




## Renormalization of interacting random graph models

Alessio Catanzaro <sup>1,2,3,\*</sup>, Diego Garlaschelli <sup>1,2,†</sup> and Subodh P. Patil <sup>1,‡</sup>

<sup>1</sup>*Instituut-Lorentz for Theoretical Physics, Leiden University, 2333 CA Leiden, The Netherlands*

<sup>2</sup>*IMT School for Advanced Studies, Piazza San Francesco 19, 55100 Lucca, Italy*

<sup>3</sup>*Dipartimento di Fisica e Chimica, Università di Palermo, Viale delle Scienze, Ed. 17, 90133 Palermo, Italy*



(Received 30 October 2025; accepted 5 February 2026; published 27 February 2026)

Random graphs offer a useful mathematical representation of a variety of real-world complex networks. Exponential random graphs, for example, are particularly suited towards generating random graphs constrained to have specified statistical moments. In this investigation we elaborate on a generalization of the former where link probabilities are conditioned on the appearance of other links, corresponding to the introduction of interactions in an effective generalized statistical mechanical formalism. When restricted to the simplest nontrivial case of pairwise interactions, one can derive a closed form renormalization group transformation for maximum coordination number two on the corresponding line graph. Higher coordination numbers do not admit exact closed form renormalization group transformations, a feature that paraphrases the usual absence of exact transformations in two or more dimensional lattice systems. We introduce disorder and study the induced renormalization group flow on its probability assignments, highlighting its formal equivalence to time reversed anisotropic drift diffusion on the statistical manifold associated with the effective Hamiltonian. We discuss the implications of our findings, stressing the long wavelength irrelevance of certain classes of pairwise conditioning on random graphs, and conclude with possible applications. These include modeling the scaling behavior of preferential effects on social networks, opinion dynamics, and reinforcement effects on neural networks, as well as how our findings offer a systematic framework to deal with data limitations in inference and reconstruction problems.

DOI: [10.1103/34n8-pw8x](https://doi.org/10.1103/34n8-pw8x)

### I. PRELIMINARY REMARKS

One of the most utilized approaches in modeling and studying the properties of random networks is the so-called Exponential Random Graph (ERG) framework [1–8]. Its formulation builds on the realization that in a precise sense, the formalism of statistical physics is ontologically equivalent to statistical inference, and can be formalized more broadly in terms of a *generalized statistical mechanics* [9,10]. The workhorse of this generalization lies in viewing the Gibbs entropy formula as a use case of Shannon information entropy [11], which facilitates applications beyond physical systems towards any system where observable quantities are drawn from a random ensemble. The implications of this are as immediate as they are profound. One can define a probability measure that reproduces the observed statistics of any given random system and associate it with the “Hamiltonian” of a generalized statistical mechanical ensemble. This is done by demanding that the measure replicate all observed moments in a manner that maximizes its Shannon information, and

therefore being minimally presumptive of any other priors.<sup>1</sup> One recovers the usual Boltzmann weight when applied to thermodynamic systems, and in the context of random graphs, the result is the formalism of exponential random graphs [1,3].

Given that the properties of a random graph can be derived from its adjacency matrix, the ERG formalism associates a Hamiltonian to any random graph ensemble such that the probability of realizing any given draw for the adjacency matrix  $A_{ij}$  is given by

$$P(\mathbf{A}) \equiv \frac{e^{-H(\mathbf{A})}}{Z}; Z := \sum_{\{A_{ij}\}} e^{-H(\mathbf{A})}, \quad (1)$$

where  $Z$  is referred to as the partition function, and where the Hamiltonian is in general a multilinear function of the elements of the adjacency matrix:

$$H(\mathbf{A}) = -h^{(0)} - \sum_{ij} h_{ij}^{(1)} A_{ij} - \sum_{ij \neq kl} h_{ijkl}^{(2)} A_{ij} A_{kl} \\ - \sum_{\substack{ij \neq kl, \\ mnkl \neq mn}} h_{ijklmn}^{(3)} A_{ij} A_{kl} A_{mn} - \dots \quad (2)$$

\*Contact author: [catanzaro@lorentz.leidenuniv.nl](mailto:catanzaro@lorentz.leidenuniv.nl)

†Contact author: [garlaschelli@lorentz.leidenuniv.it](mailto:garlaschelli@lorentz.leidenuniv.it)

‡Contact author: [patil@lorentz.leidenuniv.nl](mailto:patil@lorentz.leidenuniv.nl)

*Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.*

<sup>1</sup>In his seminal work, E. T. Jaynes credits Shannon as supplying the missing ingredient required to actualize Laplace’s Principle of Insufficient Reason for assigning probabilities to account for all possible draws from a random ensemble constrained to reproduce a fixed set of moments, and possibly supplemented with additional structural priors [9,10].

The form of the Hamiltonian parameterizes generic interactions among links related to various graph observables, such as the number and distribution of triangles or wedges. Its couplings up to some order can be fixed by demanding expectation values associated with the probability assignment [defined by Eq. (1)] for a sufficient number of observables  $\{\mathcal{O}_a\}$  be constrained to assigned values:

$$\langle \mathcal{O}_a \rangle = \sum_{\mathbf{A}} P(\mathbf{A}) \mathcal{O}_a(\mathbf{A}), \quad (3)$$

subject to maximization of the Gibbs-Shannon entropy

$$S = - \sum_{\mathbf{A}} P(\mathbf{A}) \log P(\mathbf{A}). \quad (4)$$

In practice, one only ever has access to a finite number of observables, themselves subject to sample variance and potential estimator bias. This will not be an issue for the intents and purposes of a predictive framework for reasons that we will elaborate upon shortly.

The  $\{A_{ij}\}$  sum over the set  $\{0, 1\}$  for unweighted graphs, and otherwise integrate over arbitrary values for weighted graphs. In what follows, we restrict ourselves to the unweighted case. The simplest example of the latter is furnished by the Erdős-Rényi random graph, for which all couplings in Eq. (2) vanish except for  $h_{ij}^{(1)} = h_{ji}^{(1)} = \log[p/(1-p)]$  for  $i \neq j$  (see, e.g., [12]). A disordered variant of the Erdős-Rényi graph is obtained by allowing the  $h_{ij}^{(1)}$  to vary linkwise, so that

$$h_{ij}^{(1)} = \log \left( \frac{p_{ij}}{1 - p_{ij}} \right). \quad (5)$$

The exponential random graph approach has been successful in a number of contexts, including network inference problems, and is considered as the benchmark for null models of random graphs. However, this approach comes up against a number of difficulties in wider application. One is that while straightforwardly solvable for the separable case that is, when

$$Z \equiv Z_1 Z_2 \cdots Z_N, \quad \mathcal{N} := N(N-1)/2, \quad (6)$$

where each individual factor corresponds to the partition function for each possible individual link<sup>2</sup>—when correlations and higher order conditionings are introduced, analytic solutions are no longer so straightforwardly obtainable. Moreover, in phenomenological approaches where one attempts to reconstruct graph Hamiltonians that reproduce various features inferred from data, one is not only hindered by sample variance, but also by data that is limited to network realizations of specific sizes. Certain features corresponding to graph generative processes, i.e., involving the dynamics of graph formation and growth, might elude an inference approach based on the standard formulation of the exponential random graph framework.

Perhaps the best known mechanism of graph formation in network science is furnished by the Barabasi-Albert network model, in which a “rich-get-richer” process drives graph

evolution towards particular asymptotic scaling properties [13,14]. The conceptual utility of these models derives from their simplicity, which also inevitably limits their applicability in modeling complex systems where multiple processes can shape connections. To account for the latter, one might consider introducing dynamical features that depend on various building blocks of the graph. One such class of models are often referred to as fitness (or hidden variable) models [8,15–19] where one assigns node-specific variables, with graph interactions dependent on these quantities. The assignment and distribution of the fitness variables end up determining the macroscopic properties of the graph and its associated network structures, demanding answers to questions such as which features are relevant at the largest scales, which are irrelevant, and how the presence of disorder in their assignment is reflected in the statistics of macroscopic observables.

### A. The (effective) generalized statistical mechanics of random graphs

In what follows, we develop the means to address some of the aforementioned issues by taking the generalized statistical mechanical framework in which the ERG formalism originates to its natural completion, from both a formal and an inferential point of view. We do this by embedding it within the epistemic framework furnished by the paradigm of *effective theories* [20–23]. The idea is simple to state—one simply views the Hamiltonian in Eq. (2) as a moment generating functional that has been operator Taylor expanded without loss of generality. Each specification of the coefficients in its expansion parameterizes a specific point in theory space, where the only prior restrictions on the coefficients stem from the symmetries and structural constraints that define the system.

Although at first glance it might seem that one has to specify a tower of coefficients in order to have a predictive or retrodictive framework, this is operationally not the case given that one only ever has access to limited resolution and sample sizes in observations. Consequently, only a finite number of coefficients have to be fixed (which is to work up to a certain order in the operator expansion) in order to have a predictive framework for all subsequent observations to the stipulated accuracy.<sup>3</sup> The reason for this can be understood both from a renormalization group (RG) and a statistical inference point of view. When viewed as an *effective* Hamiltonian, the terms in the expansion of Eq. (2) represent an ordered series in terms of relevance: *Power counting relevance* from the perspective of a renormalization group analysis, and *parameter relevance* from the perspective of statistical inference. The former can be viewed as a special case of the latter through the portal of Fisher information [24,25], in a manner that is straightforward to articulate.

<sup>3</sup>This presumes that one works in an operator basis such that the analog of Eq. (2) represents an expansion whose leading order terms correspond to an exactly solvable system (however non-linear or strongly correlated), where the higher order terms parametrize collective modes that perturb away from this solution in a convergent series.

<sup>2</sup>For example, the partition function for a graph with  $N$  nodes factorizes as per Eq. (6) when one truncates the expansion beyond terms linear in  $A_{ij}$  in Eq. (2).

Sequentially higher order terms in the effective Hamiltonian [Eq. (2)] have an increasingly smaller effect on observable quantities, caveating footnote 3. That is, attempting to fix the parameters of the Hamiltonian with inevitably limited data will result in a likelihood function that will exhibit progressively weaker parametric dependence on higher order couplings in a pronounced hierarchy. Leading order couplings will correspond to the largest eigenvalues of the associated Fisher-Rao information metric [26–28] (reviewed in Sec. IC), or have the greatest parameter relevance. Higher order couplings will correspond to successively smaller eigenvalues, or be increasingly *irrelevant*. This is perhaps best illustrated through an example that ought to be familiar to many readers.

For continuum physical systems, the renormalization group notion of power counting relevance translates directly into a derivative expansion. By expanding in a series of derivative operators that is consistent with the symmetries of the system (e.g., rotational symmetry), one can effectively model arbitrary dynamical processes up to some resolution. Since derivatives have physical dimension, successively higher derivative terms require compensating dimensional factors in order to book keep the expansion. For example, one can consider modeling an isotropic diffusion process via the following effective equations:

$$\frac{d\rho}{dt} = \sigma\rho + D_2\nabla^2\rho + D_4\ell^2\nabla^4\rho + D_6\ell^4\nabla^6\rho + \dots, \quad (7)$$

where as written, all the higher order coefficients  $D_4, D_6$ , etc. have the same physical dimension as the diffusion coefficient  $D_2$ . At spatial resolutions  $L \gg \ell$ , so that the Fourier modes associated with these scales have wave numbers that satisfy  $k_L/k_\ell \ll 1$ , we see that Eq. (7) represents a convergent operator Taylor expansion where successive terms are subleading in even powers of  $k_L/k_\ell$ . We'll return to the physical significance of the so far arbitrary scale  $\ell$  shortly.

In order for the quartic operator to have a comparable effect to the leading diffusion term, the  $D_4$  coefficient would have to be dialed up to  $D_4 \sim D_2k_\ell^2/k_L^2 \gg D_2$ . Similarly, one would need  $D_6 \sim D_4k_\ell^2/k_L^2 \gg D_4$  in order for the sextic operator to compete with the quartic, and so on. Conversely, unless the  $D_4$  coefficient starts to approach  $D_4 \approx D_2k_\ell^2/k_L^2$ , the quartic operator will have negligible effects on long wavelength observables in comparison to the quadratic diffusion operator. In fact, there is a wide range within the parameter interval defined by the condition  $|D_4| \ll |D_2k_\ell^2/k_L^2|$  where  $D_4$  can vary and have parametrically negligible effects on long wavelength observables relative to the leading diffusion term. That is, the likelihood function constructed from repeated observation will have a much weaker dependence on  $D_4$  than  $D_2$ , and will hence be associated with a smaller eigenvalue of the associated Fisher information metric.

In the context of a renormalization group analysis, when one successively coarse grains over small scales and reexpands the effective equations of motion as in Eq. (7) after each coarse graining, one will find that the value of  $D_2$  remains roughly the same, whereas  $D_4$  decreases. For this reason,  $D_2$  is classified as *marginally relevant*, and  $D_4$  is classified as *irrelevant*, with both notions exactly corresponding to their parameter relevance as determined by their associated Fisher

metric eigenvalues [24,25,29].<sup>4</sup> All that remains is to clarify the nature of the scale  $\ell$  in Eq. (7): In a precise operational sense, it is the resolution scale of our effective description. Were one to average over all fluctuations with wavelengths smaller than the length scale  $\ell$ , the resulting equations of motion for *all* modes with wavelengths longer than  $\ell$  would be reproduced by an operator expansion corresponding to Eq. (7), parametrically weighted by factors of  $\ell^2$  with order unity numerical prefactors. If one views diffusion processes as having coarse grained an underlying kinetic theory description [32,33], one can explicitly demonstrate this via an influence kernel parametrization of interparticle forces (see, e.g., [34] in the context of broken symmetries).

The realization that one can parametrize one's ignorance of small scale processes via an operator expansion and still end up with a tractable and predictive framework for macroscopic observables is what freed particle physicists from having to work within the confines of *renormalizable* field theories,<sup>5</sup> and ushered in the era of *effective field theories*. The latter represents the state of the art in particle phenomenology, and it features an underlying logic that transcends its original application towards continuum systems. One simply writes down all possible operators that could appear in the moment generation functional not forbidden by symmetries or other kinematic or structural constraints (i.e. the effective action in the context of particle physics, or the effective Hamiltonian in the context of a generalized statistical mechanical system), and classifies them according to their parameter relevance.

## B. Universality classes of graph Hamiltonians

It is from the vantage of effective theories that renormalization group analyses take on a categorical importance of their own. Not only do RG methods give us the tools to determine parameter relevance, their complete implementation allows us to unambiguously determine which quantities are legitimate observables at long wavelengths, and what may be artifacts of various approximation schemes. An important corollary of this classification is that all models flow under renormalization towards a handful of equivalence classes at large scales, and given the special designation of *universality classes*. Which universality class one flows to is determined entirely by the invariances or symmetries of the system, along with how these symmetries are broken as one changes scale.

In what follows, we attempt the first steps in understanding the behavior of a generalized statistical mechanical

<sup>4</sup>The classification of operators into categories of relevance, marginal relevance, and irrelevance under the action of the renormalization group is owed to Kadanoff and Wegner [30,31].

<sup>5</sup>In particle physics effective field theory, the adjective *renormalizable* is reserved for those theories which form a closed subspace under renormalization group flow in the space of all theories, parameterized via some operator basis. Effective theories in general need not be renormalizable by this definition, but nevertheless remain predictive as discussed, as only a finite subset of operators corresponding to the leading order terms in any derivative expansion remain relevant or marginal as one flows to the infrared (see, e.g., [20,21]).

parametrization of unweighted random graph models under renormalization group flow. Our starting point is the operator expansion (2), whose most immediate consequence is that the probability for any link to appear will generically be conditioned on the appearance of other links. As a consequence, the partition function for the graph will not be factorizable into a product of independent partition functions for each link, that is,

$$Z \neq Z_1 Z_2 \cdots Z_{\mathcal{N}}, \quad (8)$$

where once again  $\mathcal{N}$  counts the number of links in a graph with  $N$  nodes, and we omit the possibility of loops (edges that begin and end on the same node) in the discussion to follow. The simplest nontrivial case of this is when one restricts to bilinear interactions, which can be written without loss of generality as

$$H(\mathbf{A}) = - \sum_{i \leq j} \varepsilon_{ij} A_{ij} - \frac{1}{2} \sum_{\substack{i \leq j, k \leq l \\ (i,j) \neq (k,l)}} \beta_{ijkl} A_{ij} A_{kl}, \quad (9)$$

where the pairs  $(i, j)$  and  $(k, l)$  denote the links connecting the respectively indexed nodes. Tripartite and higher order conditioning can similarly be introduced via cubic and higher order interactions via a generating function formalism, which we review in Appendix A. If this higher order conditioning is weak enough, one can avail of the diagrammatic methods of statistical field theory and work in a perturbative series in the corresponding small cubic and higher couplings (see, e.g., [35]). Although we won't pursue this possibility in the present investigation, a more complete study of parameter relevance in random graph models would necessitate it. Instead, we focus only on bilinear interactions in what follows and defer a more comprehensive study to a future investigation. A rich structure already emerges within this class, where moreover, the existence of a closed subspace under exact RG transformations can be demonstrated.

Truncating the effective Hamiltonian to bilinear interactions might seem like an academic oversimplification at first glance. However, as elaborated in the next section, Eq. (2) can also be viewed as the Hamiltonian of an equivalent disordered spin system with couplings of arbitrary order (i.e., in terms of degrees of nearest neighbors), whose effective dimensionality is determined by the number of nearest neighbors each site has. We refer to this quantity as the coordination number of a link, which is encoded in all the  $h_{ijkl}^{(2)}$  terms being different from zero. Were one to presume a hierarchy for these terms over higher order couplings, the result of successive RG transformations for a large class of bare effective Hamiltonians would be to flow towards the form (9), which therefore represents a universality class of equivalent random graph models, and therefore an important object of study on its own terms. Moreover, from a practical and application intensive point of view, Eq. (9) can also effectively model temporal Markov network processes, where RG transformations would correspond to temporal coarse graining. We elaborate on this in more detail in an accompanying study [36], and commence our formal study of random graph renormalization in the following sections.

### C. Outline of the paper

The goal of this paper is to set up a general framework through which one might quantify the scaling of interactions in random graph models. That is, to determine the relevance, or irrelevance of the conditionings of different graph variables on each other. Although this is highly nontrivial problem in full generality, we focus in what follows on a restricted class of transformations for which closed form renormalization group transformations are available, which we argue to represent a universality class of graph Hamiltonians.

In the sections that follow, we first set up the formalism of random graph renormalization in the line graph representation in Sec. II, elaborating on the existence of closed form RG transformations for maximum coordination number two pairwise interactions. We then consider how these interactions renormalize when the couplings are homogeneously assigned in Sec. III, and generalize to the disordered setting in Sec. IV, where we immediately find that the problem can be recast in terms of determining the the induced flow on the disorder probability assignments. We understand this flow to be akin to time reversed drift diffusion, as detailed in Appendix C, with additional details relevant to our discussion collected in Appendixes A and B. We conclude our investigation with a discussion of the implications of our findings, the most pertinent of which is that within the universality class of bilinear graph Hamiltonians, all models are *infrared* trivial, meaning that all pairwise conditions flow to irrelevance at large enough scales. We translate the ramifications of this for a number of applications, and close with a discussion of open questions and future directions.

## II. LINE GRAPHS, SPIN CHAINS, AND CLOSED RG TRANSFORMATIONS

The graph Hamiltonian (9) can be reexpressed as the Hamiltonian of a disordered spin chain via mapping to the line graph:

$$H(\mathbf{A}) = - \sum_a \varepsilon_a A_a - \frac{1}{2} \sum_{a \leq b} \beta_{ab} A_a A_b, \quad (10)$$

where distinct pairs of indices  $(i, j)$  are mapped to a single index  $a$ , so that edges have been mapped to nodes, and where  $1 \leq a \leq \mathcal{N}$ . We stress that *a priori* there is no natural notion of nearest neighbor on a random graph at the level of the effective Hamiltonian Eq. (10). Metric notions of nearest neighbors and local neighborhoods—key to the usual implementation of the renormalization group—are only something that can be defined for a specific instance of a random graph.<sup>6</sup> Moreover, the couplings  $\varepsilon_a$  and  $\beta_{ab}$  are *a priori* completely arbitrary and need not even correspond to “local” interactions. That is, without additional restrictions, the probability of any given link appearing can in principle be conditioned on an arbitrary number of other links.

Nevertheless, a class of bilinear interactions can be distinguished by the fact that they remain closed under an exact

<sup>6</sup>For a recent review on progress and open issues regarding network renormalization, see [37].

RG transformation that corresponds to decimation on the line graph. Specifically, we consider interactions of the form

$$\beta_{ab} = g_a \delta_{a,b-1} + g_b \delta_{a,b+1}, \quad (11)$$

so that

$$H(\mathbf{A}) = - \sum_{a=1}^{\mathcal{N}} [\varepsilon_a + g_a A_{a+1}] A_a, \quad (12)$$

which formally looks like the Hamiltonian of a paramagnetic spin glass with nearest neighbor interactions, in spite of the absence of such notions on a random graph. The coefficients  $\varepsilon_a$  and  $g_a$  themselves could be drawn from independent ensembles, corresponding to arbitrary “local” disorder.

As discussed in Appendix A, the formal process of integrating out one or more of the edge degrees of freedom in the context of Eq. (12) is equivalent to statistical marginalization, and effects a decimation renormalization group transformation on the line graph. It should be stressed at the outset that there is no a priori compulsion to have imposed this transformation uniformly, and one can consider determining the parameters of the effective Hamiltonian obtained by marginalizing over a single link in isolation (in [36] we take an approach to network renormalization that is premised on iterating this procedure). The result of having decimated the  $n$ th link is a locally coarse-grained line graph with an effective Hamiltonian with the following renormalized couplings:<sup>7</sup>

$$\begin{aligned} \tilde{g}_{n-1} &= \log \left[ \frac{(1 + e^{g_{n-1} + g_n + \varepsilon_n})(1 + e^{\varepsilon_n})}{(1 + e^{g_n + \varepsilon_n})(1 + e^{g_{n-1} + \varepsilon_n})} \right], \\ \tilde{\varepsilon}_{n-1} &= \varepsilon_{n-1} + \log \left[ \frac{(1 + e^{g_{n-1} + \varepsilon_n})}{(1 + e^{\varepsilon_n})} \right], \\ \tilde{\varepsilon}_{n+1} &= \varepsilon_{n+1} + \log \left[ \frac{(1 + e^{g_n + \varepsilon_n})}{(1 + e^{\varepsilon_n})} \right], \end{aligned} \quad (13)$$

with all others left unchanged. On the other hand, the effect of uniform decimation, where every alternate node on the line graph is successively marginalized over is to effect the following transformations:

$$\begin{aligned} \tilde{g}_n &= \log \left[ \frac{(1 + e^{g_n + g_{n+1} + \varepsilon_{n+1}})(1 + e^{\varepsilon_{n+1}})}{(1 + e^{g_n + \varepsilon_{n+1}})(1 + e^{g_{n+1} + \varepsilon_{n+1}})} \right], \\ \tilde{\varepsilon}_n &= \varepsilon_n + \log \left[ \frac{(1 + e^{g_n + \varepsilon_{n+1}})(1 + e^{g_{n-1} + \varepsilon_{n-1}})}{(1 + e^{\varepsilon_{n+1}})(1 + e^{\varepsilon_{n-1}})} \right], \end{aligned} \quad (14)$$

where we’ve relabeled the indices to run over the integers again after each successive uniform decimation. A simple derivation of these equations can be found in the context of the Ising model with sparse disorder (see, e.g., [38]) or in our accompanying paper [36].

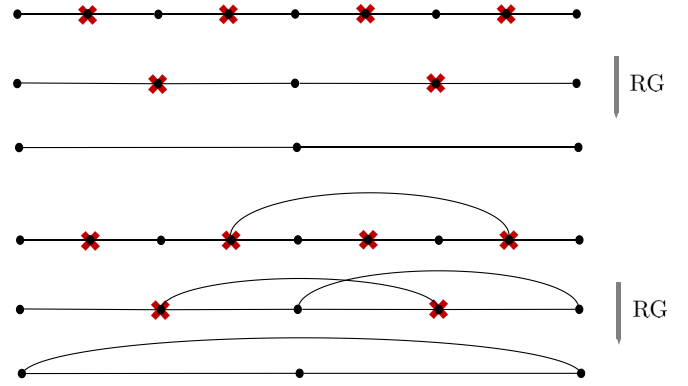


FIG. 1. Schematic rendering of maximum coordination two couplings (top) and a single bare coordination number three coupling (bottom) under RG transformations. The red crosses indicate the sites that are decimated at each RG step. More generally, successive RG steps increase the maximum coordination number, and flow towards all-to-all couplings (bottommost diagram).

The closure of these RG transformations—meaning that each RG step returns an effective Hamiltonian of the same operator form up to rescaled couplings and an overall rescaling of the partition function—results from the fact that Eq. (11) represents bilinear interactions of maximum coordination number two in the line graph representation. Consequently, the associated Hamiltonian Eq. (12) maps onto that of a one-dimensional spin glass with localized disorder and an inhomogeneous external magnetic field, inheriting the closure and exact implementability of renormalization group transformations specific to one dimensional systems. This is no longer the case the instance we add even a single interaction with coordination number three or higher, with successive RG transformations driving the system towards highly heterogeneous all-to-all couplings (the homogeneous limit of which would be a Curie-Weiss-like model).

Specifically, consider replacing some of the terms in Eq. (12) with the coordination number three interaction

$$H \supset [\varepsilon_m + g_m A_{m+1} + g'_m A_{m+k}] A_m \quad (15)$$

for some fixed  $m$  and some integer  $k$ . In the context of the analog lattice system where one can avail of notions of metric distance,  $k = 2$  would correspond to a next to next to nearest neighbor interaction. In the present context, we leave  $k$  arbitrary, and view it as corresponding to conditioning the  $m$ th and  $m + k$ -th links on each other. Under repeated decimation, the maximum coordination number on the line graph would successively increase in such a manner such that when one considers randomly distributing higher coordination interactions across the graph, the effective Hamiltonian would flow towards all-to-all couplings, schematically depicted in Fig. 1.

#### A. Maximum coordination number two-disordered spin chain

The fact that maximum coordination number two and all-to-all couplings represent the two universality classes that admit closed form renormalization group transformations is a transcription of the closure of RG transformations in one dimensional lattice spin systems to the context of random

<sup>7</sup>Up to an overall rescaling of the partition function that is equivalent to a localized additive free energy contribution given by  $F_n = \log[1 + \varepsilon_n]$ , corresponding to the individual contribution of the marginalized over link in Eq. (A12).

graph models. We can also make a straightforward identification for higher coordination numbers by noting that the lack of closure for RG transformations in general higher dimensional lattice systems necessitate approximation methods to force closure, such as bond shifting [39,40], variational, or mean field methods (see, e.g., [41] for an overview).<sup>8</sup> In the absence of these approximations, summing over the contributions of short distance degrees of freedom generates higher order and beyond nearest neighbor interactions with each successive RG transformation. In spite of this, the availability of a predictive effective description via an effective Hamiltonian stems from the fact that these higher order and higher point interactions correspond to increasingly *irrelevant* operators as one flows to the infrared, and is the basis of the derivative expansion in continuum descriptions (cf. footnote 5). Hence, RG transformations generate trajectories in theory space where all possible interactions are generated, albeit arranged in a hierarchy of relevance for long wavelength observables. The limit of this flow for an initial line graph with bilinear interactions of arbitrary coordination number is heterogeneous all-to-all couplings, unless the bare couplings form a closed subspace to begin with, as is the case for maximum coordination number two.

### III. HOMOGENEOUS RG FLOW

A simple but informative case to get our bearings with is the setting where each of the  $\{g_a, \varepsilon_a\}$  couplings in Eq. (12) are, respectively, fixed to be the same, and we successively decimate, or marginalize over alternate links on the line graph (see for a thorough derivation [36]). In this case, Eqs. (14) reduce to the following system of equations:

$$\begin{aligned}\tilde{g} &= \log \left[ \frac{(1 + e^{2g+\varepsilon})(1 + e^\varepsilon)}{(1 + e^{g+\varepsilon})^2} \right], \\ \tilde{\varepsilon} &= \varepsilon + 2 \log \left[ \frac{(1 + e^{g+\varepsilon})}{(1 + e^\varepsilon)} \right],\end{aligned}\quad (16)$$

where tildes denote renormalized quantities, and where the discrete generalization of the beta functions, defined as  $Dg := \tilde{g} - g \equiv \beta_g(g, \varepsilon)$ , and  $D\varepsilon := \tilde{\varepsilon} - \varepsilon \equiv \beta_\varepsilon(g, \varepsilon)$ , are given by

$$\begin{aligned}\beta_g(g, \varepsilon) &= \log \left[ \frac{(1 + e^{2g+\varepsilon})(1 + e^\varepsilon)}{e^g(1 + e^{g+\varepsilon})^2} \right], \\ \beta_\varepsilon(g, \varepsilon) &= 2 \log \left[ \frac{(1 + e^{g+\varepsilon})}{(1 + e^\varepsilon)} \right].\end{aligned}\quad (17)$$

The RG flow corresponding to these beta functions is depicted in Fig. 2, where we see that successive transformations drive one towards the  $g \equiv 0$  fixed line, so the links are no longer conditioned on each other and we recover an Erdős-Rényi random graph at long wavelengths.

The various features of the flow diagram lend themselves to straightforward interpretation. We first note that the part of the diagram below the  $g = 0$  axis is only meaningful for a

single RG step. The reason is that implicit in Eq. (16) is the fact that were we to begin with a negative value for the bare  $g$ , a single decimation iteration will return a positive value for  $g$ . Although the structure is far richer when we turn our attention to inhomogeneous and disordered interactions, the latter feature of considering only homogeneous interactions can be readily understood from the fact that on the analog spin system, negative values of  $g$  corresponds to uniform antiferromagnetic couplings. When one considers the effect of marginalizing over two adjacent antiferromagnetic couplings, the result is a effective ferromagnetic coupling (i.e.,  $\tilde{g} > 0$ ) that maintains its ferromagnetic nature under further RG iterations.

The separatrix that is apparent along the line  $\varepsilon = -g$  in the upper plane, along with the fixed line at  $g = 0$  can also be readily understood by referring to the analog spin system. Recalling Eq. (12) and making the transformation to the spin variables

$$A_a = \frac{1 + 2\sigma_a}{2}, \quad \sigma_a = \pm \frac{1}{2}, \quad (18)$$

the result will be a Hamiltonian of the form

$$H(\sigma) = - \sum_{a=1}^{\mathcal{N}} g_a \sigma_a \sigma_{a+1} - \sum_{a=1}^{\mathcal{N}} \sigma_a \left( \varepsilon_a + \frac{g_a + g_{a+1}}{2} \right), \quad (19)$$

where we've neglected to include additive contributions to the free energy. When restricted to the case  $\varepsilon_a \equiv \varepsilon$  and  $g_a \equiv g$ , we see that this Hamiltonian maps on to that of a one dimensional Ising model in the presence of an external magnetic field  $h$  with the identification  $h := \varepsilon + g$ . Given that decimation in zero external field does not generate any one point terms in the renormalized Hamiltonian, the locus represented by  $h \equiv 0$  or  $\varepsilon = -g$  for represents an (unstable) fixed line in the upper  $g$  plane, and an attractor in the lower  $g$  plane.<sup>9</sup> Furthermore, the fact that the overall flow is driven to a fixed line at  $g = 0$  follows from fact that in one dimension, RG flow drives one towards the infinite temperature fixed point under the identification  $g \equiv J/(k_B T)$ , where  $J$  denotes nearest neighbor couplings on the analog spin chain.

Finally, the shearing structure on either side of this separatrix also admits a straightforward interpretation. When one positively or negatively conditions the probability of a given link to appear on the appearance of another link, it effectively increases, or respectively decreases the probability for that link to appear for homogeneous probability assignments, provided that the conditioning is strong enough. Therefore coarse graining will steadily drive the effective probability up or down depending on whether one is in the upper or lower  $g$  plane until one is driven to the (analog infinite temperature) fixed line at  $g \equiv 0$ .

A far richer structure appears when one allows for the couplings to vary by site on the line graph. Although it would be meaningful to consider the effects of an arbitrary initial specification of the full set of couplings  $\{\varepsilon_a, g_a\}$ , a more general framing with an eye towards further applications would

<sup>8</sup>Caveating the availability of exact solutions in the specific case of  $d = 2$  in the absence of disorder [42,43].

<sup>9</sup>This is made explicit by considering the beta function for the linear combination  $h = \varepsilon + g$ , which is given by  $\beta_h = \log\{(e^{g+\varepsilon} + e^{-g})/(1 + e^\varepsilon)\}$ , from which we see that  $\beta_h \equiv 0$  on the locus  $\varepsilon = -g$ .

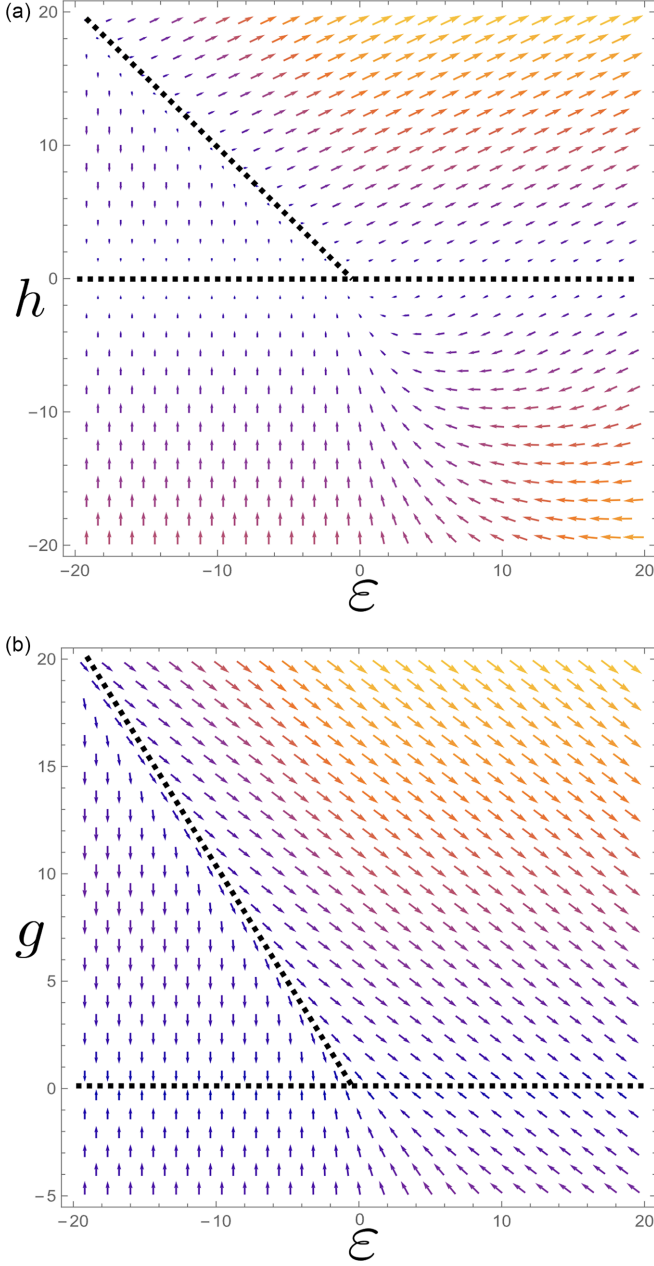


FIG. 2. RG flow for homogeneous coordination number two interactions towards an Erdős-Rényi fixed line, where arrows indicate direction of flow and color determine the its intensity, lighter being stronger. (a) A shearing behavior for positively correlated conditioning is apparent, along with a separatrix along the line  $\varepsilon = -g$  for  $g > 0$ . This corresponds to the vanishing of the beta function of external magnetic field  $h := \varepsilon + g$  along the line  $h \equiv 0$  for the analog spin glass system, illustrated in panel (b).

be to draw these from some probability distribution function (PDF), which is to introduce disorder on the line graph.

#### IV. RANDOM DISORDER AND INDUCED PROBABILITY FLOWS

A useful reformulation of the RG transformations (14) in terms of the variables  $y_n := e^{g_n}$  and  $x_n := e^{\varepsilon_n}$ , in which they take an algebraic form:

$$\begin{aligned}\tilde{y}_n &= \frac{(1 + y_n y_{n+1} x_{n+1})(1 + x_{n+1})}{(1 + y_n x_{n+1})(1 + y_{n+1} x_{n+1})}, \\ \tilde{x}_n &= x_n \frac{(1 + y_n x_{n+1})(1 + y_{n-1} x_{n-1})}{(1 + x_{n+1})(1 + x_{n-1})}.\end{aligned}\quad (20)$$

We consider each of these variables to be drawn from some statistical ensemble. Although the most general possibility has these being conditional probability assignments, a rich structure is nevertheless presented when restricting to cases where they are independently drawn from sitewise independent distributions.

We denote this set of independent probability distribution functions as  $\{\psi_n(y_n)\}$  for the  $y_n$  variables, and  $\{\chi_n(x_n)\}$  for the  $x_n$  variables. Given that  $\tilde{y}_n = \tilde{y}_n(y_n, y_{n+1}, x_{n+1})$ , and  $\tilde{x}_n = \tilde{x}_n(x_{n-1}, x_n, x_{n+1}, y_{n-1}, y_n)$ , it must be the case that under decimation, successive RG transformations induce a flow on the probability assignments given by

$$\begin{aligned}\tilde{\psi}_n(\tilde{y}_n \equiv y) &= \int_0^\infty \dots \int_0^\infty dx_{n+1} dy_n dy_{n+1} \psi_n(y_n) \\ &\quad \times \psi_{n+1}(y_{n+1}) \chi_{n+1}(x_{n+1}) \delta(\tilde{y}_n - y)\end{aligned}\quad (21)$$

and

$$\begin{aligned}\tilde{\chi}_n(\tilde{x}_n \equiv x) &= \int_0^\infty \dots \int_0^\infty dx_{n-1} dx_n dx_{n+1} dy_{n-1} dy_n \delta(\tilde{x}_n - x) \\ &\quad \times \psi_{n-1}(y_{n-1}) \psi_n(y_n) \chi_{n-1}(x_{n-1}) \chi_n(x_n) \\ &\quad \times \chi_{n+1}(x_{n+1}),\end{aligned}\quad (22)$$

where the dependent variables  $\tilde{y}_n$  and  $\tilde{x}_n$  that appear in the arguments of the delta functions are given by Eqs. (20). The renormalized distributions are normalized to unity by default, as seen by integrating both sides of the above with respect to  $y$  and  $x$ , respectively.

We first consider Eq. (21), where we note that the delta function integral can be performed by changing the integration variable to  $\tilde{y}_n$ . To do this, we must first invert  $\tilde{y}_n$  given by Eq. (20) in terms of the variable we look to transform, which we choose to be  $y_n$ . The result of this inversion is given by

$$y_n = \frac{(1 + x_{n+1}) - (1 + y_{n+1} x_{n+1}) \tilde{y}_n}{[x_{n+1} (1 + y_{n+1} x_{n+1}) \tilde{y}_n - (1 + x_{n+1}) y_{n+1} x_{n+1}]},\quad (23)$$

so that after accounting for the Jacobian factor and performing the subsequent  $\tilde{y}_n$  integral, one ends up with

$$\tilde{\psi}_n(y) = \int_0^\infty \int_0^\infty dy_{n+1} dx_{n+1} \left\{ \Theta(y_n) \psi_n(y_n[x_{n+1}, y_{n+1}, y]) \frac{\psi_{n+1}(y_{n+1}) \chi_{n+1}(x_{n+1}) (1 + x_{n+1}) (1 + y_{n+1} x_{n+1}) |y_{n+1} - 1|}{x_{n+1} [(1 + y_{n+1} x_{n+1}) y - (1 + x_{n+1}) y_{n+1}]^2} \right\},\quad (24)$$

where the Heaviside step function enforces the domain of integration to be such that Eq. (23) is positive for  $y = \tilde{y}_n$ . Proceeding analogously for Eq. (22), one obtains

$$\begin{aligned} \tilde{\chi}_n(\tilde{x}_n \equiv x) &= \int_0^\infty \cdots \int_0^\infty dy_{n-1} dy_n dx_{n-1} dx_{n+1} \left\{ \chi_n(x_n[\cdots]) \right. \\ &\times \psi_{n-1}(y_{n-1}) \psi_n(y_n) \chi_{n-1}(x_{n-1}) \chi_{n+1}(x_{n+1}) \\ &\times \left. \frac{(1+x_{n+1})(1+x_{n-1})}{(1+y_n x_{n+1})(1+y_{n-1} x_{n-1})} \right\}, \end{aligned} \quad (25)$$

where the implicit dependence of  $x_n$  on the remaining independent variables is given by

$$x_n = \tilde{x}_n \frac{(1+x_{n+1})(1+x_{n-1})}{(1+y_n x_{n+1})(1+y_{n-1} x_{n-1})}. \quad (26)$$

One can either specify the probability assignments  $\chi(x)$  and  $\psi(y)$ , or one can specify them for the original variables  $C(\varepsilon)$  and  $P(g)$  via the relations

$$\begin{aligned} \psi(y) &= P(g[y]) \frac{dg}{dy} \equiv \frac{P(g[y])}{y}, \\ \chi(x) &= C(\varepsilon[x]) \frac{d\varepsilon}{dx} \equiv \frac{C(\varepsilon[x])}{x}, \end{aligned} \quad (27)$$

where we have used that  $g = \log y$  and  $\varepsilon = \log x$ . Repeating the same steps that led to Eqs. (24) and (25), we obtain the induced RG flow for  $P(g)$  and  $C(\varepsilon)$  as

$$\begin{aligned} \tilde{P}_n(\tilde{g}_n \equiv g) &= \int_{-\infty}^\infty \int_{-\infty}^\infty dg_{n+1} d\varepsilon_{n+1} \left\{ \Theta(g_n) P_{n+1}(g_{n+1}) \right. \\ &\times \frac{e^g (1+e^{\varepsilon_{n+1}}) (1+e^{g_{n+1}+\varepsilon_{n+1}}) |e^{g_{n+1}} - 1|}{e^{\varepsilon_{n+1}} [(1+e^{g_{n+1}+\varepsilon_{n+1}}) e^g - (1+e^{\varepsilon_{n+1}}) e^{g_{n+1}}]^2} \\ &\times \left. C_{n+1}(\varepsilon_{n+1}) P_n(g_n[\varepsilon_{n+1}, g_{n+1}, \tilde{g}_n \equiv g]) \right\}, \end{aligned} \quad (28)$$

where  $g_n$  in the argument of  $P_n$  above is obtained by taking the log of Eq. (23) and substituting the definitions for the  $y$  and  $x$  variables. Similarly, one finds

$$\begin{aligned} \tilde{C}_n(\tilde{\varepsilon}_n \equiv \varepsilon) &= \int_{-\infty}^\infty \cdots \int_{-\infty}^\infty dg_{n-1} dg_n d\varepsilon_{n-1} d\varepsilon_{n+1} \\ &\times \left\{ P_n(g_n) P_{n-1}(g_{n-1}) C_{n-1}(\varepsilon_{n-1}) \right. \\ &\times \frac{(1+e^{\varepsilon_{n+1}})(1+e^{\varepsilon_{n-1}})}{(1+e^{g_n+\varepsilon_{n+1}})(1+e^{g_{n-1}+\varepsilon_{n-1}})} C_{n+1}(\varepsilon_{n+1}) \\ &\times \left. C_n(\varepsilon_n[g_{n-1}, g_n, \varepsilon_{n-1}, \varepsilon_{n+1}, \tilde{\varepsilon}_n \equiv \varepsilon]) \right\}. \end{aligned} \quad (29)$$

Although the convolutions in both sets of Eqs. (24) and (25) and Eqs. (28) and (29) are amenable to numerical integration for arbitrary probability assignments, it is informative to consider certain special cases.

### A. Sparse disorder

The simplest nontrivial example one could consider is the case where disorder is localized to a single site  $n$ , so that the bare probability assignments for the  $x$  and  $y$  variables are given by

$$\begin{aligned} \{\chi_a^{(0)}(x_a)\}_{a \neq n} &= \delta(x_a - x^{(0)}), \quad \chi_n^{(0)}(x_n) = \Omega(x_n), \\ \{\psi_a^{(0)}(y_a)\}_{a \neq n} &= \delta(y_a - y^{(0)}), \quad \psi_n^{(0)}(y_n) = \Pi(y_n), \end{aligned} \quad (30)$$

where  $\Omega(x_n)$  and  $\Pi(y_n)$  are the bare distributions of the  $x$  and  $y$  variables for the site where the disorder is localized. Inserting these into Eqs. (24) and (25) yields the following induced renormalization group transformations on the probability assignments upon iteration:

$$\begin{aligned} \chi_n^{(k+1)}(x) &= \int_0^\infty dy \left\{ \frac{\psi_n^{(k)}(y)(1+x^{(k)})^2}{(1+yx^{(k)})(1+y^{(k)}x^{(k)})} \right. \\ &\times \left. \chi_n^{(k)} \left[ \frac{x(1+x^{(k)})^2}{(1+yx^{(k)})(1+y^{(k)}x^{(k)})} \right] \right\}, \quad (31) \\ \psi_n^{(k+1)}(y) &= \frac{(1+x^{(k)})|y^{(k)} - 1|(1+x^{(k)}y^{(k)})}{x^{(k)}\{y+y^{(k)}[x^{(k)}(y-1)-1]\}^2} \\ &\times \psi_n^{(k)} \left[ \frac{1+x^{(k)} - y(1+x^{(k)}y^{(k)})}{x^{(k)}[y(1+x^{(k)}y^{(k)}) - y^{(k)}(1+x^{(k)})]} \right]_+, \end{aligned} \quad (32)$$

where dummy variables have been relabeled for compactness, and where

$$\begin{aligned} \{\chi_a^{(k+1)}(x)\}_{a \neq n} &= \delta(x - x^{(k+1)}), \\ \{\psi_a^{(k+1)}(y)\}_{a \neq n} &= \delta(y - y^{(k+1)}), \end{aligned} \quad (33)$$

where the arguments of the delta functions  $x^{(0)}$  and  $y^{(0)}$  in Eq. (30) get recursively renormalized to  $x^{(k)}$  and  $y^{(k)}$  via Eqs. (20) as

$$\begin{aligned} y^{(k+1)} &= \frac{(1+y^{(k)2}x^{(k)})(1+x^{(k)})}{(1+y^{(k)}x^{(k)})^2}, \\ x^{(k+1)} &= x^{(k)} \frac{(1+y^{(k)}x^{(k)})^2}{(1+x^{(k)})^2}. \end{aligned} \quad (34)$$

The effects of successive RG transformations are particularly straightforward to evaluate for the link disorder assignment  $\psi_n^{(k)}$  given that no convolutions are involved in Eq. (32). The subscript on Eq. (32) is to denote an implicit factor of the Heaviside step function that enforces the positivity of the argument of  $\psi_n^{(k)}$ .

For illustrative purposes, we choose the bare distributions for both  $x$  and  $y$  to be Gamma distributed over the interval  $[0, \infty)$ , so that

$$\Pi(y) = \frac{\lambda_\pi^{\alpha_\pi}}{\Gamma(\alpha_\pi)} y^{\alpha_\pi-1} e^{-\lambda_\pi y}, \quad \Omega(x) = \frac{\lambda_\omega^{\alpha_\omega}}{\Gamma(\alpha_\omega)} x^{\alpha_\omega-1} e^{-\lambda_\omega x}, \quad (35)$$

where the mean and the variance, given by  $\alpha/\lambda$  and  $\alpha/\lambda^2$ , respectively, which one can specify arbitrarily. Note that the  $x$  variables relate to a probability assignment for a

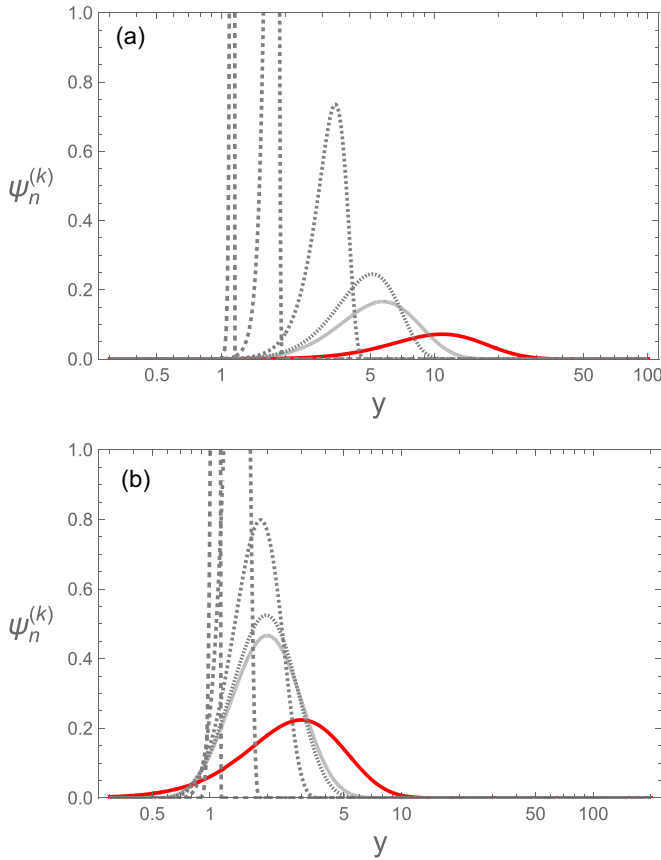


FIG. 3. (a) Renormalization group induced deformation of the bare probability assignment  $\Pi(y)$  given by Eq. (35) for  $\alpha_\pi = 5$ ,  $\lambda_\pi = e^{-1}$  (red line) for  $k$  from 1 to 5 (gray dashed lines with successively coarser dasheding). (b) The same, but for  $\Pi(y)$  given with parameters  $\alpha_\pi = 4$ ,  $\lambda_\pi = e^0$ .

corresponding Erdős-Rényi random graph via the relation [cf. Eq. (5)]

$$x = e^\varepsilon = \frac{p}{1-p}. \quad (36)$$

Upon repeated decimation transformations, we see that the disorder probability assignment would asymptote towards a delta function distribution centered around  $y = 1$  in Fig. 3, corresponding to the  $g = 0$  fixed line in Fig. 1. This is not particularly surprising, as the latter is a global attractor, and every realization within the disorder ensemble would flow towards this fixed line under repeated decimation.

One can also adapt these results to the case where one has disorder localized to multiple, but sufficiently sparsely situated sites.<sup>10</sup> The renormalization of the local probability assignments would proceed according to Eqs. (31) and (32) up until the RG step where any two sites where disorder has been assigned would coalesce under decimation. At this step, one has to revert to the expressions for the general transformation (24) and (25), iterating again from there with Eqs. (31) and (32) until the next step where coalescence occurs [cf. [38]].

<sup>10</sup>In more precise terms, this would be localized disorder that vanishes in the thermodynamic limit.

## B. Uniform disorder

A more general scenario would be to consider the  $\{y_a\}$  and  $\{x_a\}$  variables to each be independently and identically distributed. Although the general convolutions expressed in Eqs. (24) and (25) can be iterated in this case, it is useful to recast the former in simplified and more symmetrical form by transforming to the variables  $z_n, x_n$ , where  $z := x_n y_n = e^{\varepsilon_n + g_n}$ , so that Eqs. (20) transform to

$$\begin{aligned} \tilde{z}_n &= \frac{(x_n + z_n z_{n+1})(1 + z_{n-1})}{(1 + x_{n-1})(1 + z_{n+1})}, \\ \tilde{x}_n &= \frac{(x_n + z_n x_{n+1})(1 + z_{n-1})}{(1 + x_{n+1})(1 + x_{n-1})}. \end{aligned} \quad (37)$$

We now presume the disorder to be uniformly assigned as

$$\{\chi_n^{(0)}(x_n)\}_{\forall n} = \chi(x_n), \quad \{\psi_n^{(0)}(z_n)\}_{\forall n} = \psi(z_n), \quad (38)$$

for bare probability distributions  $\chi(x)$ ,  $\psi(z)$ , where the latter is given by the product distribution for the variables  $x$  and  $y$ , which can in principle be specified independently. Repeating the same steps that led to Eqs. (24) and (25), one arrives at the following induced RG transformations for the probability assignments at any given site:<sup>11</sup>

$$\begin{aligned} \chi^{(k+1)}(x) &= \int_0^\infty \cdots \int_0^\infty du dv dy dw \left\{ \frac{(1+w)(1+v)}{(y+uw)} \right. \\ &\quad \times \psi^{(k)} \left[ \frac{x(1+w)(1+v)}{(y+uw)} - 1 \right]_+ \\ &\quad \left. \times \psi^{(k)}(u) \chi^{(k)}(v) \chi^{(k)}(y) \chi^{(k)}(w) \right\} \end{aligned} \quad (39)$$

and

$$\begin{aligned} \psi^{(k+1)}(z) &= \int_0^\infty \cdots \int_0^\infty du dv dy dw \left\{ \psi^{(k)}(v) \psi^{(k)}(y) \right. \\ &\quad \times \psi^{(k)}(w) \frac{(u+yw)(1+v)}{z^2(1+w)} \chi^{(k)} \\ &\quad \left. \times \left[ \frac{(u+yw)(1+v)}{z(1+w)} - 1 \right]_+ \chi^{(k)}(u) \right\}, \end{aligned} \quad (40)$$

where again the  $+$  subscripts denote an implicit Heaviside theta function in the integrand to ensure the positivity of the arguments of  $\psi^{(k)}$  and  $\chi^{(k)}$  in the above. The two equations represent a set of coupled nonlinear convolutions that can be iterated repeatedly to evaluate the induced renormalization group flow on the bare probability assignments  $\chi^{(0)}$  and  $\psi^{(k)}$ . Their evaluation can be simplified

<sup>11</sup>Here, and in all analogous derivations that follow, one writes down the analogous equations to Eqs. (21) and (22), identifies a convenient integration variable which admits a unique inverse for the RG transformations for this variable alone analogous to Eqs. (20), and performs the delta function integration accounting for the appropriate Jacobian factor and integration limits.

in some cases by noticing that both convolutions can be expressed as

$$\begin{aligned} \chi^{(k+1)}(x) &= \frac{d}{dx} \int_0^\infty \cdots \int_0^\infty du dv dy dw \{\Psi^{(k)} \\ &\quad \times [f(x|u, v, y, w)] \psi^{(k)}(u) \chi^{(k)}(v) \chi^{(k)}(y) \chi^{(k)}(w)\}, \end{aligned} \quad (41)$$

$$\begin{aligned} \psi^{(k+1)}(z) &= -\frac{d}{dz} \int_0^\infty \cdots \int_0^\infty du dv dy dw \\ &\quad \times \{\Xi^{(k)}[g(z|u, v, y, w)] \chi^{(k)}(u) \psi^{(k)}(v) \\ &\quad \times \psi^{(k)}(y) \psi^{(k)}(w)\}, \end{aligned} \quad (42)$$

where  $\Psi^{(k)}(s)$  is the cumulative distribution function  $\Psi^{(k)}(s) = \int_0^s dt \Theta(t) \psi^{(k)}(t)$  where the step function under the integrand enforces the vanishing of  $\Psi$  where the integrand has no support, where  $\Xi^{(k)}(s)$  is defined analogously, and where the respective functional arguments are defined as

$$\begin{aligned} f(x|u, v, y, w) &:= \frac{x(1+w)(1+v)}{(y+wu)} - 1, \\ g(z|u, v, y, w) &:= \frac{(u+yw)(1+v)}{z(1+w)} - 1. \end{aligned} \quad (43)$$

Equation (41) has the advantage of recasting the induced renormalization group flow on the probability assignments as a gradient flow. One can also repeat the derivation to obtain the analogous expressions for the distribution of  $\varepsilon$  and the joint distribution for  $h = \varepsilon + g$ , denoted  $C^{(k)}(\varepsilon)$  and  $P^{(k)}(h)$ , respectively:

$$\begin{aligned} C^{(k+1)}(\varepsilon) &= \int_{-\infty}^\infty \cdots \int_{-\infty}^\infty du dv dy dw \\ &\quad \times \left\{ \left[ 1 - \frac{(e^y + e^{u+w})}{e^\varepsilon (1+e^w)(1+e^v)} \right]_+^{-1} P^{(k)} \right. \\ &\quad \times \left[ \log \left( \frac{e^\varepsilon (1+e^w)(1+e^v)}{(e^y + e^{w+u})} - 1 \right) \right] \\ &\quad \left. \times P^{(k)}(u) C^{(k)}(v) C^{(k)}(y) C^{(k)}(w) \right\} \end{aligned} \quad (44)$$

and

$$\begin{aligned} P^{(k+1)}(h) &= \int_{-\infty}^\infty \cdots \int_{-\infty}^\infty du dv dy dw \\ &\quad \times \left\{ \left[ 1 - \frac{(1+e^w)e^h}{(e^u + e^{y+w})(1+e^v)} \right]_+^{-1} C^{(k)} \right. \\ &\quad \times \left[ \log \left( \frac{(e^u + e^{y+w})(1+e^v)}{(1+e^w)e^h} - 1 \right) \right] \\ &\quad \left. \times C^{(k)}(u) P^{(k)}(v) P^{(k)}(y) P^{(k)}(w) \right\}, \end{aligned} \quad (45)$$

where the subscripts now on the Jacobian factors again restrict the integration domain to be such that the respective factors

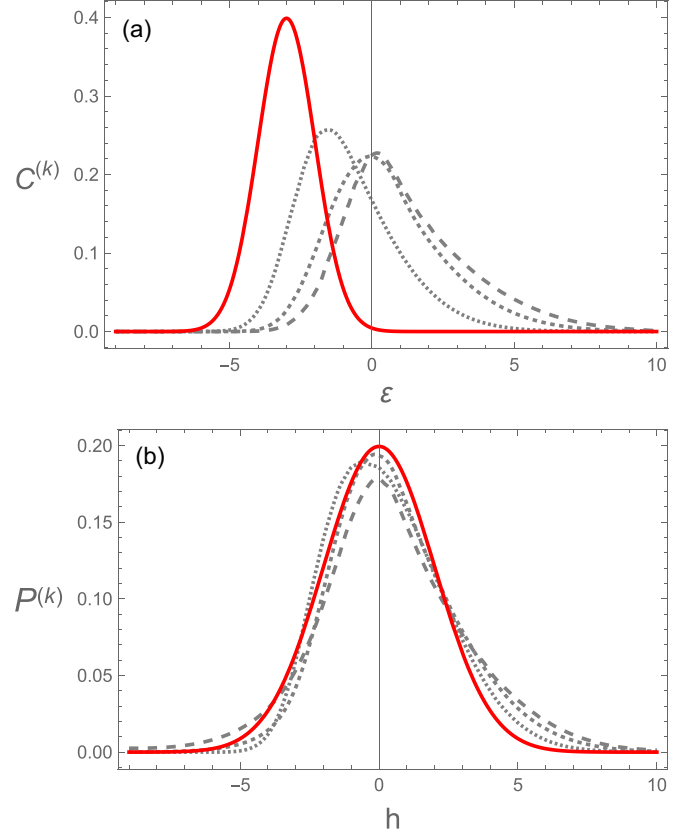


FIG. 4. Renormalization group induced deformation of the uniform probability assignments  $C^{(k)}(\varepsilon)$  (a) and  $P^{(k)}(h)$  (b). Bold red lines denote the bare probability assignments, taken to be Gaussians with means and variances given by  $\mu_\varepsilon = -3$ ,  $\sigma_\varepsilon = 1$  and  $\mu_h = -3$ ,  $\sigma_h = 1$ , respectively, and the gray dashed lines represent successive RG convolutions running from  $k = 1$  to 3 with increasingly coarse dashing.

are both positive. As before, one can recast both convolutions as an analogous gradient flow:

$$\begin{aligned} C^{(k+1)}(\varepsilon) &= \frac{d}{d\varepsilon} \int_{-\infty}^\infty \cdots \int_{-\infty}^\infty du dv dy dw \\ &\quad \times \{\Pi^{(k)}[F(\varepsilon|u, v, y, w)] \\ &\quad \times P^{(k)}(u) C^{(k)}(v) C^{(k)}(y) C^{(k)}(w)\}, \end{aligned} \quad (46)$$

$$\begin{aligned} P^{(k+1)}(h) &= -\frac{d}{dh} \int_{-\infty}^\infty \cdots \int_{-\infty}^\infty du dv dy dw \{\Xi^{(k)} \\ &\quad \times [G(h|u, v, y, w)] \\ &\quad \times C^{(k)}(u) P^{(k)}(v) P^{(k)}(y) P^{(k)}(w)\}, \end{aligned} \quad (47)$$

where  $\Pi^{(k)}(s) = \int_0^s dt \Theta(t) P^{(k)}(t)$ , and  $\Xi^{(k)}(s) = \int_0^s dt \Theta(t) C^{(k)}(t)$ , where we've defined

$$\begin{aligned} F(\varepsilon|u, v, y, w) &:= \log \left[ \frac{e^\varepsilon (1+e^w)(1+e^v)}{(e^y + e^{w+u})} - 1 \right], \\ G(h|u, v, y, w) &:= \log \left[ \frac{(e^u + e^{y+w})(1+e^v)}{(1+e^w)e^h} - 1 \right]. \end{aligned} \quad (48)$$

We illustrate the induced RG flow on the bare uniform disorder probability assignments in Fig. 4. Visual inspection of

both Figs. 3 and 4 seems to suggest the induced renormalization group flow to be akin to some sort of active transport process for the disorder probability assignments. This can in fact be made precise as we do in Appendix C, where we identify this as time-reversed drift diffusion, which is simply the cumulative result of pointwise deformations of the PDFs generated by the beta functions that generate the flow.

## V. IMPLICATIONS AND APPLICATIONS

Over the preceding chapters, we have shown how restricting, or truncating the effective Hamiltonian to quadratic order [cf. Eq. (2)] results in closed form renormalization group transformations whose repeated iterations result in pairwise interactions flowing towards increasing irrelevance. For infinite, or very large networks, one would say that such interactions flow to *triviality* at large scales. We defer the discussion as to whether this feature persists once higher order interactions are introduced to our concluding remarks, and discuss the implications of these facts for various applications in what follows. We do this by first distilling our findings to their bare elements, and then contextualizing them in a number of applied settings. These elements are the following:

(i) Disordered Erdős-Rényi random graphs with pairwise conditioning of the individual links admit closed form renormalization group transformations for maximum coordination number two.

(ii) Repeated RG transformations render these interactions irrelevant, flowing towards a disordered noninteracting Erdős-Rényi random graph.

(iii) Were one to draw the couplings of the effective Hamiltonian from a disorder ensemble, the effects of pairwise conditioning results in localized skewing of the macroscopic probability assignments relative to their microscopic (bare) assignments.

(iv) The induced RG flow on the disorder probability assignment corresponds to time reversed drift diffusion on the associated statistical manifold.

We now detail the ramifications of these elements for a number of applications. What follows is only a partial survey, with further applications, especially when higher order conditionings are incorporated and suggestive of future lines of inquiry revisited in our concluding remarks.

### A. Peer, preferential, susceptibility, and reinforcement effects on random networks

Random graphs are often invoked as a mathematical model for social networks, contact networks, or relationship networks more generally. Allowing the probabilities to form links depending on graph properties associated to a given node can lead to a variety of interesting network structures and topologies. Preferential attachment models [13,14] are a widely studied class of network growth models that associate the probability of nodes sequentially added to a graph to connect with existing nodes in a manner that depends on their degree (see, e.g., [44,45] for a survey). This is implemented via an attachment kernel, where the link formation probability is given by  $p_i = f(k_i) / \sum_j f(k_j)$ , where the sum is over the existing nodes, and where  $f(k_i)$  is some function of the degree

$k_i$  of the nodes. Within the context of a generalized statistical mechanical formalism, preferential attachment models are possible to realize for linear attachment kernels provided we go beyond maximum coordination number two. This can be seen by the fact that any conditioning of a given element of the adjacency matrix on the degree of a node that it might connect to would necessitate operators of the form

$$H(\mathbf{A}) \supset A_{ik} \sum_{j \neq k} A_{ij}. \quad (49)$$

Such terms would correspond to terms in an effective Hamiltonian of the general form Eq. (10) as opposed to Eq. (12) on the corresponding line graph representation. Nonlinear attachment kernels of the form  $p_i \propto f(k_i)$  would necessitate the addition of a series of higher order interactions of the form

$$H(\mathbf{A}) \supset A_{ik} \left( \sum_{j \neq k} A_{ij} \right)^n \quad (50)$$

that depend on the Taylor expansion of the kernel function  $f(k_i)$ , for which the bilinear truncation Eq. (10) would no longer be fit to purpose. Although for unweighted networks, the fact that the  $A_{ij} = \{0, 1\}$  implies that expressions of the form (50) can always be brought into the form (49), the resulting expressions will not be readily translatable back into a function of the degree, which could hinder inference problems that utilize the corresponding operator expansion in the effective Hamiltonian. That, along with the fact that operators of the form (50) are unavoidable for any generalization to unweighted networks, necessitates their consideration in generality.

It is, of course, incumbent upon any complete treatment to make precise the manner in which existing classes of random graph models transcribe into the generalized statistical mechanical formalism, which we expand upon in our concluding remarks. In what follows, we wish to proceed unencumbered by precedent and explore how qualitatively similar peer and preferential effects can naturally be modeled within the class of bilinear graph effective Hamiltonians. In order to do this, it first helps to visualize the action of the different operators that make up the effective Hamiltonian on a given graph.

In Fig. 5 we schematically represent a random graph with a fixed maximum number of nodes via the elements of the associated adjacency matrix, with each site representing a possible link.<sup>12</sup> In this representation, a link realized in a given draw is denoted by a circle, whose probabilities are determined by the set of  $\{\varepsilon_a\}$  in Eq. (12), with the remaining unconnected nodes discarded. We represent the maximum coordination number two pairwise interactions between links (i.e., the set of  $\{g_a\}$ ) as dashed lines that couple the sites in Fig. 6, caveating the footnote 12. From this figure, we see that we can actualize preferential attachment in a complimentary manner by assigning relatively larger positive couplings for all pairwise interactions that index a specific node, depicted with bold lines in the bottom panel of Fig. 6. We can also

<sup>12</sup>With the understanding that the ordering is inconsequential and the following discussion applicable to an arbitrary ordering.

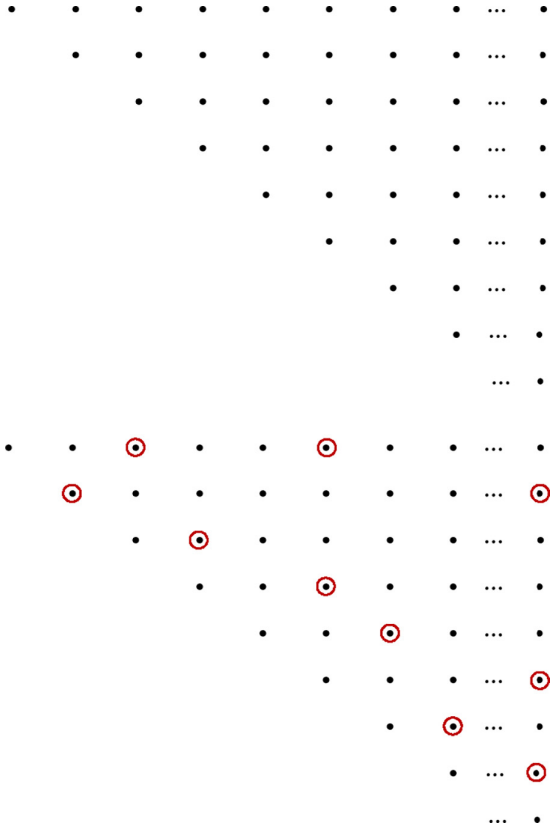


FIG. 5. Schematic representation of a random graph, where each dot represents a possible link (i.e., an off diagonal entry of the adjacency matrix). The realization of a given link in a specific draw corresponding to the relevant entry of the adjacency matrix are denoted by circles.

realize preferential *avoidance* via negative values for the bilinear couplings that index a particular set of nodes. The random assignment and distribution of nodes that have either preferential or avoidant conditioning can be implemented via a suitable choice for the disorder ensemble from which the couplings are drawn.

One can see immediately infer other ways in which one can induce qualitatively similar preferential effects on an ensemble of random graphs without invoking pairwise conditioning. One could, for instance, set all the  $\{g_a\}$  to be vanishing and specify the resulting disordered Erdős-Rényi probability assignments, i.e., the set of  $\{\varepsilon_a\}$ , to be arranged so that the desired distribution of degrees is attained, as implemented in [3,18,46,47]. In more precise terms, one can in principle reproduce a fixed number of summary statistics that derive from correlation functions of the graph degrees of freedom up to a given order from a continuous family of effective Hamiltonians. This degeneracy is a feature common to all effective theories, whereby one needs to determine up to a certain number of quantities in order to specify the couplings of the effective Hamiltonian completely. The degeneracy gets broken once one demands that a sufficient number of additional summary statistics be reproduced by the effective Hamiltonian, typically receiving contributions from higher order correlation functions, such as local clustering coefficients and the number and distribution of triangles, for example.

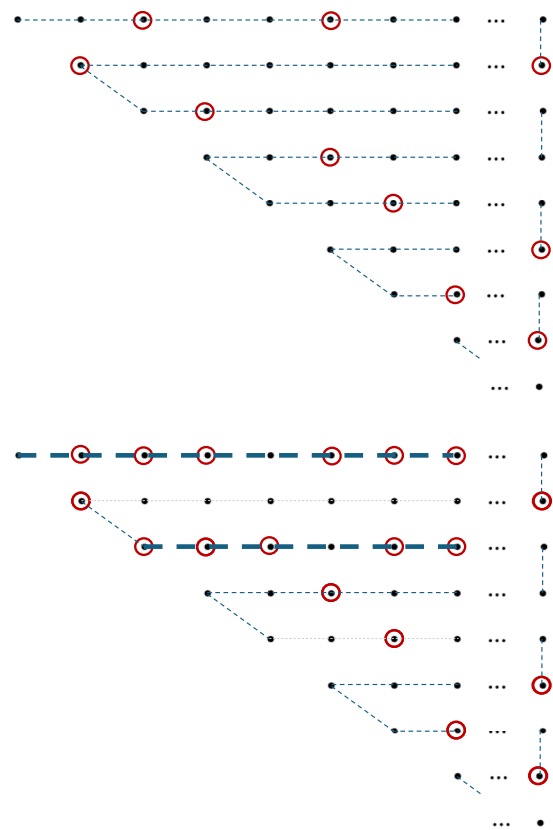


FIG. 6. Representation of a random graph with maximum coordination two interactions (top). Enhanced or diminished attachment and preference effects can be implemented via the sign and magnitude of the relevant bilinear couplings  $\{g_a\}$  (bottom: bold and faint lines, respectively).

One might also wonder how decimation on the line graph (where we decimate over links rather than nodes) maps back onto a random graph representation. Given that the marginalization effected on the line graph by each decimation reduces the number of link degrees of freedom by a factor of two, the result of two consecutive decimations would be to render a graph with  $\mathcal{N} = N(N - 1)/8$  links, which in the limit of large graphs corresponds to the number of links for a graph with  $\tilde{N} = N/2$  nodes.<sup>13</sup> One can therefore view the ensemble of networks generated by the effective Hamiltonian obtained by decimation on the line graph after an even number of RG transformations to admit a random network representation with  $\tilde{N} = N/2^m$  nodes, for some power  $m$ . As we have stressed in the introduction, this is by no means the only possibility as one can generate arbitrary RG transformations through the composition of transformations that decimate over individual links or sites. A more comprehensive treatment of this would fall within the ambit of a broader study of scheme dependence when renormalizing effective graph Hamiltonians, which deserves a separate detailed investigation.

The scaling behavior of preferential effects can be inferred from Figs. 1, 3, and 4, all of which illustrate the progressive

<sup>13</sup>There will be a number of hermit nodes that will result from this process, whose number go as  $1/N$ .

weakening and eventual vanishing of pairwise link conditioning at large enough scales. This is of course, when restricted to maximum coordination number two interactions, which merely transcribes the infinite temperature fixed point on a one-dimensional disordered Ising spin chain onto a random graph model, as discussed at the end of Sec. III. One can translate the implications of this into various applications straightforwardly.

Although utilized through some variant in a number of applications that model the structure and diffusion processes over social, contact, or population networks, it is widely appreciated that Erdős-Rényi random graphs by themselves fail to accurately model a variety of real-world networks. This is in part due to the Poisson nature of their degree distributions, which is seldom observed in application, necessitating additional structural requirements or restrictions [48–53]. One variant of particular utility is a random graph model of Molloy and Reed, which is constructed to reproduce an arbitrarily specified degree distribution [54], which nevertheless still comes up short in modeling the distance distributions observed in number of data sets of a given degree distribution [55]. Whether this is because of biased estimation in the given data sets, or because of unknown relations among elements of the graph which the model fails to capture is something that the effective generalized statistical mechanical framework offers the means to address and rectify, by allowing for either the possibility of arbitrary pairwise conditioning, or the possibility of modeling data limitations via assigning distribution priors on the relevant couplings.

In the context of social or contact networks, Fig. 1 offers a useful mnemonic for the scaling behavior of individual couplings encoded in Eq. (14) even if it restricts to the case of homogeneously assigned couplings. One sees that any localized positive or negative conditioning results in skewing the probabilities for random links to appear at macroscopic scales towards greater or lesser values, respectively. Positive conditioning can model, for instance, peer pressure or preferential effects on social networks, where the existence of a link between any two nodes makes it likelier for other links to form, or the converse. Similarly for contact networks, the probability that any two agents have been in contact with one another are certainly not independent, and positively conditioned to varying and inhomogeneously assigned degrees in general. Our findings show that such conditioning will eventually result in a disordered Erdős-Rényi assignment at the most coarse-grained level. One could therefore conceive of utilizing this conditioning in the context of temporal networks to model susceptibility effects in the context of opinion dynamics and reinforcement effects in the context of artificial and biological neural networks.

Dynamical processes on networks come in two broad categories. One can either take the underlying random graph as a fixed background on which dynamical processes occur, or one can view the structure of the graph itself to evolve over time. Both processes can also occur simultaneously and codependently. By assigning state variables to every node, one can associate a variety of dynamical structures with these variables that will depend on the network structure. Diffusion processes, for example, are naturally modeled via the graph Laplacian. More generally, one can model arbitrary dynamical

processes via operators constructed from the adjacency and degree matrices, as the graph analog of the derivative expansion (7). One might also consider the structure of the network to evolve over time as a temporal Markov process on a layered network, where each layer represents the state of the network at any given time step. Dynamical processes that interpolate between the layers can be derived from the properties of the graph structure, or state variables at the nodes, where coordination number two conditioning between links in adjacent layers can model arbitrary temporal correlations between layers.

In this way, the generalized statistical mechanical formalism developed here can readily be adapted towards classifying and studying the behavior of temporal and spatial correlations and conditioning, and their effects on late time or terminal summary statistics (such as the outcomes of elections) in opinion dynamic models based on random graphs [56–63]. Similarly, reinforcement effects at the cortical level in biological and artificial neural networks [64–70], when represented as layered temporal networks can also be modeled via pairwise, disordered, and inhomogeneously assigned pairwise conditioning, although any meaningful applications of relevance would almost certainly have to go beyond the maximum coordination number two approximation invoked in this investigation.

## B. Network inference with limited data

Yet another application of the framework developed in the preceding sections constitutes something of an inversion of the usual manner in which RG methods are utilized, which is towards inference and reconstruction problems [71,72] in the context of limited data resolution. The workhorse for this derives from the observation that in any line graph representation of a random network, decimation renormalization group transformations are formally equivalent to the process of statistical marginalization over the degrees of freedom that have been integrated out (cf. Appendix A, caveating Appendix B), which can also be viewed as marginalizing over variables we are unable to access or measure. This seemingly simple paraphrasing opens a portal for addressing how one can infer or reconstruct network representations with limited data. The most basic instance of this is when we simply have no access to, or have unreliable or noisy estimators for a finite subset of graph variables (see Fig. 7 for a schematic depiction of how uncertainty localizes on a network), in attempting to determine whether a link exists between any two nodes, or whether any two links are conditioned on each other. In statistical inference, one proceeds by viewing these as “nuisance parameters,” where by assigning prior probability assignments to model the uncertainty, and subsequently marginalizing over these parameters, one arrives at an effective description for the remaining observable degrees of freedom. This same process would translate into a local renormalization transformation on a graph representation that incorporates these unknown or inaccessible graph parameters via corresponding uncertainty priors.

A more sophisticated application of the machinery developed in the previous sections would go further, and consider the partial local inversion utilized in preceding sections to try

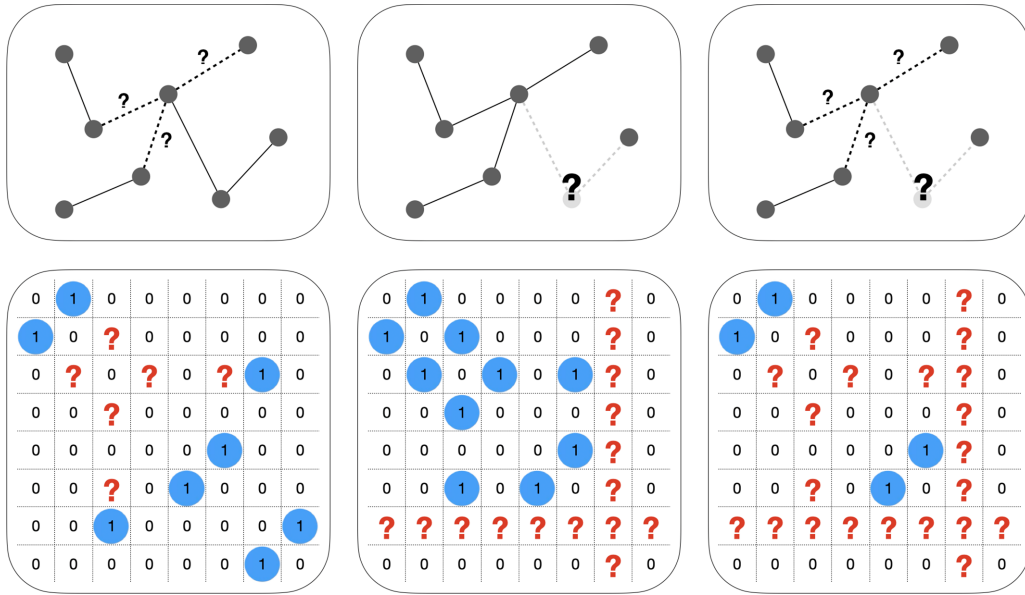


FIG. 7. Schematic depiction of incomplete network data. From left to right: individual links are missing or unresolvable; individual nodes are either unobserved, unresolvable, or of unknown connectivity; both possibilities simultaneously.

and “solve” for the family of underlying, or “fine-grained” parameters in terms of those that are inferred or measured at macroscopic scales. As we discuss in a companion paper [36], in situations where one has closed RG equations—as we do in the limited context of maximum coordination number two quadratic effective graph Hamiltonians—and where all parameters of a network are known except for those corresponding to a given link and its corresponding conditioning, one obtains two equations in two unknowns that can be locally inverted, allowing us to “reconstruct” the missing information up to the accuracy of graph data at the observed scale and the extent to which higher coordination number conditioning can be neglected.

If one is dealing with a system where the RG equations do not close, as will be the case when higher order coordination numbers and interactions are incorporated (which will be the case in the majority of practical applications), additional higher-order interactions will be generated under RG flow. On the other hand, the transformations themselves restrict the set of parameters that are consistent with renormalization group flow to have mapped to the accessible parameters. In this case, submanifolds in parameter space can be defined as level curves of the RG transformation equation with the observed, or macroscopic parameters. One can then impose additional conditions and priors to further restrict the “fine-grained” submanifold consistent with what is observed. Given that higher-order couplings often run to irrelevance under certain conditions, one has heuristic restrictions on parameter space dictated by the resolution at hand. The error one is willing to accept on the higher-order terms not being identically zero, coupled with the respective RG equation for that parameters become equations that determine the subspace of microscopic parameters consistent with what is observed, or inferred. A more in depth expansion on this approach can be found in [36], where we furthermore apply this formalism to economic and trade data of limited resolution.

## VI. CONCLUDING REMARKS AND FUTURE DIRECTIONS

The work presented in the preceding sections represents a first attempt towards addressing the question of what aspects of interactions, conditioning, and correlations across a complex network are relevant for its behavior at macroscopic scales. A more complete treatment—necessary if any of the insights gained are to apply across the breadth of possible applications—would have to incorporate higher order conditioning and coordination numbers in its analysis. To do this in full generality is of course intractable, but the subtle yet highly advantageous shift in operational perspective offered by the paradigm of effective theories focuses this problem into the more tractable question of classifying parameter (ir)relevance in any effective generalized statistical mechanical formalism. By allowing for higher order and coordination number interactions in the effective Hamiltonian (2), one can conceive of proceeding via perturbative techniques to explore the flow of couplings within certain truncations.

For a limited class of interactions that represent parametrized deviations from the maximum coordination number two case studied in the previous sections, one can be assured of the irrelevance of such higher order and higher coordination number couplings from analogous results on lattice systems. On the other hand, it will certainly be the case that novel and qualitatively different phenomena will arise when one no longer assumes higher order couplings to be perturbatively small. Nevertheless, any incremental gains in understanding the parameter relevance and possible phase structure of higher order and coordination number interactions will be of great use in any attempt to classify universality classes of effective graph Hamiltonians. It is to be stressed that within the generalized statistical mechanical framework, extensions to weighted and directed graphs can readily be accommodated via the operator expansion (2) and the configuration space integrated or summed over.

Another important application, if not consistency check on the generalized effective statistical mechanical formalism would be to construct and identify the explicit operator expansions that correspond to existing network models. However, perhaps the most important direction to pursue from the present vantage would be towards applications of topical relevance. Even with the preliminary results for pairwise coordination number two interactions, a number of extensions towards models reliant on random graph representations was partly surveyed in previous sections, and will be elaborated upon in further detail in a followup investigation [36].

### ACKNOWLEDGMENTS

This publication is part of the project “Network renormalization: from theoretical physics to the resilience of societies,” with File No. NWA.1418.24.029 of the research program NWA L3–Innovative projects within routes 2024, which is (partly) financed by the Dutch Research Council (NWO) under the grant [73], and the project “Redefining renormalization for complex networks,” with File No. OCENW.M.24.039 of the research program Open Competition Domain Science Package 24-1, which is (partly) financed by the Dutch Research Council (NWO) under the grant [74]. This work is also supported by the European Union–NextGenerationEU–National Recovery and Resilience Plan (Piano Nazionale di Ripresa e Resilienza, PNRR), projects “SoBigData.it–Strengthening the Italian RI for Social Mining and Big Data Analytics,” Grant No. IR0000013 (no. 3264, 28 Dec. 2021) [75], and “Reconstruction, Resilience and Recovery of Socio-Economic Networks,” RECON-NET EP\_FAIR\_005 - PE0000013 “FAIR”-PNRR M4C2 Investment 1.3. Work by A.C. is also supported by Grant No. NEXTGENERATIONEU (NGEU) and funded by the Italian Ministry of University and Research (MUR), National Recovery and Resilience Plan (NRRP), project MNESYS (PE0000006): A Multiscale integrated approach to the study of the nervous system in health and disease (DN. 1553 11.10.2022). S.P. wishes to thank Paolo Benincasa, La Ricotta Project, and the Asociación Cultural La Ricotta for hospitality over the writing of this manuscript as well as many useful discussions, for which he is also grateful to Kevin Grosvenor.

### DATA AVAILABILITY

The data are not publicly available. The code and data are available from the authors upon reasonable request.

### APPENDIX A: GENERATING FUNCTIONALS, AND WHEN RENORMALIZATION $\equiv$ MARGINALIZATION

Given any set of  $\mathcal{N}$  random variables denoted  $X = \{x_1, \dots, x_{\mathcal{N}}\}$  that are drawn from an ensemble defined by the probability measure  $P(x_1, \dots, x_{\mathcal{N}})$ , all statistical moments can be obtained via the corresponding generating functional  $Z[J]$ , defined as

$$Z[J] = \int d^{\mathcal{N}}x P(x_1, \dots, x_{\mathcal{N}}) e^{-(J,x)}, \quad (\text{A1})$$

where the scalar product  $(J, x) := \sum_{i=1}^{\mathcal{N}} J_i x_i$ . From this definition, it follows straightforwardly that

$$\langle x_1^{a_1}, \dots, x_{\mathcal{N}}^{a_{\mathcal{N}}} \rangle = (-1)^M \frac{\partial^M Z[J]}{\partial J_1^{a_1} \partial J_2^{a_2} \dots \partial J_{\mathcal{N}}^{a_{\mathcal{N}}}} \Big|_{J=0}, \quad (\text{A2})$$

where we’ve defined  $M := \sum_{i=1}^{\mathcal{N}} a_i$ , and furthermore require  $\int d^{\mathcal{N}}x P(x_1, \dots, x_{\mathcal{N}}) = 1$ . Similarly, given a probability measure  $P(\mathbf{A})$  defining an ensemble of  $\mathcal{N} \times \mathcal{N}$  real matrices, where

$$\int \mathcal{D}\mathbf{A} P[\mathbf{A}] = 1; \quad \mathcal{D}\mathbf{A} := \prod_{i,j=1}^{\mathcal{N}} dA_{ij}, \quad (\text{A3})$$

one can define the analogous generating functional [76]:

$$Z[J] = \int \mathcal{D}\mathbf{A} P[\mathbf{A}] e^{-\text{Tr} \mathbf{A} J^T}. \quad (\text{A4})$$

Expectation values of arbitrary powers of the matrix  $\mathbf{A}$  can be obtained from successive derivatives of  $Z[J]$

$$\langle \mathbf{A}_{ij}^m \rangle = (-1)^m \frac{\delta^m Z[J]}{\delta J_{i'j'} \delta J_{i''j''} \dots \delta J_{i'j'} \delta J_{i''j''}} \Big|_{J=0}. \quad (\text{A5})$$

Were one to consider the random matrix ensemble to correspond to adjacency matrices of a random graph, additional restrictions on the probability measure will be required depending on whether the network is weighted or not, or directed or not. Unweighted, undirected graphs correspond to symmetric matrices with entries that are either vanishing, or unity. For example, the Gilbert-Erdős-Rényi model [77,78] for  $N$  nodes where links can appear independently with probability  $p$  would have an associated probability measure given by

$$\begin{aligned} (i < j) P_{ij} &= p \delta(A_{ij} - 1) + (1 - p) \delta(A_{ij}), \\ (i = j) P_{ij} &= \delta(A_{ii}), \\ (i > j) P_{ij} &= \delta(A_{ij} - A_{ji}). \end{aligned} \quad (\text{A6})$$

Since the probability measure factorizes into identical independently drawn probabilities, the generating functional  $Z[J]$  is straightforwardly evaluated as

$$\begin{aligned} Z[J] &= \prod_{i,j} \int dA_{ij} P_{ij} e^{-A_{ij} J_{ij}} \\ &= (1 - p)^{\frac{N(N-1)}{2}} \prod_{i < j} (1 + e^{[\varepsilon - J_{(ij)}]}), \end{aligned} \quad (\text{A7})$$

where we’ve defined

$$\varepsilon := \log \left( \frac{p}{1-p} \right), \quad J_{(ij)} := J_{ij} + J_{ji}. \quad (\text{A8})$$

One can also define the generating function for all connected correlation functions  $W[J]$  as

$$e^{-W[J]} := Z[J], \quad (\text{A9})$$

which is given by

$$W[J] = \frac{N(N-1)}{2} \log \left( \frac{1}{1-p} \right) - \sum_{i < j} \log(1 + e^{[\varepsilon - J_{(ij)}]}). \quad (\text{A10})$$

Discarding the source independent term above [which corresponds to discarding the source-independent normalization factor in Eq. (A7)], one obtains

$$\begin{aligned} Z[J] &= \prod_{i<j} (1 + e^{\varepsilon - J_{ij}}), \\ W[J] &= - \sum_{i<j} \log(1 + e^{\varepsilon - J_{ij}}). \end{aligned} \quad (\text{A11})$$

Identifying  $W[J]|_{J=0}$  as the free energy associated with the corresponding exponential random graph model, one sees that it is equivalent to the thermodynamic free energy of a gas of noninteracting fermions at zero chemical potential localized at the links

$$W[J]|_{J=0} = - \sum_{i<j} \log [1 + e^{\varepsilon_{ij}}], \quad (\text{A12})$$

where the above is the generalization to the case where the individual link probabilities  $p$  vary across the graph [i.e.,  $p \rightarrow p_{ij}$  in Eq. (A6)]. This is particularly manifested when one evaluates the link occupation number:

$$\langle \mathbf{A}_{ij} \rangle = - \frac{1}{Z} \frac{\delta Z}{\delta J_{ij}} \Big|_{J=0} \equiv n_{ij} = \frac{1}{e^{\varepsilon_{ij}} + 1} = p_{ij}, \quad (\text{A13})$$

where one makes the identification

$$\varepsilon_{ij} \equiv \frac{(E - \mu)_{ij}}{k_B T}. \quad (\text{A14})$$

By taking additional derivatives of  $Z[J]$  with respect to the  $J_{ij}$ , one can calculate expectation values of any observable derivable from the adjacency matrix. For example, the average number of paths of length  $k$  connecting nodes  $i$  and  $j$ , defined as  $\#_{ij}(k)$  is given by

$$\#_{ij}(k) = \langle \mathbf{A}_{ij}^k \rangle, \quad (\text{A15})$$

and the expectation value of the degree of the  $i$ th node  $\langle k_i \rangle$ , or the average degree of the entire network  $\langle k \rangle$  can be obtained as

$$\langle k_i \rangle = \sum_{j=1}^N \langle \mathbf{A}_{ij} \rangle, \quad \langle k \rangle = \frac{1}{N} \sum_{i,j=1}^N \langle \mathbf{A}_{ij} \rangle, \quad (\text{A16})$$

and so on. The analogs of the above expressions for general heterogeneous networks can also be found, for instance, in [8,79].

Consider now the situation where we have a partition of our set of random variables into two disjoint subsets  $X = \{\mathbf{y}, \mathbf{x}\}$ , where without loss of generality, we order them such that  $\{\mathbf{y}\} := \{x_1, \dots, x_k\}$  and  $\{\mathbf{x}\} := \{x_{k+1}, \dots, x_N\}$ , with joint probability distribution given by  $\Pi(\mathbf{y}, \mathbf{x})$ . We furthermore presume that we either have no means of accessing or measuring the  $\mathbf{x}$  degrees of freedom, or we are simply uninterested in computing any statistical moments that involve them. In that case, any observable  $O(\mathbf{y})$  that is constructed exclusively out of the  $\mathbf{y}$  variables can be reexpressed as

$$\begin{aligned} \langle O(\mathbf{y}) \rangle &= \int d\mathbf{y} d\mathbf{x} O(\mathbf{y}) \Pi(\mathbf{y}, \mathbf{x}) \\ &= \int d\mathbf{y} O(\mathbf{y}) \int d\mathbf{x} \Pi(\mathbf{y}, \mathbf{x}) := \int d\mathbf{y} O(\mathbf{y}) \Pi_{\text{eff}}(\mathbf{y}), \end{aligned}$$

where

$$\Pi_{\text{eff}}(\mathbf{y}) := \int d\mathbf{x} \Pi(\mathbf{y}, \mathbf{x}). \quad (\text{A17})$$

In the context of statistical inference, this corresponds to the well-worn procedure of marginalization. When viewed from the vantage of all correlation functions being generated by a moment generating functional (or partition function), where the probability measure can be expressed as a (graph) Hamiltonian via the equivalence  $\Pi(\mathbf{y}, \mathbf{x}) = e^{-\beta H(\mathbf{y}, \mathbf{x})}$ , Eq. (A17) becomes

$$\int d\mathbf{x} e^{-\beta H(\mathbf{y}, \mathbf{x})} := e^{-\beta' H_{\text{eff}}(\mathbf{y})}, \quad (\text{A18})$$

which is formally the process of computing probabilities with an effective Hamiltonian that has integrated out the  $\{\mathbf{x}\}$  degrees of freedom.

In the context of physical systems, the resulting effective probability measure (A18) goes by various names depending on the criteria used to separate the  $\{\mathbf{x}\}$  and  $\{\mathbf{y}\}$  variables. In the case where  $\{\mathbf{x}\}$  represents short wavelength (or high energy) degrees of freedom that we have no access to, and the  $\{\mathbf{y}\}$  variables represent the long wavelength degrees of freedom that we construct all observable quantities out of, the resulting coarse grained probability measure goes by the name of the (Euclidean) *Kadanoff-Wilson effective action*. In certain cases, one might instead be interested in integrating out, or marginalizing over *all* modes of a given field or set of degrees of freedom regardless of their energetics. In the context of statistical field theory, such an object would be known as the *one particle irreducible effective action* [35]. In both cases, the formally exact implementation of renormalization can be viewed as the moment generating functional equivalent of statistical marginalization.

## APPENDIX B: WHEN RENORMALIZATION $\neq$ MARGINALIZATION

The process of separating degrees of freedom by some criteria, and integrating out the subset designated to be marginalized over can be iterated repeatedly. Each step constitutes a renormalization group transformation, and corresponds to coarse graining if the separation criteria is a resolution scale.<sup>14</sup> Although the complete execution of this procedure as described would indeed be equivalent to statistical marginalization, there are a number of situations where this will not be the case. For instance, it may be desirable to redefine the variables after every RG step, so that after each transformation, we have

$$\mathbf{y}^{(n)} \rightarrow \mathbf{y}'^{(n)}[\mathbf{y}^{(n)}]. \quad (\text{B1})$$

Such transformations may be necessitated in order to render subsequent integrations of the form (A18) to be tractable. They may also implement redefinitions in to "coarse-grained" variables (such as the block averaged variables familiar from spin systems) that one might deem to be more useful aggregate

<sup>14</sup>This, of course, presumes prior notions of metric distance, which may not always be available in the context of random networks.

variables according to some criteria. In both cases, one is not forced to make such transformations when dealing with finite systems, which typically facilitate approximations whose artifacts one must take care to identify when determining macroscopic observables. The result of these transformations would be to induce a transformation on the effective Hamiltonian obtained at each step as

$$H_{\text{eff}}^{(n)}(\mathbf{y}^{(n)}) \rightarrow \tilde{H}_{\text{eff}}^{(n)}(\mathbf{y}^{(n)}). \quad (\text{B2})$$

It is here that the nontrivial notion of *scheme dependence* creeps into the discussion, which is perhaps best illustrated through example.

Consider a regular Erdős-Rényi random graph defined by the probability measure (A6) with the associated generating functional (A7). Given the normalization of the probability measure for a finite graph with  $N$  nodes, integrating out all links associated to a given node would simply return the probability measure for the links among the remaining  $N - 1$  nodes, which is also described by the measure (A6), but omitting indices corresponding to the node that has been integrated out. In this case, the result of each successive RG transformation is the trivial mapping

$$p \rightarrow p' = p \text{ (marginalization)}. \quad (\text{B3})$$

On the other hand, one could instead coarse grain a specific draw of the ensemble into “block-averaged” variables via the prescription depicted in Fig. 8. In this case, the probability of there being a link between any of the links within the blocked nodes is given by one minus the probability that no links appear. Therefore,

$$p \rightarrow p' = 1 - (1 - p)^4 \text{ (block averaging)}, \quad (\text{B4})$$

which has an unstable fixed point at  $p = 0$ , and a stable fixed point at  $p = 1$ . It is clear that one gets a result different than in Eq. (B3) because of the nontrivial transformation induced by block averaging at each step.

On the other hand, one can also imagine averaging in blocks larger than two, as nothing forces us a priori to consider coarse graining degrees of freedom via a halving procedure. Repeating the steps that led to Eq. (B4) for blocks of three nodes at a time, one obtains the transformation  $p \rightarrow 1 - (1 - p)^9$ . More generally, one obtains  $p \rightarrow 1 - (1 - p)^{m^2}$  when averaging over blocks consisting of  $m$  nodes. In all cases regardless of the prescription, one finds the following fixed point structure:

$$p_* = 0 \text{ (unstable)}, \quad p_* = 1 \text{ (stable)}.$$

That is, although there is scheme dependence in the precise nature of the RG transformations and any intermediate quantities one may derive from them, *the existence and values of the fixed points is identical in all schemes*. This is an example of a familiar phenomenon in particle physics effective theories, whereby scheme dependence itself is a diagnostic that one hasn't yet computed an unambiguously measurable quantity, rather, an artifact of one's approximations. Scheme *independence* therefore provides a useful classification tool to identify quantities that represent legitimate macroscopic

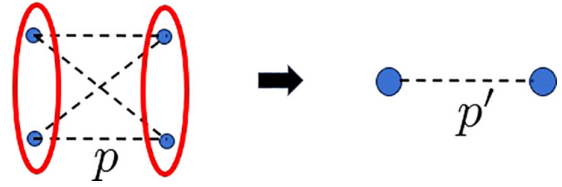


FIG. 8. Schematic depiction of a “block-averaging” renormalization of an Erdős-Rényi random graph, where each successive RG step involves grouping nodes into blocked pairs, where the renormalized probability is defined as the probability that there exists at least one link connecting nodes across blocks.

observables.<sup>15</sup> In the context of the generalized statistical mechanics of random graph models, it is the existence of fixed points and the critical parameter values taken at them that represent such scheme independent quantities when block averaging.

Although we are under no prior compulsion to make successive transformations of the form (B1) in the context of finite systems, various approximation schemes, or the need to impose additional conditions derived from constraints on observational or data resolution may necessitate them in practice. On the other hand, in the context of very large systems, i.e. in the idealized limit where the number of nodes  $N$  is taken to infinity (sometimes referred to as the thermodynamic limit), successive changes of variable, or “field redefinitions,” are *essential* in order to have a well defined framework to calculate probabilities. One of the reasons for this can immediately be inferred from the generalization of Eq. (A18) to a system with a continuum, or an infinite number of degrees of freedom:

$$\int d\mathbf{x} e^{-\beta H(\mathbf{y}, \mathbf{x})} := \tilde{\mathcal{N}} e^{-\beta' H_{\text{eff}}(\mathbf{y})}, \quad (\text{B5})$$

where  $\tilde{\mathcal{N}}$  is the requisite normalization factor, that can itself be potentially divergent. Such a factor was already implicit in Eq. (A18) but equal to unity by construction, as the probability measure can always be taken for granted as appropriately normalized when dealing with a finite number of degrees of freedom.

One has to take additional care when confronted with infinite systems due to the potential divergences that arise can from the continuum or thermodynamic idealizations [20,35]. For present purposes, however, we simply note that the normalization in Eq. (B5) drops out of all observables derived from derivatives of the free energy  $W = -\log Z$ , and that successive *renormalizations* can be absorbed by redefinitions

<sup>15</sup>For example, among the various quantities whose renormalization group running one may attempt to compute the running of in the context of continuum field theories, the form of the free energy (or effective potential) away from its critical points is generally scheme dependent, but the existence and values attained at critical points is not [80]. Furthermore, one can show that although the structure of divergent quantities that need regularization can be scheme dependent, the coefficients of logarithmic divergences, which dictate how physically observable quantities run with scale is scheme independent [81].

of the remaining variables and the effective Hamiltonian at each step according to Eq. (B1) (wave function renormalization) and Eq. (B2) (renormalization of couplings). These redefinitions themselves can be formally divergent. Nevertheless, a well-worn prescription for calculating probabilities is still possible after a process that begins with first introducing an arbitrary *regularization scale*. This can, for instance, be a minimum resolution scale  $\ell$ , or a finite cutoff on the number of degrees of freedom  $N$ , where one takes  $\ell \rightarrow 0$  or  $N \rightarrow \infty$  at the end of the computation. The result will be formal divergences in either limit, which we then subtract with appropriate *counterterms* in the effective Hamiltonian and measure factors. This process won't be without ambiguities, as subtracting infinities can result in differing finite remainders depending on the precise subtraction method. This is dealt with by fixing the finite remainder via demanding that the effective Hamiltonian reproduce the values of certain measurements at some fixed scale. This final step—imposing *renormalization conditions*—only has to be done for a finite number of terms when working to a given order in parameter relevance.

The totality of the aforementioned steps—regularization, subtraction, and the imposition of renormalization conditions—constitutes the entire renormalization procedure in systems with an infinite number of degrees of freedom. After performing each one to completeness, one is left with a well defined formalism to unambiguously compute all relevant statistical moments, from which all observable quantities can be derived.

### APPENDIX C: INDUCED RG FLOW AS TIME REVERSED DRIFT DIFFUSION

Consider the space of couplings of an effective Hamiltonian that is closed under action of the renormalization group.<sup>16</sup> In the context of Eq. (12), this would be the space  $\mathbb{R}^{2\mathcal{N}}$ , but more generally it could be an arbitrary manifold  $\mathcal{M}$  (possibly with a boundary), with points that can be identified via local coordinates  $\theta^a$ , where  $a$  runs from 1 to  $d$ , the dimension of  $\mathcal{M}$ . The probability assignments for random draws of the adjacency matrix defined in Eq. (1) implicitly depend on the couplings of the effective Hamiltonian, so that more explicitly

$$P(\mathbf{A}|\theta) \equiv \frac{e^{-H(\mathbf{A}|\theta)}}{Z(\theta)}, \quad Z(\theta) := \sum_{\{\mathbf{A}\}} e^{-H(\mathbf{A}|\theta)}. \quad (\text{C1})$$

Each point in  $\mathcal{M}$  parameterized by the coordinates  $\theta^a$  corresponds to a specific effective Hamiltonian that generates random graphs according to the probability assignment above.  $\mathcal{M}$  is therefore a *statistical manifold*, where every point corresponds to a probability distribution function, which

furthermore comes equipped with a natural metric tensor given by the Fisher-Rao metric [26–28]:

$$G_{ab}(\theta) = \left\langle \frac{\partial \log P}{\partial \theta^a} \frac{\partial \log P}{\partial \theta^b} \right\rangle, \\ = \sum_{\{\mathbf{A}\}} P(\mathbf{A}) \frac{\partial \log P(\mathbf{A}|\theta)}{\partial \theta^a} \frac{\partial \log P(\mathbf{A}|\theta)}{\partial \theta^b}. \quad (\text{C2})$$

Although many interesting structures relating to parameter relevance and renormalization group flow can be inferred from and for the Fisher-Rao metric alone,<sup>17</sup> it will play only a spectator role in what follows.

We are interested in computing the RG induced flow on the disorder probability assignments, which we represent by a probability measure over the statistical manifold itself. Denoting this measure as  $\Pi(\theta) \equiv \Pi(\theta^1, \theta^2, \dots, \theta^d)$ , the renormalization of the couplings  $\theta^a$  induce an RG flow for  $\Pi(\theta)$  according to

$$D\Pi(\theta) = \beta^a \partial_a \Pi(\theta), \quad (\text{C3})$$

where  $\partial_a := \partial/\partial\theta^a$  and repeated indices summed over, and where we've defined the beta functions

$$D\theta^a = \beta^a(\theta). \quad (\text{C4})$$

Here the notation  $D\theta^a$  is shorthand either for  $D\theta^a := \partial\theta^a/\partial t$  where  $t$  is the renormalization scale in the context of continuum systems,<sup>18</sup> or  $D\theta^a := \theta_{k+1}^a - \theta_k^a$  for discrete RG transformations, where the subscript denotes the relevant RG step, with  $\theta_0$  being the bare (i.e., initial) values of the couplings. In the present context, one can interchangeably consider the  $\theta^a$  to be the  $d = 2\mathcal{N}$  dimensional vectors  $\theta^a := (\varepsilon_n, g_n), (x_n, y_n)$ , or  $(x_n, z_n)$ , whose corresponding beta functions would follow from Eqs. (14), (20), or (37), respectively. By recasting Eq. (C3) as

$$(D - \beta^a \partial_a)\Pi(\theta) = 0, \quad (\text{C5})$$

one obtains the general solution

$$\Pi(\theta) = F[\theta^a(\tau)], \quad F[\theta^a(0)] \equiv \Pi(\theta_0^a), \quad (\text{C6})$$

where  $F$  is any arbitrary function over  $\mathcal{M}$ , and where  $\theta^a(\tau)$  is the solution to Eq. (C4) with initial conditions  $\theta^a(0) = \theta_0^a$ . In

<sup>17</sup>In addition to the direct correspondence between the eigenvalues of the Fisher-Rao metric and RG parameter relevance discussed earlier [24], one can also show that the metric undergoes a geometric flow under coarse graining induced by the beta function vector field [82]. The information metric also geometrizes the study of the phase structure and critical exponents of a statistical mechanical system, as can be seen via Eqs. (C1) and (C2), so that  $G_{ab}(\theta) = -\partial_a \partial_b \log Z$ . Recognizing the latter as the generalized free energy, we see immediately that singularities on the statistical manifold identify phase transitions [83–85], and allows one to define and compute distances between theories [86].

<sup>18</sup>It is sometimes convenient to define continuum RG transformations as  $\mu d\theta^a/d\mu$ , where  $\mu$  is some energy or length scale, in which case we identify  $t/t_0 = \log \mu/\mu_0$ , where the subscripted quantities refer to a reference scale, typically where renormalization conditions are imposed [20,81] (cf. Appendix B).

<sup>16</sup>The following can be generalized to situations where this assumption is relaxed with additional care.

the above,  $\tau$  denotes either the continuum variable  $t$ , or discrete variable  $k$  that indexes the flow, and the bare probability assignment is denoted  $\Pi(\theta_0)$ . In the continuum case, the  $\theta^a$  coordinatize the integral curves of the vector field generated by the beta functions.

The induced renormalization group flow on the probability assignments can be straightforwardly interpreted via Eqs. (C4) and (C6): the bare PDF gets a pointwise functional deformation, with each point flowing along its integral curves defined by the respective beta functions. As seen from both Figs. 3 and 4, the induced RG flow exhibits behavior akin to time reversed drift diffusion. This is in fact a general property of induced RG flow on the disorder probability assignments. In order to see this, we first make some preliminary observations, which for concreteness we frame in terms of the continuum limit.<sup>19</sup> The first is that conservation of probability under the action of RG transformations (cf. discussion below Eq. (22)) implies that

$$\frac{\partial}{\partial t} \int_{\mathcal{M}} d^d \theta \Pi(\theta) = \int_{\mathcal{M}} d^{d-1} \theta \beta^a \partial_a \Pi(\theta) = 0, \quad (\text{C7})$$

so that

$$\langle\langle \partial_a \beta^a \rangle\rangle = \int_{\partial \mathcal{M}} d^{d-1} \theta \Pi(\theta) n_a \beta^a, \quad (\text{C8})$$

where  $n_a$  denotes a normal vector at the boundary of  $\mathcal{M}$ , and where double angled brackets denote an average over the disorder:

$$\langle\langle f(\theta) \rangle\rangle := \int_{\mathcal{M}} d^d \theta \Pi(\theta) f(\theta). \quad (\text{C9})$$

This is to be contrasted with averaging over the random variables, which we denote with single-angled brackets as in Eq. (C2). For statistical manifolds without boundary, or whose boundary is at infinity (where the probability measure has to decay sufficiently rapidly to ensure a well defined normalization), the right-hand side of Eq. (C8) vanishes so that

$$\langle\langle \partial_a \beta^a \rangle\rangle = 0, \quad (\text{C10})$$

which represents a globally (as opposed to locally) conserved disorder probability current. The lack of a locally conserved current is consistent with induced RG flows generally admitting local sources and sinks, corresponding to unstable and stable fixed points of the flow, even as disorder probability is globally conserved under action of the renormalization group.

At this stage, it behooves us to be more precise about the disorder probability measure. The freedom to reparameterize the disorder variables (as we have done freely in the previous sections), which is to allow for arbitrary coordinate transformations on  $\mathcal{M}$ , would require us to work covariantly and incorporate a volume factor  $\sqrt{G}$  (with  $G$  being the determinant of  $G_{ab}$ ) into the measure, with all derivatives corresponding to covariant derivatives with metric

compatible connections. Although the derivation that follows proceeds analogously for arbitrary connections, we restrict ourselves to flat connections over a specific coordinatization such that the measure is as indicated in Eqs. (C7) and (C9). This is not just to simplify the discussion, but also because a more comprehensive treatment would also admit more general possibilities for connections on the statistical manifold (e.g., the so-called one parameter family of  $\alpha$  connections [27,28]), which would deserve a separate discussion.<sup>20</sup>

Acting twice on the probability measure with the operator  $\partial_t = \beta^a \partial_a$  results in the expression

$$\partial_t^2 \Pi(\theta, t) = (\beta^c \partial_c \beta^a) \partial_a \Pi(\theta, t) + \beta^a \beta^c \partial_a \partial_c \Pi(\theta, t). \quad (\text{C11})$$

The first term in parentheses on the right-hand side can be recognized as the directional derivative of the beta function vector field tangent to itself, and can without loss of generality be decomposed as

$$\beta^c \partial_c \beta^a = \frac{\beta^a}{\kappa_{\parallel}} + \frac{\hat{v}_{\perp}^a}{\kappa_{\perp}}, \quad (\text{C12})$$

where  $\hat{v}_{\perp}^a$  is defined to be a unit norm orthogonal drift vector that itself depends on  $\theta$ , as do  $\kappa_{\parallel}$  and  $\kappa_{\perp}$ , where the latter can be viewed as analogous to the inverse radius of curvature of the vector flow of  $\beta^a$  in a Frenet-Serret formalism [87,88]. As defined,  $\kappa_{\parallel}$  will always be a positive nonzero number. With the decomposition above and recalling the definition  $\partial_t = \beta^a \partial_a$ , when one presumes the hierarchy  $\partial_t \Pi \gg \kappa_{\parallel} \partial_t^2 \Pi$  (or enforces it by rescaling the parameter  $t$ ), one arrives at

$$-\frac{d}{dt} \Pi(\theta, t) = v_d^a \partial_a \Pi(\theta, t) + \kappa_{\parallel} \beta^a \beta^b \partial_a \partial_b \Pi(\theta, t) \quad (\text{C13})$$

with  $v_d^a := \frac{\kappa_{\parallel}}{\kappa_{\perp}} \hat{v}_{\perp}^a$ , which formally describes time reversed anisotropic drift diffusion for the probability assignments under renormalization group flow.

We stress in closing the present discussion that in the context of continuum systems, the so-called Exact Renormalization Group implementation of Wilsonian renormalization [89–92] can be also be recast as functional diffusion [93–95] for probability assignments of the form  $P(\mathbf{x}) = e^{-H(\mathbf{x})}/Z$  in analogy with Eq. (C1). In particular, Polchinsky's formulation of the ERG [90], which can be viewed as a particular case of the Wegner-Morris flow equation [96,97], has a finite dimensional analog that can be recast as

$$\frac{d}{dt} P(\mathbf{x}, t) = \partial_i [v^i P(\mathbf{x}, t)] + g^{ij}(\mathbf{x}, t) \partial_i \partial_j P(\mathbf{x}, t), \quad (\text{C14})$$

where  $v^i$  is drift term that depends of the particular regularization scheme [93,94] (cf. Appendix B for a recap of the

<sup>19</sup>Which can also be viewed as the  $N \rightarrow \infty$  limit in the discrete case, so that summations can be approximated by integrals and RG transformations can be iterated arbitrarily.

<sup>20</sup>In the present context, the probability assignments Eq. (C1) correspond to an exponential family of distributions in the information geometric framework, and our choice to use the flat metric connections corresponds to working with an  $\alpha$  connection with  $\alpha = 1$  [27,28].

renormalization procedure for continuum systems). Although Eq. (C14) appears to be analogous to Eq. (C13), their physical content is very different as  $P$  refers to the probability

assignment for the stochastic degrees of freedom Eq. (C1), whereas  $\Pi$  refers to the disorder probability assignment over the statistical manifold.

- 
- [1] D. Strauss, On a general class of models for interaction, *SIAM Rev.* **28**, 513 (1986).
- [2] O. Häggström and J. Jonasson, Phase transition in the random triangle model, *J. Appl. Probab.* **36**, 1101 (1999).
- [3] J. Park and M. E. J. Newman, Statistical mechanics of networks, *Phys. Rev. E* **70**, 066117 (2004).
- [4] J. Park and M. E. J. Newman, Solution of the two-star model of a network, *Phys. Rev. E* **70**, 066146 (2004).
- [5] J. Park and M. E. J. Newman, Solution for the properties of a clustered network, *Phys. Rev. E* **72**, 026136 (2005).
- [6] A. Coolen, A. Annibale, and E. Roberts, *Generating Random Networks and Graphs* (Oxford University Press, Oxford UK, 2017).
- [7] T. Squartini and D. Garlaschelli, *Maximum-Entropy Networks: Pattern Detection, Network Reconstruction and Graph Combinatorics* (Springer, Cham, Switzerland, 2017).
- [8] G. Cimini, T. Squartini, F. Saracco, D. Garlaschelli, A. Gabrielli, and G. Caldarelli, The statistical physics of real-world networks, *Nat. Rev. Phys.* **1**, 58 (2019).
- [9] E. T. Jaynes, Information theory and statistical mechanics, *Phys. Rev.* **106**, 620 (1957).
- [10] E. T. Jaynes, Information theory and statistical mechanics. II, *Phys. Rev.* **108**, 171 (1957).
- [11] C. E. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J.* **27**, 379 (1948).
- [12] A. Catanzaro, S. Patil, and D. Garlaschelli, A general solution for network models with pairwise edge coupling, [arXiv:2503.20805](https://arxiv.org/abs/2503.20805).
- [13] A.-L. Barabási and R. Albert, Emergence of scaling in random networks, *Science* **286**, 509 (1999).
- [14] R. Albert and A.-L. Barabási, Statistical mechanics of complex networks, *Rev. Mod. Phys.* **74**, 47 (2002).
- [15] G. Caldarelli, A. Capocci, P. De Los Rios, and M. A. Munoz, Scale-free networks from varying vertex intrinsic fitness, *Phys. Rev. Lett.* **89**, 258702 (2002).
- [16] D. Garlaschelli and M. I. Loffredo, Fitness-dependent topological properties of the world trade web, *Phys. Rev. Lett.* **93**, 188701 (2004).
- [17] R. van der Hofstad, *Random Graphs and Complex Networks* (Cambridge University Press, Cambridge, UK, 2016), Vol. 1.
- [18] R. Van Der Hofstad, *Random Graphs and Complex Networks* (Cambridge University Press, Cambridge, UK, 2024), Vol. 2.
- [19] A. Catanzaro, R. S. Hazra, and D. Garlaschelli, Spectra of random graphs with discrete scale invariance, [arXiv:2509.12407](https://arxiv.org/abs/2509.12407).
- [20] C. P. Burgess, *Introduction to Effective Field Theory: Thinking Effectively About Hierarchies of Scale* (Cambridge University Press, Cambridge, UK, 2020).
- [21] A. V. Manohar, Introduction to effective field theories, [arXiv:1804.05863](https://arxiv.org/abs/1804.05863).
- [22] R. Shankar, Effective field theory in condensed matter physics, [arXiv:cond-mat/9703210](https://arxiv.org/abs/cond-mat/9703210).
- [23] A. M. J. Schakel, Boulevard of broken symmetries, [arXiv Condensed Matter](https://arxiv.org/abs/cond-mat/9703210) (1998).
- [24] B. B. Machta, R. Chachra, M. K. Transtrum, and J. P. Sethna, Parameter space compression underlies emergent theories and predictive models, *Science* **342**, 604 (2013).
- [25] K. N. Quinn, M. C. Abbott, M. K. Transtrum, B. B. Machta, and J. P. Sethna, Information geometry for multiparameter models: New perspectives on the origin of simplicity, [arXiv:2111.07176](https://arxiv.org/abs/2111.07176).
- [26] C. R. Rao, Information and accuracy attainable in the estimation of statistical parameters, in *Breakthroughs in Statistics*, Springer Series in Statistics, edited by S. Kotz and N. L. Johnson (Springer, New York, NY, 1992), pp. 235–247.
- [27] F. Nielsen, An elementary introduction to information geometry, *Entropy* **22**, 1100 (2020).
- [28] S. Amari and H. Nagaoka, *Methods of Information Geometry*, Translations of Mathematical Monographs (American Mathematical Society, 2000).
- [29] M. K. Transtrum, B. B. Machta, K. S. Brown, B. C. Daniels, C. R. Myers, and J. P. Sethna, Perspective: Sloppiness and emergent theories in physics, biology, and beyond, *J. Chem. Phys.* **143**, 010901 (2015).
- [30] L. P. Kadanoff, Scaling, universality and operator algebras, *Phase Transitions Crit. Phenom.* **1**, 240 (1976).
- [31] F. J. Wegner, The critical state, general aspects, in *Phase Transitions and Critical Phenomena*, Vol. 6, edited by C. Domb and M. S. Green (Academic, New York, 1976).
- [32] L. D. Landau and E. M. Lifshitz, *Physical Kinetics*, Course of Theoretical Physics, Vol. 10 (Pergamon Press, Oxford, 1981).
- [33] S. Succi, *The Lattice Boltzmann Equation: For Fluid Dynamics and Beyond*, Numerical Mathematics and Scientific Computation (Oxford University Press, Oxford UK, 2013).
- [34] K. T. Grosvenor, N. A. Obers, and S. P. Patil, Hydrodynamics without boost-invariance from kinetic theory: From perfect fluids to active flocks, *SciPost Phys.* **19**, 071 (2025).
- [35] M. Helias and D. Dahmen, *Statistical Field Theory for Neural Networks* (Springer International Publishing, Cham, Switzerland, 2020).
- [36] A. Catanzaro, S. P. Patil, and D. Garlaschelli (unpublished).
- [37] A. Gabrielli, D. Garlaschelli, S. P. Patil, and M. Á. Serrano, Network renormalization, *Nat. Rev. Phys.* **7**, 203 (2025).
- [38] Y. Bar-Yam and S. P. Patil, Renormalization of sparse disorder in the Ising model, [arXiv:1805.12556](https://arxiv.org/abs/1805.12556).
- [39] A. Migdal, Phase transitions in gauge and spin-lattice systems, *Zh. Eksp. Teor. Fiz.* **69**, 1457 (1975).
- [40] L. P. Kadanoff, Notes on Migdal’s recursion formulas, *Ann. Phys.* **100**, 359 (1976).
- [41] L. P. Kadanoff, *Statistical Physics: Statics, Dynamics and Renormalization* (World Scientific Publishing Co. Pte. Ltd., Singapore, 2000).
- [42] L. Onsager, Crystal statistics. I. A two-dimensional model with an order-disorder transition, *Phys. Rev.* **65**, 117 (1944).

- [43] H. J. Hilhorst, M. Schick, and J. M. J. van Leeuwen, Exact renormalization group equations for the two-dimensional Ising model, *Phys. Rev. B* **19**, 2749 (1979).
- [44] M. Newman, *Networks* (Oxford University Press, Oxford, 2018).
- [45] G. G. Piva, F. L. Ribeiro, and A. S. Mata, Networks with growth and preferential attachment: Modeling and applications, [arXiv:2007.01349](https://arxiv.org/abs/2007.01349).
- [46] E. Garuccio, M. Lalli, and D. Garlaschelli, Multiscale network renormalization: Scale-invariance without geometry, *Phys. Rev. Res.* **5**, 043101 (2023).
- [47] V. D. P. Servedio, G. Caldarelli, and P. Buttà, Vertex intrinsic fitness: How to produce arbitrary scale-free networks, *Phys. Rev. E* **70**, 056126 (2004).
- [48] C. Jiang, Y. Chen, and K. J. R. Liu, Evolutionary dynamics of information diffusion over social networks, *IEEE Trans. Signal Process.* **62**, 4573 (2014).
- [49] C. Schmeltzer, J. Soriano, I. M. Sokolov, and S. Rüdiger, Percolation of spatially constrained Erdős-Rényi networks with degree correlations, *Phys. Rev. E* **89**, 012116 (2014).
- [50] M. Alizadeh, C. Cioffi-Revilla, and A. Crooks, Generating and analyzing spatial social networks, *Comput. Math. Organ. Theory* **23**, 362 (2017).
- [51] X. Shi, L. A. Adamic, and M. J. Strauss, Networks of strong ties, *Physica A* **378**, 33 (2007).
- [52] K. Kandhway and J. Kuri, Campaigning in heterogeneous social networks: Optimal control of SI information epidemics, *IEEE/ACM Trans. Networking* **24**, 383 (2016).
- [53] C. Seshadhri, T. G. Kolda, and A. Pinar, Community structure and scale-free collections of Erdős-Rényi graphs, *Phys. Rev. E* **85**, 056109 (2012).
- [54] M. Molloy and B. Reed, A critical point for random graphs with a given degree sequence, *Random Struct. Algorithms* **6**, 161 (1995).
- [55] M. E. J. Newman, D. J. Watts, and S. H. Strogatz, Random graph models of social networks, *Proc. Natl. Acad. Sci. USA* **99**, 2566 (2002).
- [56] G. Albi, L. Pareschi, and M. Zanella, Opinion dynamics over complex networks: Kinetic modeling and numerical methods, [arXiv:1604.00421](https://arxiv.org/abs/1604.00421).
- [57] S. E. Parsegov, A. V. Proskurnikov, R. Tempo, and N. E. Friedkin, Novel multidimensional models of opinion dynamics in social networks, *IEEE Trans. Autom. Control* **62**, 2270 (2017).
- [58] J. Ghaderi and R. Srikant, Opinion dynamics in social networks with stubborn agents: Equilibrium and convergence rate, *Automatica* **50**, 3209 (2014).
- [59] G. Travieso and L. da Fontoura Costa, Spread of opinions and proportional voting, *Phys. Rev. E* **74**, 036112 (2006).
- [60] M. Z. Rácz and D. E. Rigobon, Towards consensus: Reducing polarization by perturbing social networks, *IEEE Trans. Network Sci. Eng.* **10**, 3450 (2023).
- [61] G. Schoenebeck and F.-Y. Yu, Consensus of interacting particle systems on Erdős-Rényi graphs, in *Proceedings of the 2018 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)* (SIAM, 2018), pp. 1945–1964.
- [62] J. Xie, J. Emenheiser, M. Kirby, S. Sreenivasan, B. K. Szymanski, and G. Korniss, Evolution of opinions on social networks in the presence of competing committed groups, *PLoS ONE* **7**, e33215 (2012).
- [63] F. D. Domenico, F. Caccioli, and G. Livan, Peer influence breaks ergodicity in an opinion dynamics model with external information, [arXiv:2507.06661](https://arxiv.org/abs/2507.06661).
- [64] C. Fan, L. Yao, J. Zhang, Z. Zhen, and X. Wu, Advanced reinforcement learning and its connections with brain neuroscience, *Research* **6**, 0064 (2023).
- [65] K. Doya, Reinforcement learning: Computational theory and biological mechanisms, *HFSP J.* **1**, 30 (2007).
- [66] T. Alkam, A. H. Van Benschoten, and E. Tarshizi, Reinforcement learning in artificial intelligence and neurobiology, *Neurosci. Inf.* **5**, 100220 (2025).
- [67] M. Khoshkhou and A. Montakhab, Optimal reinforcement learning near the edge of a synchronization transition, *Phys. Rev. E* **105**, 044312 (2022).
- [68] N. Rungratsameetaweemana, R. Kim, T. Chotibut, and T. J. Sejnowski, Random noise promotes slow heterogeneous synaptic dynamics important for robust working memory computation, *Proc. Natl. Acad. Sci. USA* **122**, e2316745122 (2025).
- [69] D. C. Mocanu, E. Mocanu, P. Stone, P. H. Nguyen, M. Gibescu, and A. Liotta, Scalable training of artificial neural networks with adaptive sparse connectivity inspired by network science, *Nat. Commun.* **9**, 2383 (2018).
- [70] Y. Arya and S. H. Lee, Effects of structural properties of neural networks on machine learning performance, [arXiv:2507.10005](https://arxiv.org/abs/2507.10005).
- [71] G. Cimini, R. Mastrandrea, and T. Squartini, *Reconstructing Networks* (Cambridge University Press, Cambridge, UK, 2021).
- [72] T. Squartini, G. Caldarelli, G. Cimini, A. Gabrielli, and D. Garlaschelli, Reconstruction methods for networks: The case of economic and financial systems, *Phys. Rep.* **757**, 1 (2018).
- [73] <https://doi.org/10.61686/AOIJPO5368>.
- [74] <https://doi.org/10.61686/PBSEC42210>.
- [75] <https://pnrr.sobigdata.it/>.
- [76] A. D. Jackson and S. P. Patil, Phases of small worlds: A mean field formulation, *J. Stat. Phys.* **189**, 40 (2022).
- [77] P. Erdős and A. Rényi, On random graphs I, *Publ. Math. Debrecen* **6**, 18 (1959).
- [78] E. N. Gilbert, Random graphs, *Ann. Math. Stat.* **30**, 1141 (1959).
- [79] T. Squartini, R. Mastrandrea, and D. Garlaschelli, Unbiased sampling of network ensembles, *New J. Phys.* **17**, 023052 (2015).
- [80] A. Andreassen, W. Frost, and M. D. Schwartz, Consistent use of the standard model effective potential, *Phys. Rev. Lett.* **113**, 241801 (2014).
- [81] A. Negro and S. P. Patil, An Étude on the regularization and renormalization of divergences in primordial observables, *Riv. Nuovo Cim.* **47**, 179 (2024).
- [82] A. Raju, B. B. Machta, and J. P. Sethna, Information loss under coarse graining: A geometric approach, *Phys. Rev. E* **98**, 052112 (2018).
- [83] W. Janke, D. Johnston, and R. Kenna, Information geometry and phase transitions, *Physica A* **336**, 181 (2004).
- [84] B. P. Dolan, D. A. Johnston, and R. Kenna, The information geometry of the one-dimensional Potts model, *J. Phys. A: Math. Gen.* **35**, 9025 (2002).

- [85] J. Erdmenger, K. Grosvenor, and R. Jefferson, Information geometry in quantum field theory: Lessons from simple examples, *SciPost Phys.* **8**, 073 (2020).
- [86] V. Balasubramanian, J. J. Heckman, and A. Maloney, Relative entropy and proximity of quantum field theories, [arXiv:1410.6809](https://arxiv.org/abs/1410.6809).
- [87] F. Frenet, Sur les courbes à double courbure, *J. Math. Pures Appl.* **17**, 890 (1852).
- [88] J. A. Serret, Sur quelques formules relatives à la théorie des courbes à double courbure, *J. Math. Pures Appl.* **16**, 153 (1851).
- [89] K. G. Wilson and J. Kogut, The renormalization group and the  $\epsilon$  expansion, *Phys. Rep.* **12**, 75 (1974).
- [90] J. Polchinski, Renormalization and effective lagrangians, *Nucl. Phys. B* **231**, 269 (1984).
- [91] C. Bagnuls and C. Bervillier, Exact renormalization group equations: An introductory review, *Phys. Rep.* **348**, 91 (2001).
- [92] J. Berges, N. Tetradis, and C. Wetterich, Non-perturbative renormalization flow in quantum field theory and statistical physics, *Phys. Rep.* **363**, 223 (2002).
- [93] J. Cotler and S. Rezchikov, Renormalization group flow as optimal transport, *Phys. Rev. D* **108**, 025003 (2023).
- [94] D. S. Berman, M. S. Klinger, and A. G. Stapleton, Bayesian renormalization, *Mach. Learn.: Sci. Technol.* **4**, 045011 (2023).
- [95] D. S. Berman and M. S. Klinger, The inverse of exact renormalization group flows as statistical inference, *Entropy* **26**, 389 (2024).
- [96] F. Wegner, Some invariance properties of the renormalization group, *J. Phys. C* **7**, 2098 (1974).
- [97] T. R. Morris, A gauge invariant exact renormalisation group. (i), *Nucl. Phys. B* **573**, 97 (2000).