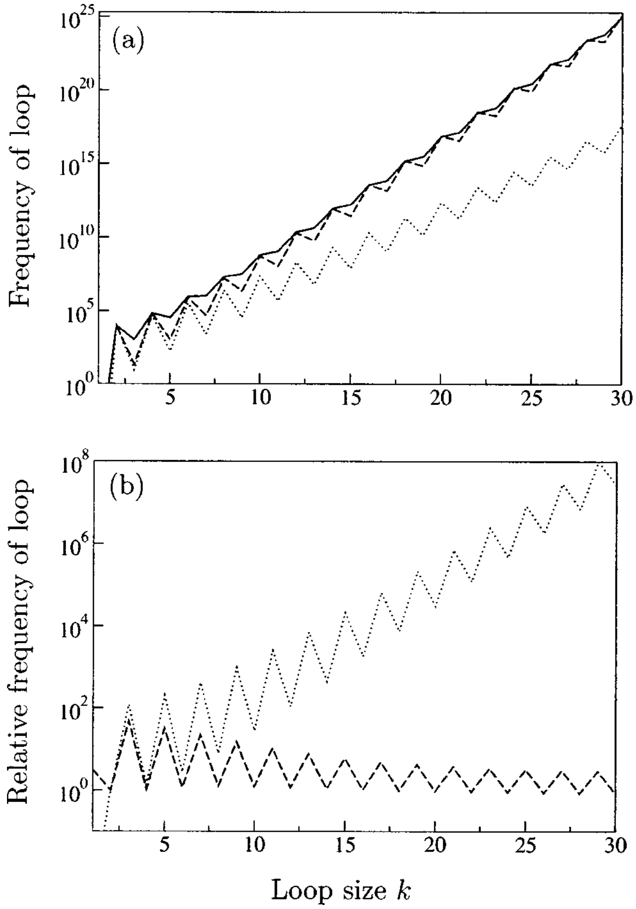

FIG. 6. (a) The frequency of loops of size k in the PIN of *Drosophila melanogaster* (solid line), a randomized PIN with identical degree sequence (dashed line), and a classical random network of identical size and average degree $\langle k \rangle = 2.05$ (dotted line). Odd cycles represents nontrivial loops, that is, deviations from a treelike structure in the networks. (b) The relative frequency of loops of size k in the PIN of *Drosophila melanogaster* with respect to the randomized PIN (dashed line) and the classical random network (dotted line).

The analysis of a network's spectrum could reveal some structural information about the network as a whole. Exact conclusions about general properties like the prevalence of loops are possible. As spectral analysis does not count predefined motifs of a given size it is not hampered by an *a priori* bias. It is a challenging question of future research to further investigate the relationship between a network's spectrum and its topological features, e.g., in terms of motifs, to get a more rigorous and *unbiased* characterization of a network's topological features.

III. FINGERPRINTS OF DUPLICATION

The evolution of many biological networks and specifically PINs is assumed to be strongly driven by duplication (and diversification) of nodes in the network [26,27]. The genomes underlying the PIN of many organisms have undergone a few whole genome duplications complemented by many single-gene duplications [28,29]. After duplication,

TABLE I. The number of nodes that are duplicate and the number of duplicate links (i.e., the number of neighbors associated with these nodes) in the original network and the two reference networks, that is, the difference in the number of nodes and links between the original network and the reduced network. Isolated nodes have been neglected.

	Duplicate nodes	Duplicate links
PIN *D. melanogaster*	686	728
Randomized PIN	626.0 ± 22.0	629.1 ± 22.0
Random network	151.0 ± 14.2	151.0 ± 14.2

one of the duplicates usually diverges from its original appearance, possibly providing new functionality. The concept of duplication has similarly been recognized to be important for functional roles in a network motif [30]. The search for fingerprints of the evolutionary history of a PIN naturally has to include an assessment of duplicate nodes, that is, those that share the same interaction partners. Each set of duplicate nodes represents an equivalence class also referred to as an orbit. The reduced network is a network in which all nodes of an orbit are reduced to one node.

Table I shows that the PIN has more duplicate nodes (with associated links) than the reference networks. This is reflected in an increased number of zero eigenvalues in the PIN relative to the randomized networks for which theoretical foundation is given in the Appendix. Figure 7 shows the frequency of orbits of a given size in the original as well as in the reference networks. The distribution of orbit sizes in the original network is very close to the one found in the randomized PIN with the same degree sequence, but much broader than that of the classical random network.

Again, spectral analysis offers a complementary approach to the topic. Using the results from the Appendix, we determined the eigenvalues of those graphs that arise from duplication of the two simplest reduced graphs: a line and a tri-

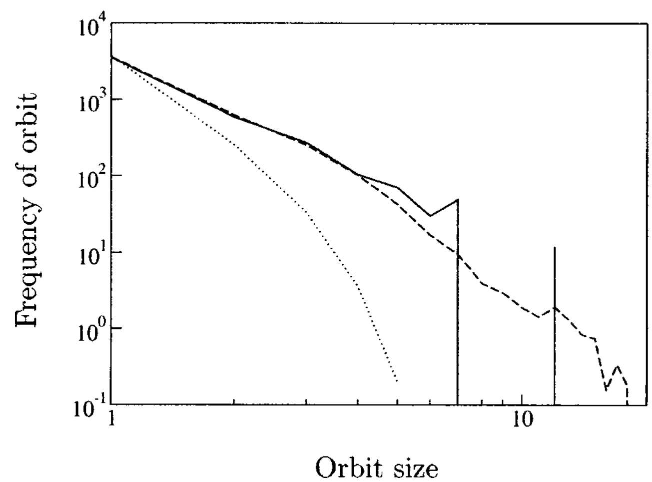

FIG. 7. The frequency of orbit size in the PIN of *Drosophila melanogaster* (solid line), a randomized PIN with identical degree sequence (dashed line), and a classical random network of identical size and average degree $\langle k \rangle = 2.05$ (dotted line). Isolated nodes have been neglected.

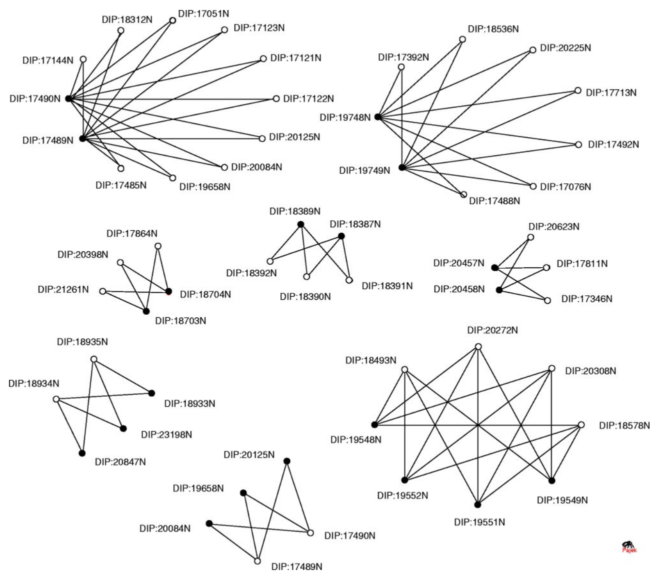

FIG. 8. Subgraphs of nodes that form orbits of size $\geqslant 2$ and that have degree $\geqslant 2$ (black); orbits of size 2 of nodes with degree 2 have been omitted. The white nodes are the neighbors of duplicate nodes that may have more neighbors than shown. The nodes' labels are their identifiers from the Database of Interacting Proteins [23].

angle (graphs 2.1 and 3.2 in Fig. 4). We allowed for up to ten duplications of each node of the reduced network and searched for the eigenvalues of the resulting subgraphs. However, spectral analysis is only consistent with the emergence of star graphs and the original triangle as well as the box or bifan structure (graph 4.3 in Fig. 4).

Considering the representation of star graphs in the spectrum of the PIN one might guess that the high frequency of large orbits mainly reflects nodes with many leaves. A look at the joint distribution of the size of an orbit and the degree of its nodes in the original and the reference networks supports this hypothesis. In 100 reference networks (of both kind) the nodes in an orbit larger than one have degree one, that is, only nodes with degree one have duplicates. Only in extremely rare cases do nodes with degree two have a single duplicate. This matches the global situation in the original network, however, there are some remarkable exceptions with nodes of high degree in large orbits shown in Fig. 8 that cannot be found in the reference networks. This is also well in accordance with the values found in Table I. Different from the original network, in both reference networks the number of duplicate links is practically the same as the number of duplicate nodes.

From the Database of Interacting Proteins [23] and FlyBase [31] we derived names and descriptions (if available) for the proteins in Fig. 8 as shown in the web page supplementary to this article [32]. We find that duplicate proteins are likely to have similar functionality in accordance with results in the yeast PIN [33].

IV. SUMMARY AND CONCLUSIONS

Recent developments in the research on complex networks have increased the understanding of a network's topology and its connection to functionality. However, a comprehensive theory of networks incorporating classical graph theory as well as recent findings into a self-consistent framework still has to be worked out. Considering spectral graph theory to be a promising ansatz for this attempt, we have done a case study on the PIN of *Drosophila melanogaster*. The eigenvalues of a network's adjacency matrix (and of related matrices) provide information about a network's structural properties like the number of connected components, its diameter or characteristics of its degree distribution. Here, we have put special emphasis on the investigation of the discrete spectrum of a sparse network relating it to prevalent substructures. The identification of these substructures, or motifs, is computationally very expensive as the number of subgraphs in a network grows exponentially with the number of nodes in the subgraph [34]. Moreover there is no algorithm known to solve graph isomorphism problem in polynomial time which is essential to establish well-defined motifs. However, together with a once generated dictionary of subgraph spectra spectral analysis can be done in polynomial time. Although it will probably not be possible to derive the densities of specific subgraphs from the spectrum of a network we could show that structural prevalences on a more abstract level are reflected in the (discrete) spectrum of the PIN under investigation. While we here focused on the appearance of loops in subgraphs as well as the whole PIN future analysis might reveal further topological features.

Considering the evolutionary history of PINs we also discussed the appearance of proteins that share their neighbors together with the fingerprints of these structures that can be found in the network's spectrum. Studying structures of duplicate proteins in more detail we find that they often have close functional relationships in accordance with earlier findings in yeast.

Members of a so-called orbit have to have exactly the same neighborhood. This is very restrictive, though required to define equivalence classes of proteins that all mutually share the same neighborhood, that is to allow for transitivity. To generalize this approach one can define a similarity measure that quantifies the overlap of the neighborhoods of two nodes. This similarity measure can be interpreted as a distance measure between nodes in which the distance between nodes is the smaller the more similar their neighborhoods are. The application of a clustering algorithm in the associated metric space might give further insight into local structures.

This case study shows connections between a network's spectral and topological properties and points to its further systematic assessment. It is a challenging task, however, it can bring important insight into a network's structure in a less biased and more systematic way than currently available.

ACKNOWLEDGMENTS

C. Kamp would like to thank A. Bunten for providing computing facilities and software and for many helpful discussions on technical problems. I am grateful to N. Farid, S.A. Teichmann, and J. Leal for the discussions and many helpful comments on the manuscript. Also, we would like to acknowledge the hospitality of the Department of Physics of the University of Oslo during the period of finalizing this manuscript. This work was supported by the German Academic Exchange Service (DAAD).

APPENDIX

Let $\mathbf{A}$ be an $N \times N$ matrix representing an undirected graph, i.e., a symmetric matrix with entries $a_{ij} \in \{0,1\}$ and $a_{ii}=0$. Let $\mathbf{D}$ be the matrix that is obtained after m perfect duplications of nodes or in other words by m duplications of rows and columns, respectively. Let $i_1,\ldots,i_l$, $l \in \{1,\ldots,N\}$, be the number (identifier) of (mutually different) nodes that have been duplicated and $m_{i_1},\ldots,m_{i_l}$ be the corresponding number of duplications per node with $m=\Sigma_{j=1}^{l} m_{i_j}$. Let $\mathbf{A}_{i_j}$ the matrix $\mathbf{A}$ but with the element $a_{i_j i_j}$ replaced by $a_{i_j i_j}+\lambda$. Analogously, the matrix $\mathbf{A}_{i_1,\ldots,i_l}$ corresponds to the matrix $\mathbf{A}$ but with $a_{i_1 i_1},\ldots,a_{i_l i_l}$ replaced by $a_{i_1 i_1}+\lambda$, …, $a_{i_l i_l}+\lambda$. Let furthermore $\mathbf{I}$ be the identity matrix. Then the following equation holds:

$$\begin{aligned}\det(\mathbf{D}-\lambda\mathbf{I}) = &\\ &(-\lambda)^m \det(\mathbf{A}-\lambda\mathbf{I})\\ &+(-\lambda)^m \sum_{r\leq l} m_{i_r} \det(\mathbf{A}_{i_r}-\lambda\mathbf{I})\\ &+(-\lambda)^m \sum_{r\leq l-1}\sum_{r<j\leq l} m_{i_r} m_{i_j} \det(\mathbf{A}_{i_r i_j}-\lambda\mathbf{I})\\ &\vdots\\ &+(-\lambda)^m \sum_{r\leq 1}\sum_{r<j\leq 2}\cdots\sum_{x<y\leq l} m_{i_r}\cdots m_{i_y} \det(\mathbf{A}_{i_r,\ldots,i_y}-\lambda\mathbf{I}).\end{aligned} \tag{A1}$$

Note that the last term is equivalent to $m_{i_1}\cdots m_{i_l} \det(\mathbf{A}_{i_1,\ldots,i_l} - \lambda\mathbf{I})$. It gets obvious from this formula that perfect duplication of nodes only adds zeros to the spectrum of the graph.

Equation (A1) can be proven by induction. Considering a graph with adjacency matrix $\mathbf{A}$ in which an arbitrary node i is duplicated m times leading a duplication matrix $\mathbf{D}$ one can show that

$$\det(\mathbf{D}-\lambda\mathbf{I}) = (-\lambda)^m[\det(\mathbf{A}-\lambda\mathbf{I}) + m\det(\mathbf{A}_i-\lambda\mathbf{I})]. \tag{A2}$$

After validating the case of $l=0$, $m=0$ of Eq. (A1) we do the induction by evaluation of the adjacency matrix $\tilde{\mathbf{D}}$ of a graph generated from the duplication graph represented by $\mathbf{D}$ by duplicating (a nonduplicate) node $i_{l+1} m_{i_{l+1}}$ times. Therefore, we apply (A2)

$$\det(\tilde{\mathbf{D}}-\lambda\mathbf{I}) = (-\lambda)^m_{i_{l+1}}[\det(\mathbf{D}-\lambda\mathbf{I}) + m_{i_{l+1}}\det(\mathbf{D}_{i_{l+1}}-\lambda\mathbf{I})]$$

and derive $\det(\mathbf{D}-\lambda\mathbf{I})$ and $\det(\mathbf{D}_{i_{l+1}}-\lambda\mathbf{I})$ using the assumption (A1) yielding the formula (A1) for $m+m_{i_{l+1}}$ duplications of $l+1$ mutually different nodes.

As an example, this formula is applied to the 2×2 matrix $\mathbf{A}$ corresponding to two connected nodes. Then, $i_1=1$, $i_2=2$, and one gets for the matrix D after $m=m_1+m_2$ duplications

$$\det(\mathbf{D}-\lambda\mathbf{I}) = (-\lambda)^m(\lambda^2 - 1 - m_1 - m_2 - m_1 m_2),$$

$$\lambda_{1;2} = \pm\sqrt{1+m+m_1 m_2}.$$

Translating this into the number of nodes per orbit $n_i=m_i+1$ leads to the eigenvalues

$$\lambda_{1;2} = \pm\sqrt{n_1 n_2}.$$

[1] S. H. Strogatz, Nature (London) **410**, 268 (2001).
[2] B. Bollobás, *Modern Graph Theory* (Springer, New York, 1998).
[3] B. Bollobás, *Random Graphs* (Cambridge University Press, Cambridge, U.K., 2001).
[4] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[5] F. Chung, *Spectral Graph Theory*, Regional Conference Series in Mathematics Vol. 92 (American Mathematical Society, Providence, RI, 1994).
[6] D. Cvetković, M. Doob, and H. Sachs, *Spectra of Graphs*, Pure and Applied Mathematics, Vol. 87 (Academic Press, New York, 1980).
[7] D. Cvetković, M. Doob, I. Gutman, and A. Torgašev, *Recent Results in the Theory of Graph Spectra*, Annals of Discrete Mathematics Vol. 36 (North-Holland, Amsterdam, 1988).
[8] N. Biggs, *Algebraic Graph Theory* (Cambridge University Press, Cambridge, U.K., 1996).
[9] Introductions to spectral analysis and spectral graph theory of varying levels can be found at http://www.sosmath.com/matrix/eigen0/eigen0.html, http://en.wikipedia.org/wiki/Eigenvalue, http://en.wikipedia.org/wiki/Adjacency_matrix, and http://www.win.tue.nl/~aeb/srgbk/node5.html
[10] M. Bauer and O. Golinelli, J. Stat. Phys. **103**, 301 (2001).
[11] I. Farkas, I. Derényi, A.-L. Barabási, and T. Vicsek, Phys. Rev. E **64**, 026704 (2001).
[12] S. Dorogovtsev, A. Goltsev, J. Mendes, and A. Samukhin, Phys. Rev. E **68**, 046109 (2003).
[13] F. Chung, L. Lu, and V. Vu, Ann. Comb. **7**, 21 (2003).
[14] F. Chung, L. Lu, and V. Vu, Proc. Natl. Acad. Sci. U.S.A. **100**, 6313 (2003).
[15] M. Mihail and C. Papadimitriou, in *Proceedings of the 6th International Workshop on Randomization and Approximation Techniques*, 2002 Lecture Notes in Comp. Sci., Vol. 2483 (Springer, New York) pp. 254–262.
[16] K.-I. Goh, B. Kahng, and D. Kim, Phys. Rev. E **64**, 051903 (2001).
[17] O. Golinelli, e-print cond-mat/0301437.
[18] S. Kirkpatrick and T. Eggarter, Phys. Rev. B **6**, 3598 (1972).
[19] S. Evangelou, Phys. Rev. B **27**, 1397 (1983).
[20] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, Science **298**, 824 (2002).
[21] S. Milo, R. Itzkovitz, N. Kashtan, R. Levitt, S. Shen-Orr, I.

Ayzenshtat, M. Sheffer, and U. Alon, Science **303**, 1538 (2004).
[22] L. Giot *et al.*, Science **302**, 1727 (2003).
[23] Database of Interacting Proteins (DIP), available at http://dip.doe-mbi.ucla.edu/
[24] S. Maslov and K. Sneppen, Science **296**, 910 (2002).
[25] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, Computer code MFINDER, available at http://www.weizmann.ac.il/mcb/UriAlon/groupDownloadableData.html
[26] S. Ohno, *Evolution by Gene Duplication* (Springer, Berlin, 1970).
[27] S. Teichmann and M. Babu, Nat. Genet. **36**, 492 (2004).
[28] G. Amoutzias, D. Robertson, S. Oliver, and E. Bornberg-Bauer, EMBO Rep. **5**, 1 (2004).
[29] G. Amoutzias, D. Robertson, and E. Bornberg-Bauer, Comput. and Funct. Genomics **5**, 79 (2004).
[30] N. Kashtan, S. Itzkovitz, R. Milo, and U. Alon, Bioinformatics **20**, 1746 (2004).
[31] FlyBase, A Database of the *Drosophila* Genome, available at http://flybase.bio.indiana.edu/
[32] http://www.cmth.ph.ic.ac.uk/people/k.christensen/PRE/supp.html
[33] M. Samanta and S. Liang, Proc. Natl. Acad. Sci. U.S.A. **100**, 12579 (2003).
[34] S. Itzkovitz, R. Milo, N. Kashtan, G. Ziv, and U. Alon, Phys. Rev. E **68**, 026127 (2003).