



Universiteit
Leiden
The Netherlands

On the renormalization of random network models

Catanzaro, A.

Citation

Catanzaro, A. (2026, September 10). *On the renormalization of random network models*. Retrieved from <https://hdl.handle.net/1887/4309543>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4309543>

Note: To cite this publication please use the final published version (if applicable).

Renormalization of 1-D Interacting Random Graph

This chapter is based on:

Alessio Catanzaro, Diego Garlaschelli, and Subodh P. Patil. "*Renormalization of Interacting Random Graph Models*". Phys. Rev. E 113, 024314 (2026).

Abstract

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 non-trivial 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 pair-wise 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.

§5.1 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 [47, 50, 90]. 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* [111, 112]. The workhorse of this generalization lies in viewing the Gibbs entropy formula as a use case of Shannon information entropy [113], 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 priors¹. 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.

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

$$P(\mathbf{A}) \equiv \frac{e^{-H(\mathbf{A})}}{Z}; \quad Z := \sum_{\{A_{ij}\}} e^{-H(\mathbf{A})}, \quad (5.1)$$

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:

$$H(\mathbf{A}) = -h^{(0)} - \sum_{ij} h_{ij}^{(1)} A_{ij} - \sum_{ij \neq kl} h_{ijkl}^{(2)} A_{ij} A_{kl} - \sum_{\substack{ij \neq kl, \\ mnkl \neq mn}} h_{ijklmn}^{(3)} A_{ij} A_{kl} A_{mn} - \dots \quad (5.2)$$

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. 5.1) for a sufficient number of observables $\{\mathcal{O}_a\}$ be constrained to assigned values:

$$\langle \mathcal{O}_a \rangle = \sum_{\mathbf{A}} P(\mathbf{A}) \mathcal{O}_a(\mathbf{A}), \quad (5.3)$$

¹In 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 [111, 112].

subject to maximization of the Gibbs-Shannon entropy

$$S = - \sum_{\mathbf{A}} P(\mathbf{A}) \log P(\mathbf{A}). \quad (5.4)$$

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-Renyi random graph, for which all couplings in Eq. 5.2 vanish except for $h_{ij}^{(1)} = h_{ji}^{(1)} = \log[p/(1-p)]$ for $i \neq j$ (see e.g. [114]). A disordered variant of the Erdős-Renyi graph is obtained by allowing the $h_{ij}^{(1)}$ to vary link-wise, so that

$$h_{ij}^{(1)} = \log \left(\frac{p_{ij}}{1 - p_{ij}} \right). \quad (5.5)$$

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

$$Z \equiv Z_1 Z_2 \dots Z_{\mathcal{N}}, \quad \mathcal{N} := N(N-1)/2, \quad (5.6)$$

where each individual factor corresponds to the partition function for each possible individual link² — 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 [7, 5]. 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

²For example, the partition function for a graph with N nodes factorizes as per Eq. 5.6 when one truncates the expansion beyond terms linear in A_{ij} in Eq. 5.2.

models are often referred to as fitness (or hidden variable) models [11, 12, 47, 95] 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.

§5.1.1 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* [51, 52, 53, 58]. The idea is simple to state – one simply views the Hamiltonian in Eq. 5.2 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 accuracy³. 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. 5.2 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 [115, 116], in a manner that is straightforward to articulate.

Sequentially higher order terms in the effective Hamiltonian (Eq. 5.2) 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

³This presumes that one works in an operator basis such that the analog of Eq. 5.2 represents an expansion whose leading order terms correspond to an exactly solvable system (however non-linear or strongly correlated), where the higher order terms parameterize collective modes that perturb away from this solution in a convergent series.

couplings will correspond to the largest eigenvalues of the associated Fisher-Rao information metric [117, 118, 119], 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:

$$\frac{d\rho}{dt} = \sigma\rho + D_2\nabla^2\rho + D_4\ell^2\nabla^4\rho + D_6\ell^4\nabla^6\rho + \dots \quad (5.7)$$

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. 5.7 represents a convergent operator Taylor expansion where successive terms are subleading in even powers of k_L/k_ℓ . We'll return to the physical significance of the so far arbitrary scale ℓ 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 $D_4 \sim D_2 k_\ell^2/k_L^2 \gg D_2$. Similarly, one would need $D_6 \sim D_4 k_\ell^2/k_L^2 \gg D_4$ in order for the sextic operator to compete with the quartic, and so on. Conversely, unless the D_4 coefficient starts to approach $D_4 \approx D_2 k_\ell^2/k_L^2$, 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 $|D_4| \ll |D_2 k_\ell^2/k_L^2|$ 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 re-expands the effective equations of motion as in Eq. 5.7 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 [115, 116, 120]⁴. All that remains is to clarify the nature of the scale ℓ in Eq. 5.7: In

⁴The classification of operators into categories of relevance, marginal relevance, and irrelevance

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 ℓ , the resulting equations of motion for *all* modes with wavelengths longer than ℓ would be reproduced by an operator expansion corresponding to Eq. 5.7, parametrically weighted by factors of ℓ^2 with order unity numerical prefactors. If one views diffusion processes as having coarse grained an underlying kinetic theory description [123, 124], one can explicitly demonstrate this via an influence kernel parameterization of inter-particle forces (see e.g. [125] in the context of broken symmetries).

The realization that one can parameterize 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 theories⁵, 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.

§5.1.2 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 parameterization of unweighted random graph models under renormalization group flow. Our starting point is the operator expansion Eq. 5.2, 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

under the action of the renormalization group is owed to Kadanoff and Wegner [121, 122].

⁵In 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 infra-red (see e.g. [51, 52]).

partition function for the graph will not be factorizable into a product of independent partition functions for each link. That is,

$$Z \neq Z_1 Z_2 \dots Z_N, \quad (5.8)$$

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 non-trivial case of this is when one restricts to bilinear interactions, which can be written without loss of generality as:

$$H(\mathbf{A}) = - \sum_{i \leq j} \varepsilon_{ij} A_{ij} - \frac{1}{2} \sum_{\substack{i \leq j, k \leq l \\ (i,j) \neq (k,l)}} \beta_{ijkl} A_{ij} A_{kl}, \quad (5.9)$$

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 and 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. [126]). 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. 5.2 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 Eq. 5.9, 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. 5.9 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 [127], and commence our formal study of random graph renormalization in the following sections.

§5.1.3 Outline

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 non-trivial 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 section 5.2, 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 5.3, and generalize to the disordered setting in 5.4, where we immediately find that the problem can be recast in terms of determining the induced flow on the disorder probability assignments. 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 *infra-red* 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.

§5.2 Line graphs, spin chains, and closed RG transformations

The graph Hamiltonian Eq. 5.9 can be re-expressed as the Hamiltonian of a disordered spin chain via mapping to the line graph:

$$H(\mathbf{A}) = - \sum_a \varepsilon_a A_a - \frac{1}{2} \sum_{a \leq b} \beta_{ab} A_a A_b, \quad (5.10)$$

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. 5.10. 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 graph⁶. Moreover, the couplings ε_a and β_{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 RG transformation that corresponds to decimation on the line graph. Specifically, we consider interactions of the form:

$$\beta_{ab} = g_a \delta_{a,b-1} + g_b \delta_{a,b+1}, \quad (5.11)$$

⁶For a recent review on progress and open issues regarding network renormalization, see [54].

so that

$$H(\mathbf{A}) = - \sum_{a=1}^{\mathcal{N}} [\varepsilon_a + g_a A_{a+1}] A_a, \quad (5.12)$$

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 ε_a and g_a themselves could be drawn from independent ensembles, corresponding to arbitrary ‘local’ disorder.

As discussed in Chapter 4, the formal process of integrating out one or more of the edge degrees of freedom in the context of Eq. 5.12 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 [127] 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 couplings⁷:

$$\begin{aligned} \tilde{g}_{n-1} &= \log \left\{ \frac{(1 + e^{g_{n-1} + g_n + \varepsilon_n})(1 + e^{\varepsilon_n})}{(1 + e^{g_n + \varepsilon_n})(1 + e^{g_{n-1} + \varepsilon_n})} \right\}, \\ \tilde{\varepsilon}_{n-1} &= \varepsilon_{n-1} + \log \left\{ \frac{(1 + e^{g_{n-1} + \varepsilon_n})}{(1 + e^{\varepsilon_n})} \right\}, \\ \tilde{\varepsilon}_{n+1} &= \varepsilon_{n+1} + \log \left\{ \frac{(1 + e^{g_n + \varepsilon_n})}{(1 + e^{\varepsilon_n})} \right\}, \end{aligned} \quad (5.13)$$

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:

$$\begin{aligned} \tilde{g}_n &= \log \left\{ \frac{(1 + e^{g_n + g_{n+1} + \varepsilon_{n+1}})(1 + e^{\varepsilon_{n+1}})}{(1 + e^{g_n + \varepsilon_{n+1}})(1 + e^{g_{n+1} + \varepsilon_{n+1}})} \right\}, \\ \tilde{\varepsilon}_n &= \varepsilon_n + \log \left\{ \frac{(1 + e^{g_n + \varepsilon_{n+1}})(1 + e^{g_{n-1} + \varepsilon_{n-1}})}{(1 + e^{\varepsilon_{n+1}})(1 + e^{\varepsilon_{n-1}})} \right\} \end{aligned} \quad (5.14)$$

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. [128]), or in our accompanying paper [127].

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

⁷Up 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.

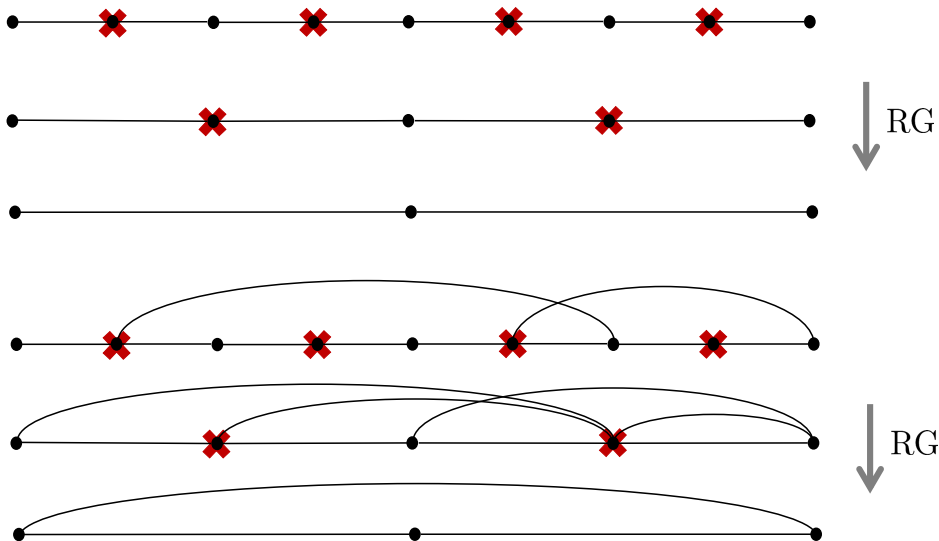


Figure 5.1: Schematic rendering of maximum coordination two couplings (top) and a single bare coordination number three coupling (bottom) under RG transformations. The red crosses indicate the sites that are decimated at each RG step. More generally, successive RG steps increase the maximum coordination number, and flow towards all-to-all couplings (bottom most diagram).

an overall rescaling of the partition function – results from the fact that Eq. 5.11 represents bilinear interactions of maximum coordination number two in the line graph representation. Consequently, the associated Hamiltonian Eq. 5.12 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 heterogenous 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. 5.12 with the following coordination number three interaction:

$$H \supset [\varepsilon_m + g_m A_{m+1} + g'_m A_{m+k}] A_a \quad (5.15)$$

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. 5.1.

§5.2.1 Maximum coordination number two $\equiv$ 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 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 [129, 130], variational, or mean field methods (see e.g. [131] for an overview)⁸. 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 infra-red, 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.

§5.3 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. 5.12 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 [127]). In this case, Eqs. 5.14 reduce to the following system of equations:

$$\begin{aligned}\tilde{g} &= \log \left\{ \frac{(1 + e^{2g+\varepsilon})(1 + e^\varepsilon)}{(1 + e^{g+\varepsilon})^2} \right\}, \\ \tilde{\varepsilon} &= \varepsilon + 2 \log \left\{ \frac{(1 + e^{g+\varepsilon})}{(1 + e^\varepsilon)} \right\},\end{aligned}\tag{5.16}$$

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

⁸Caveating the availability of exact solutions in the specific case of $d = 2$ in the absence of disorder [132, 133].

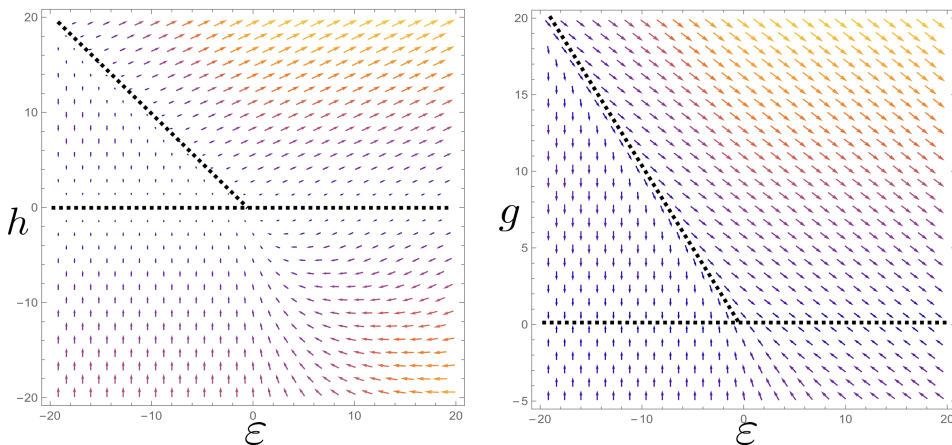


Figure 5.2: RG flow for homogeneous coordination number two interactions towards an Erdős-Rényi fixed line, where arrows indicate direction of flow and color determine the its intensity, lighter being stronger. Upper panel: A shearing behavior for positively correlated conditioning is apparent, along 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 by the bottom panel.

given by:

$$\begin{aligned}\beta_g(g, \varepsilon) &= \log \left\{ \frac{(1 + e^{2g+\varepsilon})(1 + e^\varepsilon)}{e^g(1 + e^{g+\varepsilon})^2} \right\}, \\ \beta_\varepsilon(g, \varepsilon) &= 2 \log \left\{ \frac{(1 + e^{g+\varepsilon})}{(1 + e^\varepsilon)} \right\}.\end{aligned}\tag{5.17}$$

The RG flow corresponding to these beta functions is depicted in Fig. 5.2, 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. 5.16 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 anti-ferromagnetic couplings. When one considers the effect of marginalizing over two adjacent anti-ferromagnetic 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. 5.12 and making the transformation to the spin variables

$$A_a = \frac{1 + 2\sigma_a}{2}, \quad \sigma_a = \pm \frac{1}{2}, \quad (5.18)$$

the result will be a Hamiltonian of the form:

$$H(\sigma) = - \sum_{a=1}^{\mathcal{N}} g_a \sigma_a \sigma_{a+1} - \sum_{a=1}^{\mathcal{N}} \sigma_a \left(\varepsilon_a + \frac{g_a + g_{a+1}}{2} \right) \quad (5.19)$$

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 plane⁹. 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 be to draw these from some probability distribution function (PDF), which is to introduce disorder on the line graph.

⁹This 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 $\varepsilon = -g$.

§5.4 Random disorder and induced probability flows

A useful reformulation of the RG transformations Eqs. 5.14 is in terms of the variables $y_n := e^{g_n}$ and $x_n := e^{\varepsilon_n}$, in which they take an algebraic form:

$$\begin{aligned}\tilde{y}_n &= \frac{(1 + y_n y_{n+1} x_{n+1})(1 + x_{n+1})}{(1 + y_n x_{n+1})(1 + y_{n+1} x_{n+1})}, \\ \tilde{x}_n &= x_n \frac{(1 + y_n x_{n+1})(1 + y_{n-1} x_{n-1})}{(1 + x_{n+1})(1 + x_{n-1})}.\end{aligned}\tag{5.20}$$

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 site-wise 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

$$\begin{aligned}\tilde{\psi}_n(\tilde{y}_n \equiv y) &= \int_0^\infty \dots \int_0^\infty dx_{n+1} dy_n dy_{n+1} \psi_n(y_n) \\ &\times \psi_{n+1}(y_{n+1}) \chi_{n+1}(x_{n+1}) \delta(\tilde{y}_n - y),\end{aligned}\tag{5.21}$$

and

$$\begin{aligned}\tilde{\chi}_n(\tilde{x}_n \equiv x) &= \int_0^\infty \dots \int_0^\infty dx_{n-1} dx_n dx_{n+1} dy_{n-1} dy_n \delta(\tilde{x}_n - x) \\ &\times \psi_{n-1}(y_{n-1}) \psi_n(y_n) \chi_{n-1}(x_{n-1}) \chi_n(x_n) \chi_{n+1}(x_{n+1}),\end{aligned}\tag{5.22}$$

where the dependent variables $\tilde{y}_n$ and $\tilde{x}_n$ that appear in the arguments of the delta functions are given by Eqs. 5.20. 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. 5.21, 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. 5.20 in terms of the variable we look to transform, which we choose to be y_n . The result of this inversion is given by

$$y_n = \frac{(1 + x_{n+1}) - (1 + y_{n+1} x_{n+1}) \tilde{y}_n}{[x_{n+1}(1 + y_{n+1} x_{n+1}) \tilde{y}_n - (1 + x_{n+1}) y_{n+1} x_{n+1}]},\tag{5.23}$$

so that after accounting for the Jacobian factor and performing the subsequent $\tilde{y}_n$

integral, one ends up with

$$\begin{aligned} \tilde{\psi}_n(y) = \int_0^\infty \int_0^\infty dy_{n+1} dx_{n+1} & \left\{ \Theta(y_n) \psi_n(y_n[x_{n+1}, y_{n+1}, y]) \right. \\ & \left. \frac{\psi_{n+1}(y_{n+1}) \chi_{n+1}(x_{n+1}) (1+x_{n+1}) (1+y_{n+1}x_{n+1}) |y_{n+1}-1|}{x_{n+1} [(1+y_{n+1}x_{n+1})y - (1+x_{n+1})y_{n+1}]^2} \right\}, \end{aligned} \quad (5.24)$$

where the Heaviside step function enforces the domain of integration to be such that Eq. 5.23 is positive for $y = \tilde{y}_n$. Proceeding analogously for Eq. 5.22, one obtains

$$\begin{aligned} \tilde{\chi}_n(\tilde{x}_n \equiv x) = \int_0^\infty \dots \int_0^\infty dy_{n-1} dy_n dx_{n-1} dx_{n+1} & \left\{ \chi_n(x_n[\dots]) \right. \\ & \psi_{n-1}(y_{n-1}) \psi_n(y_n) \chi_{n-1}(x_{n-1}) \chi_{n+1}(x_{n+1}) \\ & \left. \times \frac{(1+x_{n+1})(1+x_{n-1})}{(1+y_n x_{n+1})(1+y_{n-1} x_{n-1})} \right\}, \end{aligned} \quad (5.25)$$

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

$$x_n = \tilde{x}_n \frac{(1+x_{n+1})(1+x_{n-1})}{(1+y_n x_{n+1})(1+y_{n-1} x_{n-1})}. \quad (5.26)$$

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:

$$\begin{aligned} \psi(y) = P(g[y]) \frac{dg}{dy} & \equiv \frac{P(g[y])}{y}, \\ \chi(x) = C(\varepsilon[x]) \frac{d\varepsilon}{dx} & \equiv \frac{C(\varepsilon[x])}{x}, \end{aligned} \quad (5.27)$$

where we have used that $g = \log y$ and $\varepsilon = \log x$. Repeating the same steps that led to Eqs. 5.24 and 5.25, we obtain the induced RG flow for $P(g)$ and $C(\varepsilon)$ as

$$\begin{aligned} \tilde{P}_n(\tilde{g}_n \equiv g) = \int_{-\infty}^\infty \int_{-\infty}^\infty dg_{n+1} d\varepsilon_{n+1} & \left\{ \Theta(g_n) P_{n+1}(g_{n+1}) \right. \\ & \frac{e^g (1+e^{\varepsilon_{n+1}}) (1+e^{g_{n+1}+\varepsilon_{n+1}}) |e^{g_{n+1}}-1|}{e^{\varepsilon_{n+1}} [(1+e^{g_{n+1}+\varepsilon_{n+1}})e^g - (1+e^{\varepsilon_{n+1}})e^{g_{n+1}}]^2} \\ & \left. \times C_{n+1}(\varepsilon_{n+1}) P_n(g_n[\varepsilon_{n+1}, g_{n+1}, \tilde{g}_n \equiv g]) \right\}, \end{aligned} \quad (5.28)$$

where g_n in the argument of P_n above is obtained by taking the log of Eq. 5.23 and substituting the definitions for the y and x variables. Similarly, one finds:

$$\begin{aligned} \tilde{C}_n(\tilde{\varepsilon}_n \equiv \varepsilon) = \int_{-\infty}^\infty \dots \int_{-\infty}^\infty dg_{n-1} dg_n d\varepsilon_{n-1} d\varepsilon_{n+1} & \left\{ P_n(g_n) \right. \\ & P_{n-1}(g_{n-1}) C_{n-1}(\varepsilon_{n-1}) \frac{(1+e^{\varepsilon_{n+1}})(1+e^{\varepsilon_{n-1}})}{(1+e^{g_n+\varepsilon_{n+1}})(1+e^{g_{n-1}+\varepsilon_{n-1}})} \\ & \left. \times C_{n+1}(\varepsilon_{n+1}) C_n(\varepsilon_n[g_{n-1}, g_n, \varepsilon_{n-1}, \varepsilon_{n+1}, \tilde{\varepsilon}_n \equiv \varepsilon]) \right\}. \end{aligned} \quad (5.29)$$

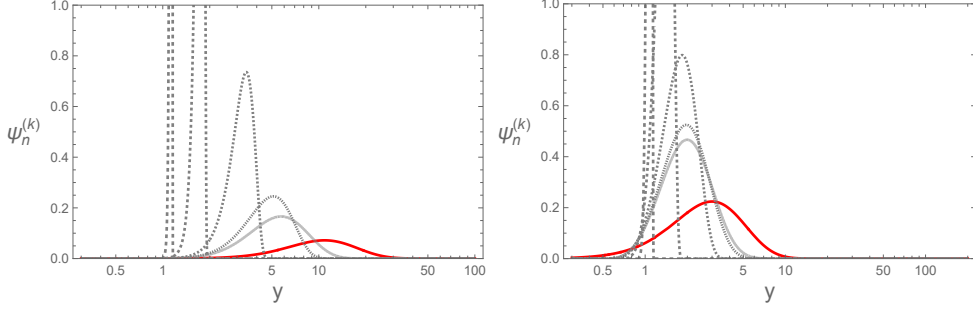


Figure 5.3: Top panel: Renormalization group induced deformation of the bare probability assignment $\Pi(y)$ given by Eq. 5.35 for $\alpha_\pi = 5$, $\lambda_\pi = e^{-1}$ (red line) for k from 1 to 5 (gray dashed lines with successively coarser dasheding). Bottom panel: The same, but for $\Pi(y)$ given with parameters $\alpha_\pi = 4$, $\lambda_\pi = e^0$.

Although the convolutions in both sets of Eqs. 5.24, 5.25 and Eqs. 5.28, 5.29 are amenable to numerical integration for arbitrary probability assignments, it is informative to consider certain special cases.

§5.4.1 Sparse disorder

The simplest non-trivial 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:

$$\begin{aligned} \{\chi_a^{(0)}(x_a)\}_{a \neq n} &= \delta(x_a - x^{(0)}), \quad \chi_n^{(0)}(x_n) = \Omega(x_n) \\ \{\psi_a^{(0)}(y_a)\}_{a \neq n} &= \delta(y_a - y^{(0)}), \quad \psi_n^{(0)}(y_n) = \Pi(y_n) \end{aligned} \quad (5.30)$$

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. 5.24 and 5.25 yield the following induced renormalization group transformations on the probability assignments upon iteration:

$$\begin{aligned} \chi_n^{(k+1)}(x) &= \int_0^\infty dy \left\{ \frac{\psi_n^{(k)}(y) (1 + x^{(k)})^2}{(1 + yx^{(k)}) (1 + y^{(k)}x^{(k)})} \right. \\ &\quad \left. \times \chi_n^{(k)} \left[\frac{x (1 + x^{(k)})^2}{(1 + yx^{(k)}) (1 + y^{(k)}x^{(k)})} \right] \right\}, \end{aligned} \quad (5.31)$$

$$\begin{aligned} \psi_n^{(k+1)}(y) &= \frac{(1 + x^{(k)})|y^{(k)} - 1|(1 + x^{(k)}y^{(k)})}{x^{(k)}(y + y^{(k)})[x^{(k)}(y - 1) - 1]^2} \\ &\quad \times \psi_n^{(k)} \left[\frac{1 + x^{(k)} - y(1 + x^{(k)}y^{(k)})}{x^{(k)}(y[1 + x^{(k)}y^{(k)}] - y^{(k)}[1 + x^{(k)}])} \right]_+, \end{aligned} \quad (5.32)$$

where dummy variables have been relabeled for compactness, and where

$$\begin{aligned}\{\chi_a^{(k+1)}(x)\}_{a \neq n} &= \delta(x - x^{(k+1)}), \\ \{\psi_a^{(k+1)}(y)\}_{a \neq n} &= \delta(y - y^{(k+1)}),\end{aligned}\tag{5.33}$$

where the arguments of the delta functions $x^{(0)}$ and $y^{(0)}$ in Eq. 5.30 get recursively renormalized to $x^{(k)}$ and $y^{(k)}$ via Eqs. 5.20 as:

$$\begin{aligned}y^{(k+1)} &= \frac{(1 + y^{(k)^2} x^{(k)}) (1 + x^{(k)})}{(1 + y^{(k)} x^{(k)})^2}, \\ x^{(k+1)} &= x^{(k)} \frac{(1 + y^{(k)} x^{(k)})^2}{(1 + x^{(k)})^2}.\end{aligned}\tag{5.34}$$

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. 5.32. The subscript on Eq. 5.32 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

$$\begin{aligned}\Pi(y) &= \frac{\lambda_\pi^{\alpha_\pi}}{\Gamma(\alpha_\pi)} y^{\alpha_\pi-1} e^{-\lambda_\pi y}, \\ \Omega(x) &= \frac{\lambda_\omega^{\alpha_\omega}}{\Gamma(\alpha_\omega)} x^{\alpha_\omega-1} e^{-\lambda_\omega x},\end{aligned}\tag{5.35}$$

where the mean and the variance, given by α/λ and α/λ^2 respectively, which one can specify arbitrarily. Note that the x variables relate to a probability assignment for a corresponding Erdős-Renyi random graph via the relation (cf. Eq. 5.5):

$$x = e^\varepsilon = \frac{p}{1-p}.\tag{5.36}$$

Upon repeated decimation transformations, we see that the disorder probability assignment would asymptote towards a delta function distribution centered around $y = 1$ in Fig. 5.3, corresponding to the $g = 0$ fixed line in Fig. 5.1. 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 sites¹⁰. The renormalization of the local probability assignments would proceed according to Eqs. 5.31 and 5.32 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 Eqs. 5.24, 5.25, iterating again from there with Eqs. 5.31, 5.32 until the next step where coalescence occurs (cf. [128]).

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

§5.4.2 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. 5.24 and 5.25 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. 5.20 transform to:

$$\begin{aligned}\tilde{z}_n &= \frac{(x_n + z_n z_{n+1})(1 + z_{n-1})}{(1 + x_{n-1})(1 + z_{n+1})}, \\ \tilde{x}_n &= \frac{(x_n + z_n x_{n+1})(1 + z_{n-1})}{(1 + x_{n+1})(1 + x_{n-1})}.\end{aligned}\quad (5.37)$$

We now presume the disorder to be uniformly assigned as:

$$\{\chi_n^{(0)}(x_n)\}_{\forall n} = \chi(x_n), \quad \{\psi_n^{(0)}(z_n)\}_{\forall n} = \psi(z_n), \quad (5.38)$$

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. 5.24 and 5.25, one arrives at the following induced RG transformations for the probability assignments at any given site¹¹:

$$\begin{aligned}\chi^{(k+1)}(x) &= \int_0^\infty \dots \int_0^\infty du dv dy dw \left\{ \frac{(1+w)(1+v)}{(y+uw)} \right. \\ &\quad \left. \psi^{(k)} \left[\frac{x(1+w)(1+v)}{(y+wu)} - 1 \right]_+ \psi^{(k)}(u) \chi^{(k)}(v) \chi^{(k)}(y) \chi^{(k)}(w) \right\},\end{aligned}\quad (5.39)$$

and

$$\begin{aligned}\psi^{(k+1)}(z) &= \int_0^\infty \dots \int_0^\infty du dv dy dw \left\{ \psi^{(k)}(v) \psi^{(k)}(y) \psi^{(k)}(w) \right. \\ &\quad \left. \frac{(u+yw)(1+v)}{z^2(1+w)} \chi^{(k)} \left[\frac{(u+yw)(1+v)}{z(1+w)} - 1 \right]_+ \chi^{(k)}(u) \right\},\end{aligned}\quad (5.40)$$

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 non-linear 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 in some

¹¹Here, and in all analogous derivations that follow, one writes down the analogous equations to Eqs. 5.21 and 5.22, identifies a convenient integration variable which admits a unique inverse for the RG transformations for this variable alone analogous to Eqs. 5.20, and performs the delta function integration accounting for the appropriate Jacobian factor and integration limits.

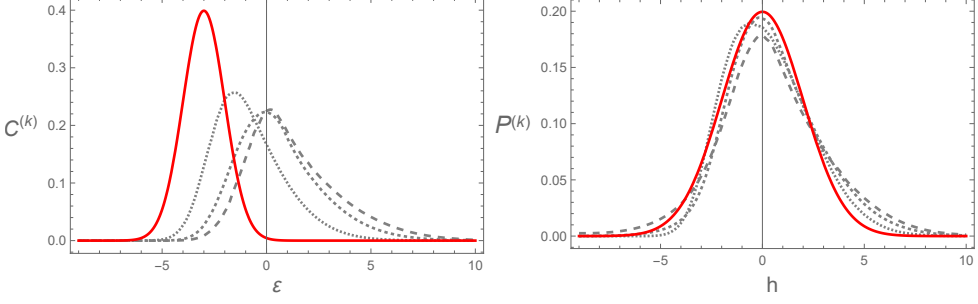


Figure 5.4: Renormalization group induced deformation of the uniform probability assignments $C^{(k)}(\varepsilon)$ (top) and $P^{(k)}(h)$ (bottom). Bold red lines denote the bare probability assignments, taken to be Gaussians with means and variances given by $\mu_\varepsilon = -3, \sigma_\varepsilon = 1$ 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 dasheding.

cases by noticing that both convolutions can be expressed as

$$\begin{aligned} \chi^{(k+1)}(x) = \frac{d}{dx} \int_0^\infty \dots \int_0^\infty du dv dy dw & \left\{ \Psi^{(k)} [f(x|u, v, y, w)] \right. \\ & \left. \times \psi^{(k)}(u) \chi^{(k)}(v) \chi^{(k)}(y) \chi^{(k)}(w) \right\}, \end{aligned} \quad (5.41)$$

$$\begin{aligned} \psi^{(k+1)}(z) = -\frac{d}{dz} \int_0^\infty \dots \int_0^\infty du dv dy dw & \left\{ \Xi^{(k)} [g(z|u, v, y, w)] \right. \\ & \left. \times \chi^{(k)}(u) \psi^{(k)}(v) \psi^{(k)}(y) \psi^{(k)}(w) \right\}. \end{aligned} \quad (5.42)$$

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 Ψ where the integrand has no support, where $\Xi^{(k)}(s)$ is defined analogously, and where the respective functional arguments are defined as:

$$\begin{aligned} f(x|u, v, y, w) &:= \frac{x(1+w)(1+v)}{(y+uw)} - 1, \\ g(z|u, v, y, w) &:= \frac{(u+yw)(1+v)}{z(1+w)} - 1. \end{aligned} \quad (5.43)$$

Eq. 5.41 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 ε and the joint distribution

for $h = \varepsilon + g$, denoted $C^{(k)}(\varepsilon)$ and $P^{(k)}(h)$ respectively:

$$\begin{aligned}
 C^{(k+1)}(\varepsilon) = & \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} du dv dy dw \left\{ \left[1 - \frac{(e^y + e^{u+w})}{e^\varepsilon (1 + e^w)(1 + e^v)} \right]_+^{-1} \right. \\
 & \times P^{(k)} \left[\log \left(\frac{e^\varepsilon (1 + e^w)(1 + e^v)}{(e^y + e^{w+u})} - 1 \right) \right] \\
 & \left. \times P^{(k)}(u) C^{(k)}(v) C^{(k)}(y) C^{(k)}(w) \right\}, \tag{5.44}
 \end{aligned}$$

and

$$\begin{aligned}
 P^{(k+1)}(h) = & \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} du dv dy dw \left\{ \left[1 - \frac{(1 + e^w)e^h}{(e^u + e^{y+w})(1 + e^v)} \right]_+^{-1} \right. \\
 & \times C^{(k)} \left[\log \left(\frac{(e^u + e^{y+w})(1 + e^v)}{(1 + e^w)e^h} - 1 \right) \right] \\
 & \left. \times C^{(k)}(u) P^{(k)}(v) P^{(k)}(y) P^{(k)}(w) \right\}, \tag{5.45}
 \end{aligned}$$

where the subscripts now on the Jacobian factors again restrict the integration domain to be such that the respective factors are both positive. As before, one can recast both convolutions as an analogous gradient flow:

$$\begin{aligned}
 C^{(k+1)}(\varepsilon) = \frac{d}{d\varepsilon} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} du dv dy dw \left\{ \Pi^{(k)} [F(\varepsilon|u, v, y, w)] \right. \\
 \left. \times P^{(k)}(u) C^{(k)}(v) C^{(k)}(y) C^{(k)}(w) \right\}, \tag{5.46}
 \end{aligned}$$

$$\begin{aligned}
 P^{(k+1)}(h) = -\frac{d}{dh} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} du dv dy dw \left\{ \Xi^{(k)} [G(h|u, v, y, w)] \right. \\
 \left. \times C^{(k)}(u) P^{(k)}(v) P^{(k)}(y) P^{(k)}(w) \right\}, \tag{5.47}
 \end{aligned}$$

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:

$$\begin{aligned}
 F(\varepsilon|u, v, y, w) &:= \log \left\{ \frac{e^\varepsilon (1 + e^w)(1 + e^v)}{(e^y + e^{w+u})} - 1 \right\}, \tag{5.48} \\
 G(h|u, v, y, w) &:= \log \left\{ \frac{(e^u + e^{y+w})(1 + e^v)}{(1 + e^w)e^h} - 1 \right\}.
 \end{aligned}$$

We illustrate the induced RG flow on the bare uniform disorder probability assignments in Fig. 5.4. Visual inspection of both Figs. 5.3 and 5.4 seems to suggest the induced renormalization group flow to be akin to some sort of active transport process for the disorder probability assignments.

§5.5 Implications and applications

Over the preceding chapters, we have shown how restricting, or truncating the effective Hamiltonian to quadratic order (cf. Eq. 5.2) 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:

- Disordered Erdős-Renyi random graphs with pairwise conditioning of the individual links admit closed form renormalization group transformations for maximum coordination number two.
- Repeated RG transformations render these interactions irrelevant, flowing towards a disordered non-interacting Erdős-Renyi random graph.
- 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.
- 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.

§5.5.1 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 an variety of interesting network structures and topologies. Preferential attachment models [?] 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. [4, 134] 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:

$$H(\mathbf{A}) \supset A_{ik} \sum_{j \neq k} A_{ij}. \quad (5.49)$$

Such terms would correspond to terms in an effective Hamiltonian of the general form Eq. 5.10 as opposed to Eq. 5.12 on the corresponding line graph representation. Non-linear 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

$$H(\mathbf{A}) \supset A_{ik} \left(\sum_{j \neq k} A_{ij} \right)^n \quad (5.50)$$

that depend on the Taylor expansion of the kernel function $f(k_i)$, for which the bilinear truncation Eq. 5.10 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 Eq. 5.50 can always be brought into the form Eq. 5.49, 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 Eq. 5.50 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. 5.5 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 link¹². 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. 5.12, 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. 5.6, caveating 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. 5.6. We can also realize preferential *avoidance* via negative values for the bilinear couplings that index

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



Figure 5.5: 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.

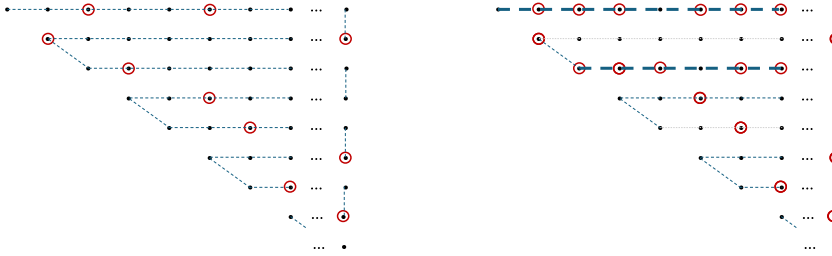


Figure 5.6: Representation of a random graph with maximum coordination two interactions (top panel). Enhanced or diminished attachment and preference effects can be implemented via the sign and magnitude of the relevant bilinear couplings $\{g_a\}$ (bottom panel: bold and faint lines, respectively).

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-Renyi probability assignments, i.e. the set of $\{\varepsilon_a\}$, to be arranged so that the desired distribution of degrees is attained, as implemented in [65, 135]. 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.

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 $\tilde{N} = N/2$ nodes¹³. 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. 5.1, 5.3, and 5.4, all of which illustrate the progressive 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 section 5.3. 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-Renyi 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 [136, 137, 138, 139]. One variant of particular utility is a random graph model of Molloy and Reed, which is constructed to reproduce an arbitrarily specified degree distribution [140], which nevertheless still comes up short in modeling the distance distributions observed in number of data sets of a given degree distribution [?]. 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.

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

In the context of social or contact networks, Fig. 5.1 offers a useful mnemonic for the scaling behavior of individual couplings encoded in Eq. 5.14 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 Eq. 5.7. 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 [141, 142, 143]. Similarly, reinforcement effects at the cortical level in biological and artificial neural networks [147, 148, 149, 150], 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.

§5.5.2 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 [41, 151] 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, 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 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 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 [127], 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, sub-manifolds 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’ sub-manifold 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 [127], where we furthermore apply this formalism to economic and trade data of limited resolution.

§5.6 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 Eq. 5.2, 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 Eq. 5.2 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 [127].

