nature physics

Article

https://doi.org/10.1038/s41567-022-01866-8

Laplacian renormalization group for heterogeneous networks

In the format provided by the authors and unedited

Supplementary Information: Laplacian Renormalization Group for heterogeneous networks

Pablo Villegas,[1] Tommaso Gili,[2] Guido Caldarelli,[3,4,5,6,7,*] and Andrea Gabrielli[1,8]

[1] *'Enrico Fermi' Research Center (CREF), Via Panisperna 89A, 00184 - Rome, Italy*
[2] *Networks Unit, IMT Scuola Alti Studi Lucca, Piazza San Francesco 15, 55100- Lucca, Italy.*
[3] *DMSN, Ca' Foscari University of Venice, Via Torino 155, 30172 - Venice, Italy*
[4] *European Centre for Living Technology (ECLT), Dorsoduro 3911, 30123 - Venice, Italy*
[5] *Institute for Complex Systems (ISC), CNR, UoS Sapienza, Piazzale Aldo Moro 2, 00185 - Rome, Italy*
[6] *London Institute for Mathematical Sciences (LIMS), W1K2XF London, United Kingdom*
[7] *Fondazione per il Futuro delle Città (FFC), Via Boccaccio 50, 50133 - Firenze Italy*
[8] *Dipartimento di Ingegneria, Università degli Studi 'Roma Tre', Via Vito Volterra 62, 00146 - Rome, Italy*

CONTENTS

* Corresponding author: guido.caldarelli@unive.it

1. REAL-SPACE RENORMALIZATION GROUP

The development of the renormalization group is a significant milestone that allowed for dealing with some of the most challenging problems of physics, ranging from relativistic quantum field theory to critical phenomena or the Kondo effect. Moreover, it allowed solving two critical problems in physics, namely, the task of solving systems with many degrees of freedom and the description of how the cooperative features of collective behavior arise [1]. Here we briefly describe the key RG concepts, making the analogy with RG diffusion on complex networks easier.

Let us consider a Hamiltonian $\mathcal{H}(\sigma(x))$, where $\sigma(x)$ represents the field of local mean spin and thus is a continuum variable. The main idea of the RG is to find an *effective Hamiltonian* at scale λ, $\mathcal{H}(\sigma) \mapsto \mathcal{H}_\lambda(\sigma)$.

One of the ideas to develop the RG process was proposed in 1966 by Kadanoff to deal with the real-space RG problem. Let us illustrate the procedure for the specific case of the Ising model. The three-step procedure (illustrated in Figure 1) to perform RG in real-space is:

1. Group the lattice points into groups of b sites.
2. Replace each block with a single one of size b, where $\mathcal{R}_g(\sigma)$ represents some renormalization group transformation as, e.g., the majority rule.
3. Rescale all scales by a factor b to return to the original lattice spacing.

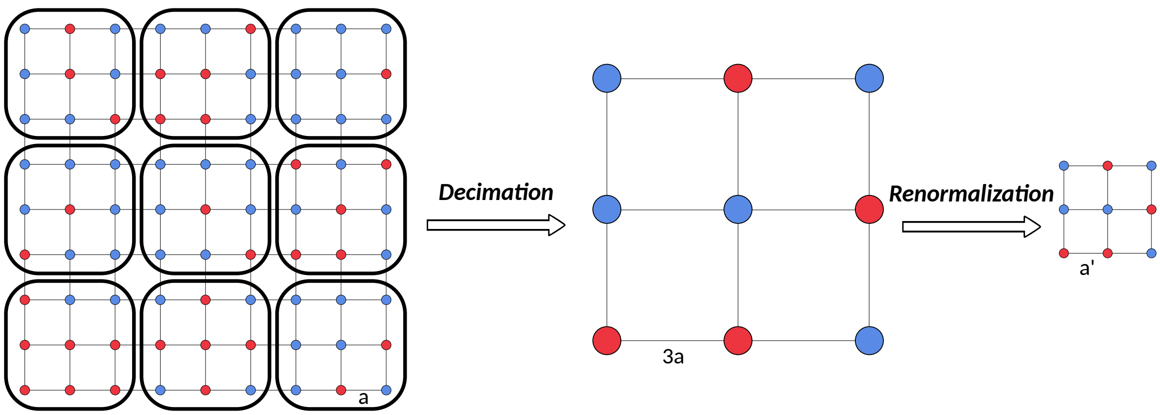

FIG. 1: Sketch the decimation process employing Kadanoff blocks on a square lattice of side a, where blue (red) points represent up (down) spins, respectively. The lattice is divided into several blocks with b^2 sites which are now coarse-grained, replacing it with a single-block following some rule $\mathcal{R}_g(\sigma)$, finally reducing all the system scales by a factor b. This scheme produces a reduced version of the original system.

In the paradigmatic case of the Ising model, the renormalization group operates at any value of the temperature, T. However, the Kadanoff procedure induces particular flows towards the trivial fixed points $T = 0$ for any spin configuration with $T < T_c$ and $T = \infty$ for any spin configuration with $T > T_c$. For $T = T_c$, there is no flow associated with RG transformations [2], and it is a repulsive fixed point of the flow along with the (tuning) control parameter, which reflects the self-similar properties of the system. Figure 2 shows the particular application of Kadanoff blocks to the Ising model for the three possible cases: $T < T_c$, $T = T_c$, and $T > T_c$.

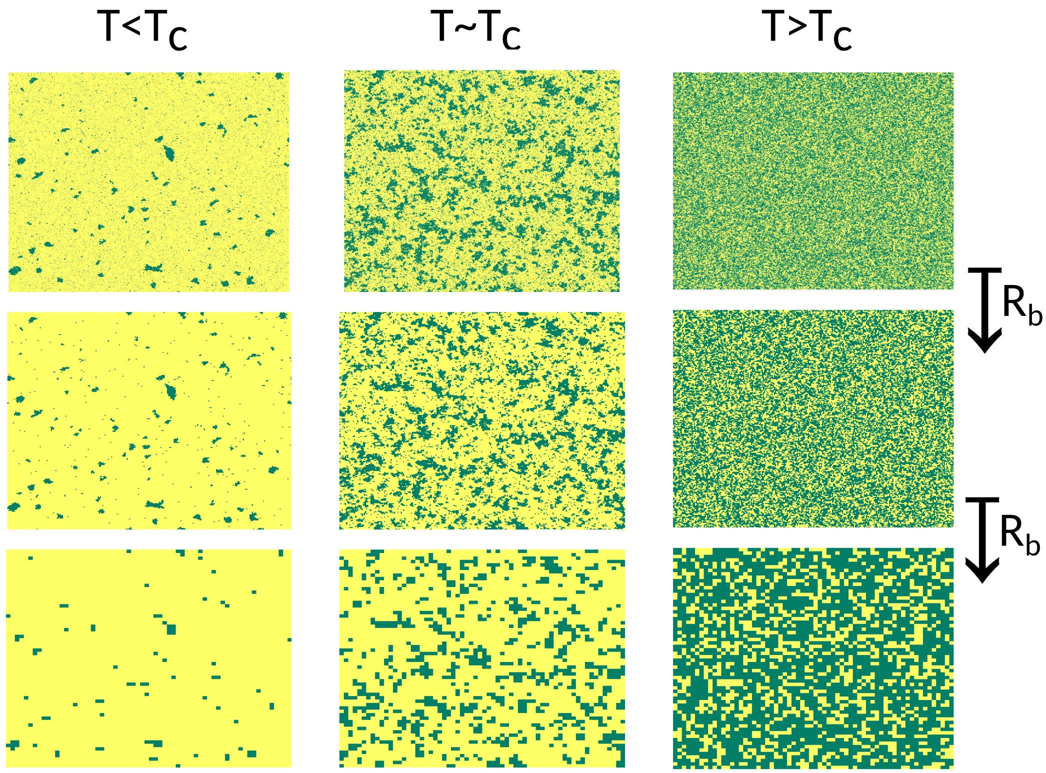

FIG. 2: (a) RG process for the Ising model simulated in a 2D lattice of $L = 1024$ (upper panels), and employing Kadanoff blocks of size $b = 3$. The three panels correspond to different temperatures $T < T_c$, $T \sim T_c$, and $T > T_c$, respectively. The RG process is performed from top to bottom scales of the figure, revealing the RG flow of the dynamical process.

1.1. The *Laplacian RG* for graphs as a *natural extension* of the ordinary RG in statistical physics

The renormalization group (RG) is the fundamental tool in statistical physics and statistical field theory to study and describe the physical behavior of systems showing a multi-scale and/or scale-invariant behavior as, for instance, it happens in the proximity of a critical point of a second-order phase transition at and out of equilibrium. It comprises three fundamental ingredients, namely:

1. A coarse-graining procedure: it can be performed either (a) in real space, through the derivation of mesoscopic collective variables (i.e., block variables) from a local resummation of the microscopic ones inside identical mesoscopic cells of size Λ tiling the whole space, or (b) in the conjugate Fourier space by rewriting the model through the Fourier modes of the microscopic variables and explicitly integrating all the modes with wavelength smaller than an arbitrary mesoscopic value Λ. In both cases, the new lower cut-off of the model is given, respectively by Λ. For instance, in the Ising model, we can either define *à la* Kadanoff mesoscopic magnetic moments summing the microscopic spins inside identical cells, or *à la* Wilson, through the Hubbard-Stratonovich transformation, we can rewrite the model in terms of a continuous spatial field whose Gaussian approximation is diagonal in Fourier space.

2. Reformulation of the model (i.e., redefinition of the interaction constants of the system) in terms of the block variables in real space and of the variable modes with wavelength larger than the new lower cut-off.

3. Spatial rescaling, so that the new lower cut-off becomes the new unit length.

If we want to extend this approach to models defined on heterogeneous graphs, the fundamental step to overcome to formulate a correct RG approach is the first one. Indeed, given a homogeneous model (i.e., with interaction constants identical for all spatial points as the Ising model with nearest-neighbor interactions, for instance), the definition of coarse-graining in identical mesoscopic cells or integrating small wavelength modes all over the space is immediate both in homogeneous spaces (e.g., $\mathbb{R}^d$) or homogeneous lattices or trees due to the continuous or discrete translational invariance of the space itself. This fundamental property is completely lost in heterogeneous networks. Consequently, both the definition of real space block variables and the resummation procedure of small wavelengths of the model in Fourier space may appear completely arbitrary or even meaningless.

Our Laplacian scheme for connected undirected graphs tackles this point as a natural extension of the coarse-graining in homogeneous spaces. First, let us notice that the Laplacian operator ∇^2 in such spaces can be strictly related to the progressive coarse-graining procedure in the RG: it is an operator with the spectrum of eigenvalues given by eigenvalues $-k^2$, with k Fourier wave-vector, i.e., the inverse of the wavelength, which has no characteristic scale. The corresponding eigenvectors are exactly given by the Fourier basis of plane waves e^{ikx}, which are also eigenfunctions of the translation group. In other words, the Laplacian operator is a sort of telescopic "scanner" of the coarse-graining scales. Second, the diffusion equation $\partial_t \rho = -\hat{L}\rho$ with $\hat{L} = -\nabla^2$ couples the spatial scales "scanned" by ∇^2 to corresponding diffusion times, which are proportional to the square of the wavelengths, or, inversely, at each value of the time. The diffusion, therefore, reaches from each point an identical length of the order of the square root of the time itself, defining in this way identical coarse-graining cells around each point *à la* Kadanoff. In other words, the coarse-graining *à la* Wilson in a homogeneous space can be performed by integrating out all wavelengths smaller than the square root of an arbitrary diffusion time and then increasing this time.

The symmetric graph Laplacian operator $\hat{L} = \hat{D} - \hat{A}$ is the natural graph discrete representation of the continuous operator $-\nabla^2$. It exactly describes the same diffusion dynamics $\dot{\rho} = -\hat{L}\rho$ on a heterogeneous connected network with the difference that for a given time, such dynamics cover different structures at different locations on the network due to the connectivity heterogeneity of the space. However, these structures share the fundamental property of being covered by diffusion simultaneously. In this strict sense, our formulation of Laplacian RG for graphs can be seen as the natural extension to heterogeneous networks of the usual RG approach in statistical physics and statistical field theory. Being impossible to define identical coarse-graining cells or an integration scheme over small wavelengths due to the lack of spatial translational invariance, i.e., due to the topological inhomogeneity of the space, we can, however, adopt the Laplacian operator and the diffusion equation as a tool to define the coarse-graining procedure both on homogeneous spaces and on inhomogeneous graphs.

Therefore, we expect to follow the LRG flow when performing LRG transformations on top of networks lacking scale-free properties. Hence, to gain analytical insight into the LRG approach employing Kadanoff supernodes, here we first analyze the case of pertinent scale-dependent networks, i.e., those who present an isolate resolution single scale being responsible for the physical behavior of macroscopic quantities as, e.g., Erdös-Renyi networks or stochastic block models (SBM). In complete analogy with Kadanoff transformations, such networks are expected to follow the LRG towards a trivial fixed point of the network structure, yet reflecting crucial mesoscales.

2. SCALE-DEPENDENT NETWORKS

2.1. Erdös-Renyi network

Figure 3(b) shows the diffusion of information over the network for simple Erdös-Rényi graphs. The particular structure of these graphs, lacking any sign of scale-invariance, is reflected in the single-peak of C, where information diffuses all over the network. Note that any Kadanoff transformation in the vicinity of that peak will automatically make collapse the whole network into a single node, as described in the main text. Just before τ^*, it is possible to perform some RG steps, even if the network properties constantly change following the LRG flow. Figure 3(a) shows a particular case for successive RG transformations with $\tau = 0.25$ (within the greyish area of Figure 3(b)), reflecting two types of effects: (i) Network properties continuously change, evidencing the LRG flow towards a trivial fixed point. (ii) ER networks converge rapidly to the trivial fixed point where the network acts as a single node.

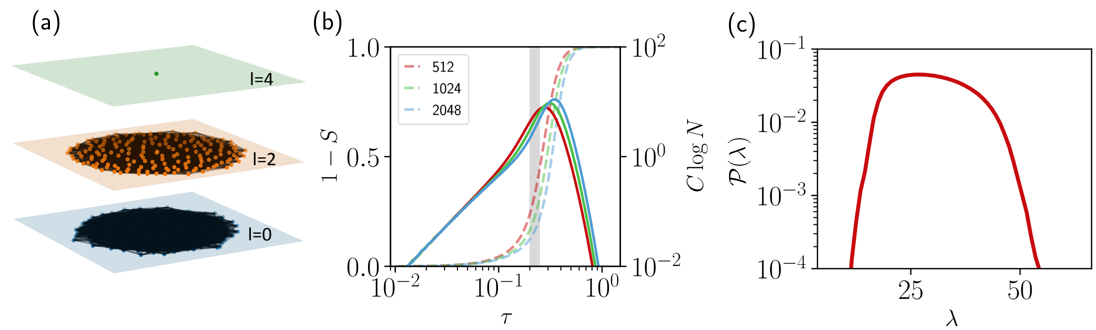

FIG. 3: (a) LRG process for Erdös-Rényi networks of $N = 256$ employing $\tau = 0.25$. Each layer represents the LRG step of order l. (b) Entropy parameter (dashed lines, $(1 - S)$), and specific heat (solid lines, C), versus the temporal resolution parameter of the network, τ for Erdös-Rényi networks of different systems sizes (see legend) with mean connectivity $\langle\kappa\rangle = 30$. (c) Probability distribution of Laplacian eigenvalues, $\mathcal{P}(\lambda)$, for an Erdös-Rényi network with $N = 4096$ and $\langle\kappa\rangle = 30$.

In particular, Figure 4 illustrates the evolution of the averaged degree distribution over network realizations, $P(\kappa)$, for subsequent LRG transformations and different resolution scales, τ. Note that, as soon as any coarse grain is performed, the mean connectivity flows undoubtedly towards smaller values, according to the results presented in the main text, until it collapses in a trivial fixed point, as shown in Figure 3(a) and Figure 4(c).

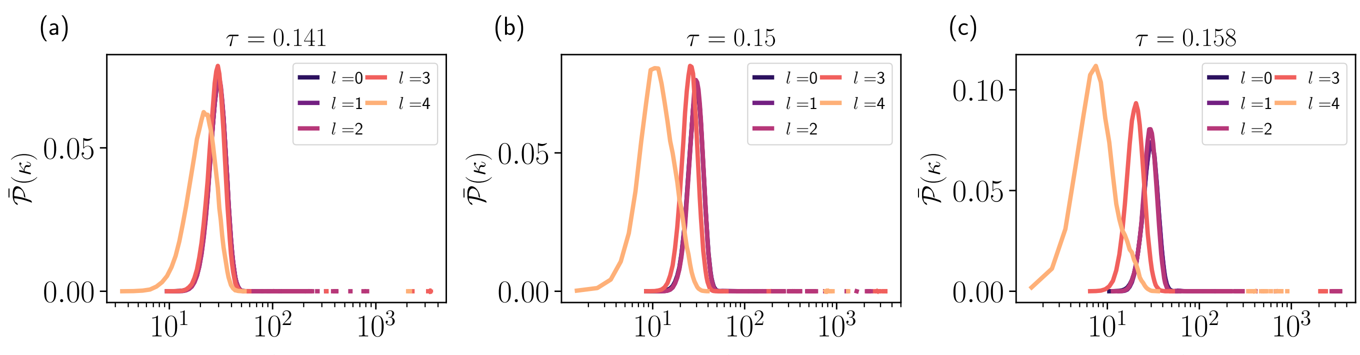

FIG. 4: Degree distribution after subsequent LRG transformations for an Erdös-Rényi network with $N = 4096$ nodes and mean connectivity $\kappa = \langle 30\rangle$ for different values of τ (see title): (a) $\tau = 0.141$, (b) $\tau = 0.15$, and (c) $\tau = 0.158$. Note that, as soon as we perform any blocking of nodes, the degree distribution, $P(\kappa)$, suddenly changes flowing through a trivial fixed point. All curves have been averaged over 10^2 network realizations.

2.2. Stochastic-block model

Despite the network convergence through a trivial fixed point due to its intrinsic lacking of scale-invariant properties, the study of LRG flow in networks with characteristic scales can shed light on its characteristic scales and functional modules. We stress here the particular case of a stochastic block model, composed of four different modules and different intra and interconnectivity between them. As shown in Figure 5, a detailed analysis close to τ^* makes emerge the four modules at a mesoscopic scale at some point of the RG process. Finally, the network will always collapse at a trivial fixed point.

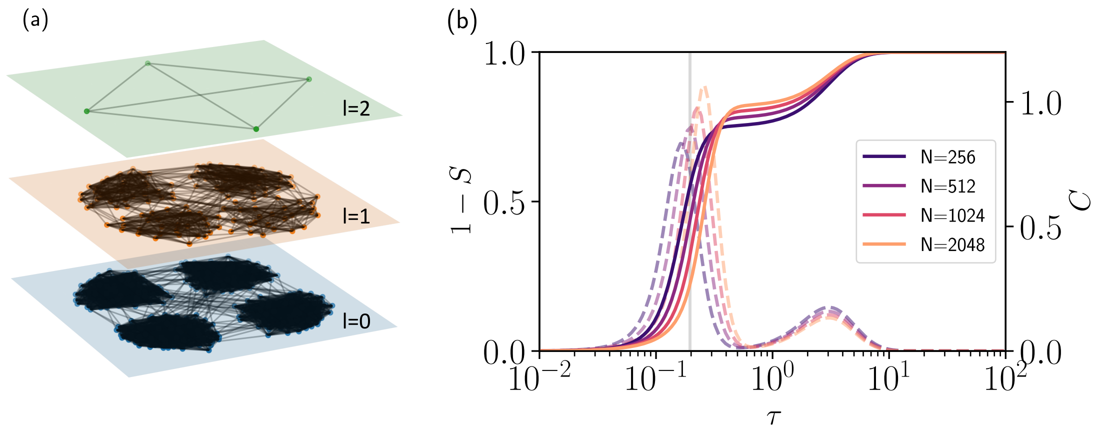

FIG. 5: (a) LRG process for a SBM of size $N = 256$ employing $\tau = 0.2$. Each layer represents the LRG step of order l. (b) Entropy parameter for the stochastic block model considering different network sizes (see legend) constituted by four equal interconnected modules $p = 128/N$ and interconnectivity probability $q = 1/N$. The two peaks in the derivative of the order parameter C indicate the two critical network resolution scales.

2.3. Small-world networks

One of the main problems in analyzing RG procedures, or network reduction, is to deal with small-world or ultra small-world problems giving place to short path length between all nodes in the network structure, thus generating an infinite spectral dimension. Here we show the capability of our method to perform networks reduction even in the worst scenario where network scales are strongly interlinked.

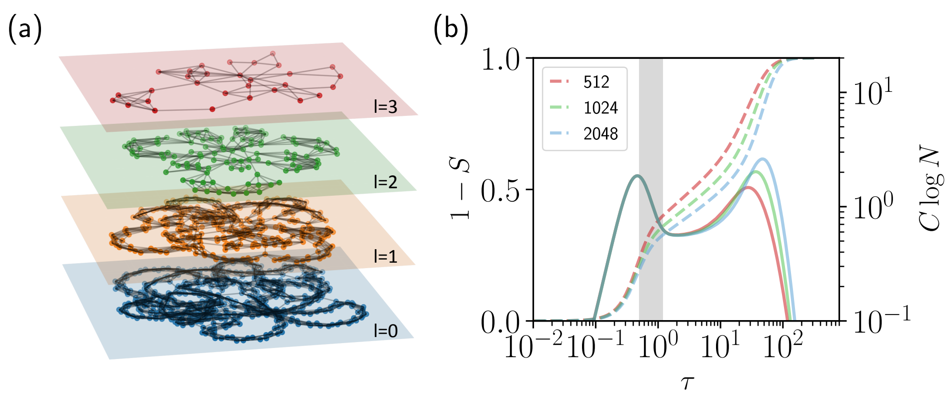

FIG. 6: (a) LRG process for a small-world network (SW) of size $N = 512$ employing $\tau = 0.8$. Each layer represents the LRG step of order l. (b) Entropy parameter for a SW network with rewiring probability $p = 0.02$ considering different network sizes (see legend). The two peaks in the specific heat, C, indicate the two critical network resolution scales: the underlying lattice and the long-range connections which allow for information to be distributed and processed across the entire network.

Figure 6(a) shows the LRG procedure for a small-world network [3] setting a certain probability of long-range connections which allow for information to be distributed and processed across the entire network. Note the particular shape of the entropy parameter at Figure 6(b). Regardless of the discussions about the scale-invariant nature of small-world networks, let us pinpoint here a critical issue. The way these networks are built strongly constrain the large-time specific heat peak. Note that its particular position reduces in time due to the effects of the long-range connections, even if the local scales of the network (a regular lattice in this case) remain unaltered. The same effect is expected to occur in more complex structures, where the local properties of the system will determine the dynamics while small-world effects facilitate the information to be distributed and processed across the entire network in shorter times. In summary, our LRG scheme allows us individuating different mesoscales and communities even in the presence of small-world substantial effects.

3. SCALE-INVARIANT NETWORKS

3.1. Random trees

We also analyze the simple case of simple trees and random trees, presenting complementary results to those of the main text. Random trees have been generated using the bijection between Prüfer sequences and trees. In particular, Figure 7 shows the LRG diffusion scheme when applied over these two network structures. As it can be seen, both structures remain scale-invariant after subsequent LRG transformations.

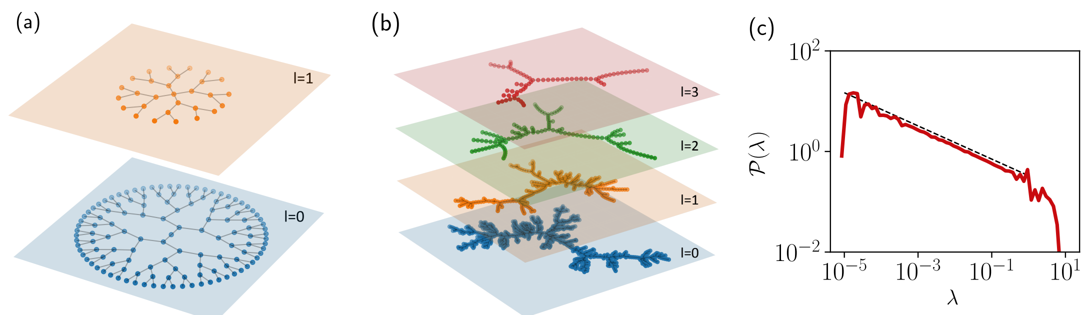

FIG. 7: (a) LRG process for a regular tree with branching factor $r = 2$, and $h = 6$ hierarchical levels. (b) Four-step LRG transformation for a random tree with $N = 1024$ nodes using $\tau = 3.5$. (c) Probability distribution of Laplacian eigenvalues for random tree with $N = 4096$. Dashed line is a guide to the eye with power law exponent $\gamma = -1/3$.

Figure 8 illustrates the evolution of the averaged degree distribution over network realizations, $P(\kappa)$, for subsequent LRG transformations and different resolution scales, τ. As we expected, the connectivity distribution remains invariant, according to the results presented in the main text.

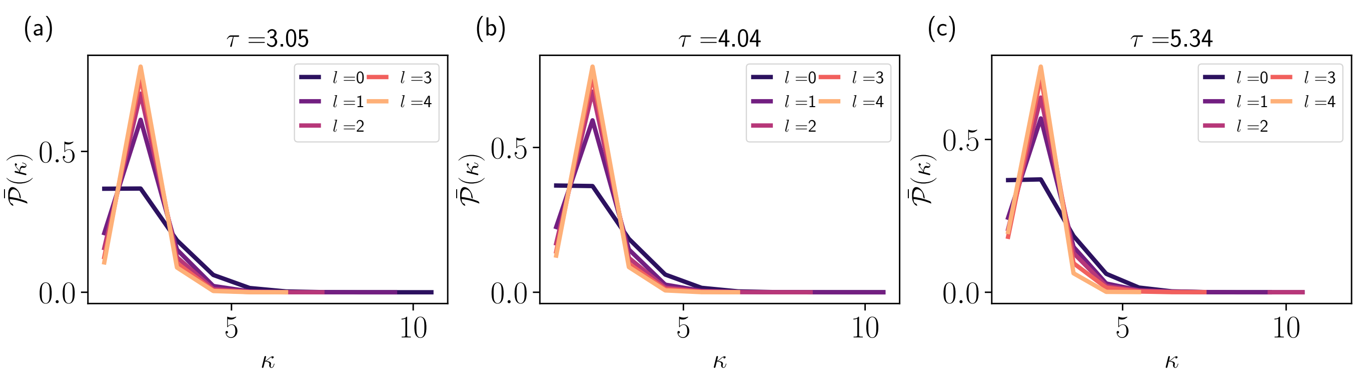

FIG. 8: Degree distribution after subsequent LRG transformations for random trees with $N = 4096$ nodes for different values of τ (see title): (a) $\tau = 3.05$, (b) $\tau = 4.04$, and (c) $\tau = 5.34$. The degree distribution, $P(\kappa)$, remains invariant under succesive LRG transformations. All curves have been averaged over 10^2 network realizations.

3.2. 2D Lattices

For the sake of consistency and to avoid spurious results, we study here the simplest trivial scale-invariant structure: a regular two-dimensional lattice. As shown in Figure 9(b), a small τ^* peak reflects the characteristic resolution scale of the system, allowing us to perform the RG procedure over the lattice reducing it to smaller lattices for several LRG steps (see Figure 9(a)). Note that LRG transformations are consistent over all the greyish areas of Figure 9(b), after which Kadanoff blocks become too large and finite-size effects can alter the renormalization process.

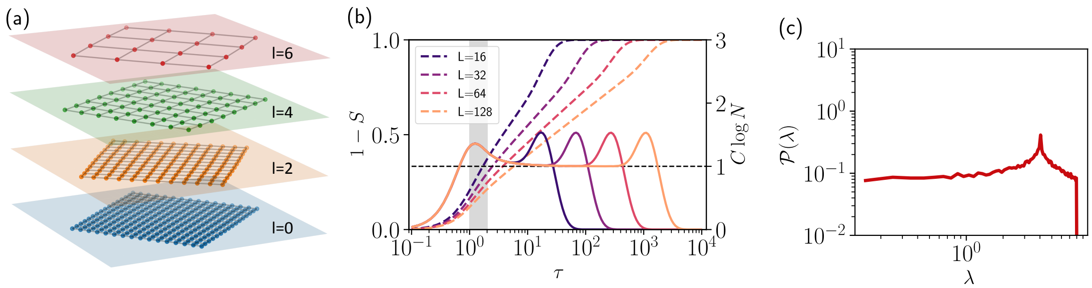

FIG. 9: (a) LRG process for a 2D lattice of size $N = 16^2$ employing $\tau = 1.5$. Each layer represents the LRG step of order l. (b) Entropy parameter for the 2D lattices considering different lattice sizes ($N = L^2$, see legend). The first peak allows us to identify the critical network scale to perform the RG reduction. (c) Probability distribution of Laplacian eigenvalues for a 2D lattice of size $N = 128^2$. Note that the spectrum presents a characteristic scale at $\lambda = 4$ (the number of neighbors), while the flat region leads to the constant value $C = 1$.

4. SCALE-FREE NETWORKS

4.1. Barabasi-Albert networks

We present here analyses of different Barabasi-Albert networks modifying the number of connecting nodes at each time-step in the building procedure of the network (m), which it is known to critically constrain the spectral dimension of the network. Figure 10(a) and (b) shows the evolution of the entropy parameter as a function of the resolution scale, τ, for different system sizes. Pay close attention to the fact that, for any value $m > 1$, information flows much faster all across the network due to small-world effects, and τ^*, the peak at short-time scales, disappears. In particular, it is yet possible to robustly perform network LRG transformations even if –for $m > 1$– the flat-region for small eigenvalues disappears, therefore, leading to an absence of scale-invariant properties in these type of networks.

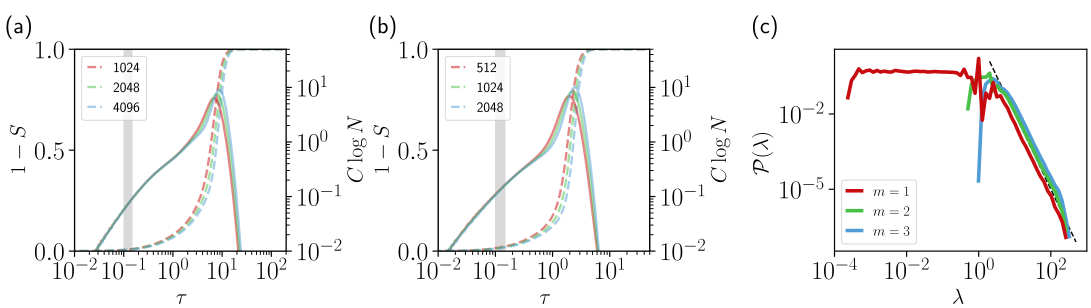

FIG. 10: (a) Entropy parameter (dashed lines, $(1 - S)$), and specific heat (solid lines, C), versus the temporal resolution parameter of the network, τ for BA networks of different systems sizes (see legend) with: (a) $m = 2$, and (b) $m = 4$. (c) Probability distribution of Laplacian eigenvalues for a BA with $N = 4096$ and different m values. Note that only for $m = 1$ the BA network exhibits a flat region leading to a constant value of the specific heat of the network. All curves have been averaged over 10^2 network realizations.

As shown in Figure 11 (a) and (b) the degree distribution, $P(\kappa)$ remains invariant when we perform subsequent RG transformations over the network, also conserving the mean-connectivity and the different intrinsic network properties.

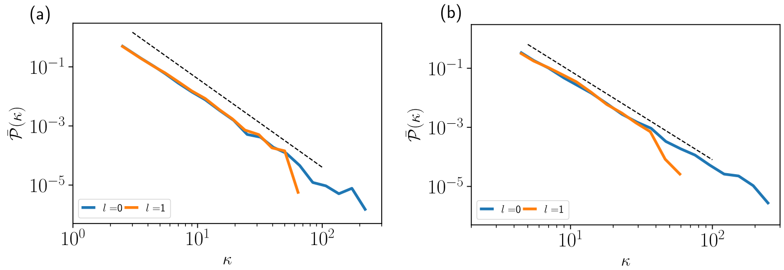

FIG. 11: Degree distribution after subsequent LRG transformations for Barabasi-Albert networks with $N = 4096$ nodes and $\tau = 0.1$ for different values of m: (a) $m = 2$, (b) $m = 4$. The degree distribution, $P(\kappa)$, remains invariant under LRG transformations. All curves have been averaged over 10^2 network realizations.

4.2. Configurational model

The different analyses on top of BA networks reveal fundamental differences on the network behavior depending on the particular value of m, even if the exponent in the degree distribution remains unchanged ($\gamma = 3$). We analyze now the effect of different scale-free networks, generated through the configurational model, the paradigmatic method for generating random networks from a given degree sequence. In particular, given a degree sequence of the general form $P(\kappa) \propto \kappa^{-\gamma}$, we choose the specific values $\gamma = 2.2$ and $\gamma = 3.5$ to ensure the robustness of the RG method under the convergence of different moments of the degree distribution (the only condition is to ensure the existence of $\langle \lambda \rangle_t$).

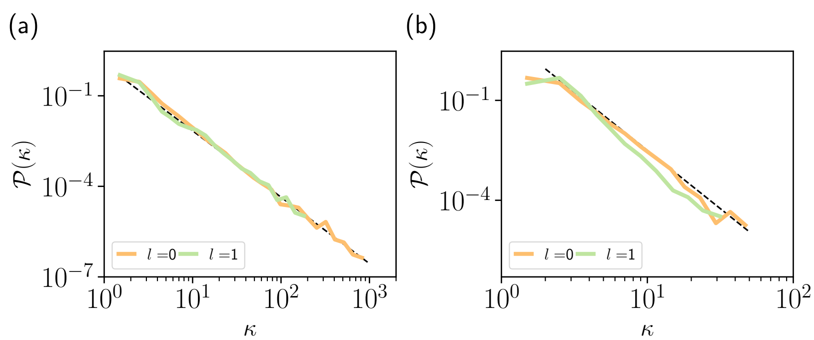

FIG. 12: Degree distribution after a single-LRG transformations for scale-free networks of the general form $P(\kappa) \propto \kappa^{-\gamma}$, generated through the configurational model with: (a) $\gamma = 2.2$ and $N_0 = 12022$. The RG step is performed using $\tau = 0.025$ which leads to $N_1 = 6805$. (b) $\gamma = 3.5$ and $N_0 = 13944$. The LRG step is performed using $\tau = 0.8$ which leads to $N_1 = 10138$. Note that the degree distribution, $P(\kappa)$, remains invariant under RG transformations.

5. REAL-WORLD NETWORKS

For the sake of completeness we present here a full analysis of different real scale-free metabolic networks of different organisms from a collection of real network data from the ICON at https://icon.colorado.edu/ as well as the KONECT at https://west.uni-koblenz.de/konect. Figure 14 shows the evolution of the entropy parameter as a function of the resolution scale, τ, evidencing the complex mesoscopic structure of these networks. We also show the Laplacian spectrum, which evidence that only the *M.Musculus* and *E.Coli* network present clean scale-invariant properties, i.e. they present constant or power-law regimes in the eigenvalues spectrum. In any case, we can produce replicas for all the networks, with consistent degree distribution for the downscaled networks.

We also analyze a set of networks of different areas, following [4], namely: (i) the Internet at the Autonomous Sustems level [4], (ii) the Airports network (from KONECT database), (iii) the human HI-II-14 interactome [5], (iv) the network of adjacency between words in Darwin's book "The Origin of Species" [6], and (v) the Drosophila connectome network [7].

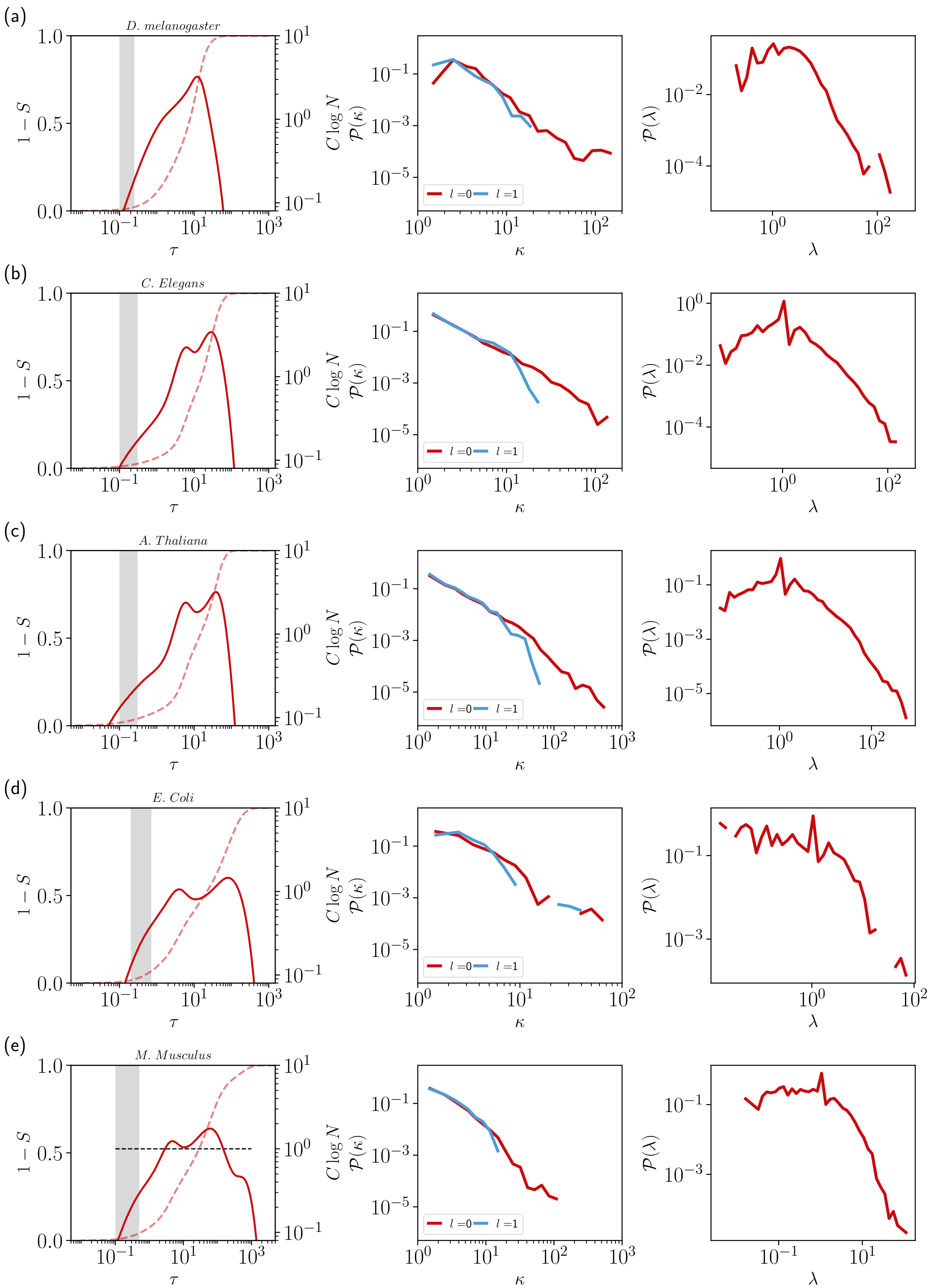

FIG. 13: **(Left column)** Entropy parameter (dashed lines, $(1-S)$), and specific heat (solid lines, C), versus the temporal resolution parameter of the network, τ for different species (see title). **(Central column)** LRG transformation for the different considered species using: (a) $\tau = 0.2$ (b) $\tau = 0.2$ (c) $\tau = 0.1$ (d) $\tau = 0.5$ (e) $\tau = 0.3$ **(Right column)** Spectral probability distribution of the Laplacian matrix for the different species.

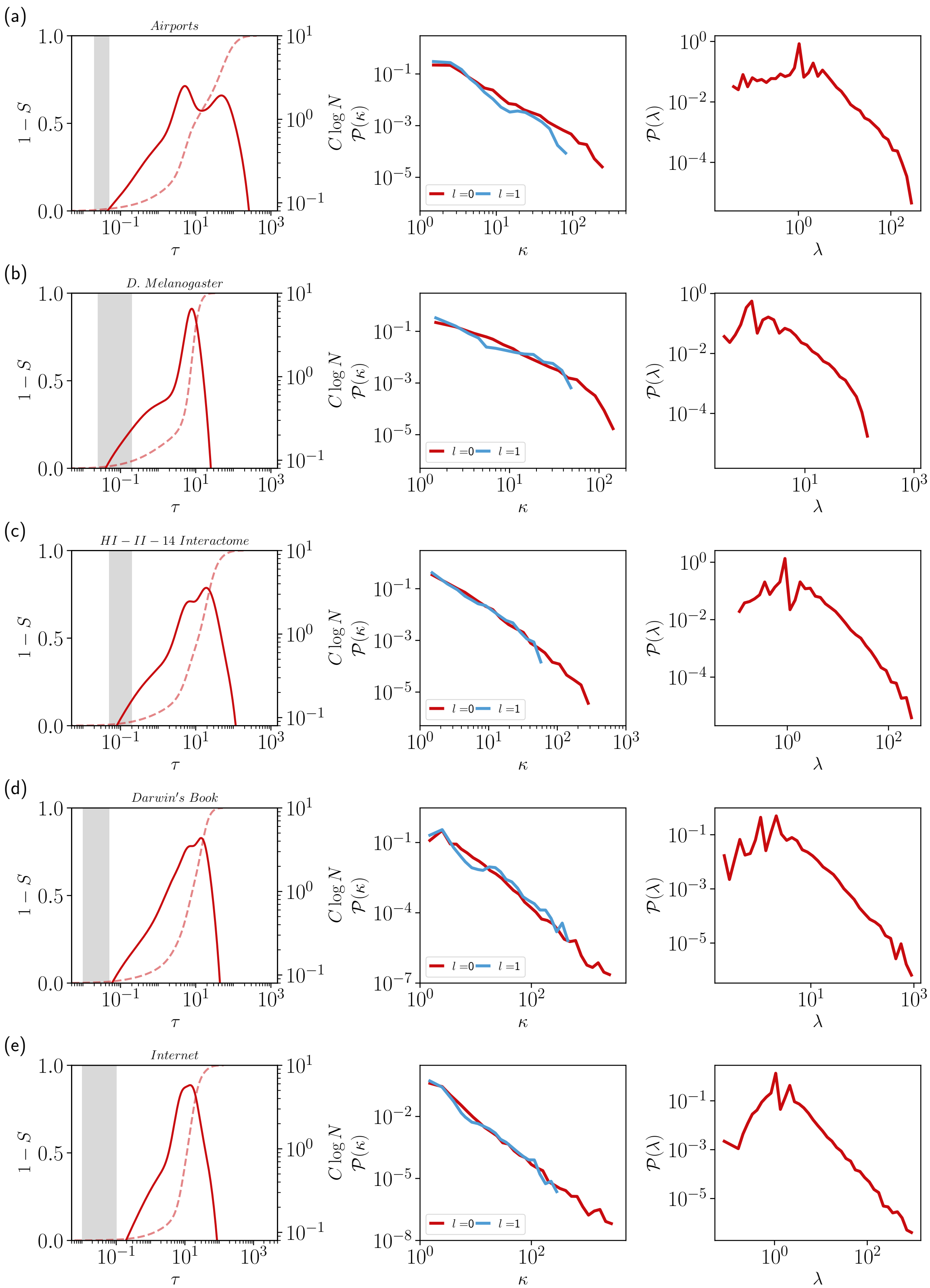

FIG. 14: **(Left column)** Entropy parameter (dashed lines, $(1-S)$), and specific heat (solid lines, C), versus the temporal resolution parameter of the network, τ for different networks (see title). **(Central column)** LRG transformation for the different considered species using: (a) $\tau = 0.045$ (b) $\tau = 0.07$ (c) $\tau = 0.07$ (d) $\tau = 0.01$ (e) $\tau = 0.02$ **(Right column)** Spectral probability distribution of the Laplacian matrix for the different networks.

REFERENCES

[1] K. G. Wilson and J. Kogut, The renormalization group and the ϵ expansion, Phys. Rep. **12**, 75 (1974).
[2] K. Christensen and N. R. Moloney, *Complexity and criticality*, Vol. 1 (World Scientific Publishing Company, 2005).
[3] D. J. Watts and S. H. Strogatz, Collective dynamics of 'small-world'networks, Nature **393**, 440 (1998).
[4] G. García-Pérez, M. Boguñá, and M. Á. Serrano, Multiscale unfolding of real networks by geometric renormalization, Nat. Phys. **14**, 583 (2018).
[5] T. Rolland, M. Taşan, B. Charloteaux, S. J. Pevzner, Q. Zhong, N. Sahni, S. Yi, I. Lemmens, C. Fontanillo, R. Mosca, *et al.*, A proteome-scale map of the human interactome network, Cell **159**, 1212 (2014).
[6] R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, S. Shen-Orr, I. Ayzenshtat, M. Sheffer, and U. Alon, Superfamilies of evolved and designed networks, Science **303**, 1538 (2004).
[7] S.-y. Takemura, A. Bharioke, Z. Lu, A. Nern, S. Vitaladevuni, P. K. Rivlin, W. T. Katz, D. J. Olbris, S. M. Plaza, P. Winston, *et al.*, A visual motion detection circuit suggested by drosophila connectomics, Nature **500**, 175 (2013).