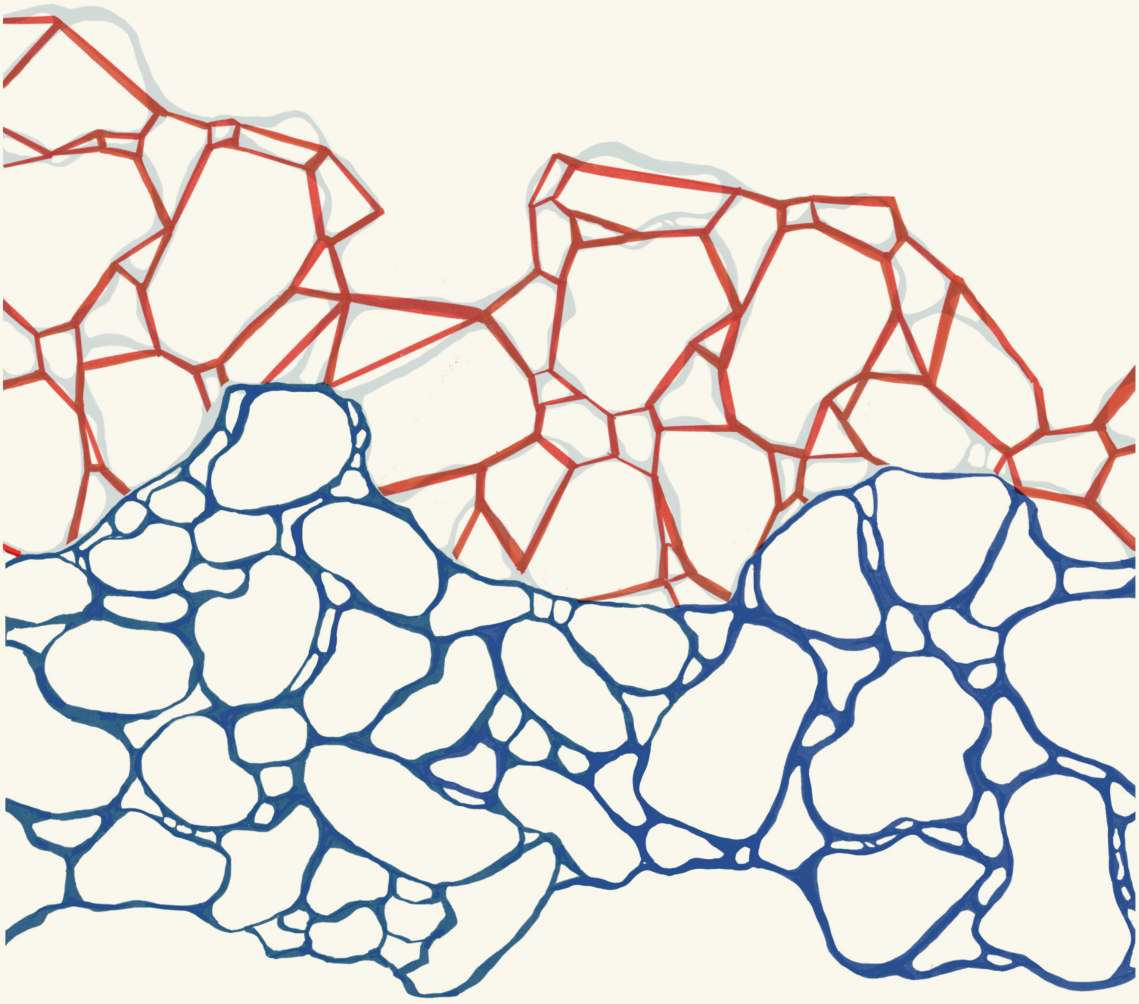
On the Renormalization of Random Network Models



Alessio Catanzaro

Stellingen

Behorende bij het proefschrift

On the Renormalization of Random Network Models

1. The understanding of renormalization for complex networks as a scaling operation akin to a magnifying lens can be complemented with the perspective of the same network being unobservable in some of its constituents. Which information is lost under this operation is often neglected, but can actually help the inference over the unattainable parts.

[Chapter 4]

2. Self-averaging of network models, and hence the assumption that a single network is representative of the whole probabilistic ensemble it was generated from, is not always true and therefore should not be taken for granted during network inference.

[Chapter 3]

3. There exists only one class of interactions between edges in a network that is closed under the flow of renormalization, namely the coupling structure with maximum coordination number equal to 2.

[Chapter 5]

4. Second order random network models can be solved in general and their partition functions computed.

[Chapter 6]

5. Embedding the nodes of a network model with independent edges in a latent, geometrical space is not a necessary condition for the emergence of a non-trivial local clustering coefficient.

[Chapter 3, Aliakbarisani, Boguñá, Serrano, Phys. Rev. Lett. 135, 197402, (2025)]

6. Log-periodicity in the eigenmodes of an adjacency matrix is a distinctive feature of discrete

scale-invariance, and can be found in the structure of eigenmodes of a self-similar network model.

[Chapter 2, Sornette Phys. Rep. 297, 239–270, (1998)]

7. There are graph models reproducing real network properties that defy the low-rank hypothesis of complex systems by having an infinite, although sublinear, rank and a spectral bulk edge growing with the same scaling of the signal.

[Chapter 2; Thibeault, Allard, Desrosiers, Nat. Phys. 20, 294—302, (2024)]

8. Irreducible higher-order correlations naturally emerge from renormalization over unobservable sections of an interacting system. One can mathematically deal with them in the paradigm of effective field theories, but to justify their claimed irreducible nature solid empirical evidence of mechanisms beyond pairwise coupling should be put forward.

[Chapter 4; Battiston et al., Nat. Phys. 17, 1093–1098 (2021); Peixoto et al., arXiv:2602.16937 (2026)]

9. "Anarchism, while perhaps [...] the most attractive political philosophy, is certainly excellent medicine for epistemology, and for the philosophy of science." However, Feyerabend's argument should not allow complex-systems scientists to abandon self-questioning regarding their own research's epistemological foundations.

[Feyerabend, Paul, 'Against Method: Outline of an Anarchistic Theory of Knowledge', London 1975]

10. The European Research Council and public research institutes in general should firmly uphold the ban on dual civilian-military use in research programmes.

Alessio Catanzaro

Leiden, 10 september 2026

On the Renormalization of Random Network Models

Alessio Catanzaro

Cover design by Dr. Beatrice Vaienti.

Permission to use the picture has been kindly granted by the artist.

On the Renormalization of Random Network Models

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof. dr. S. de Rijcke,
volgens besluit van het college voor promoties
te verdedigen op donderdag 10 september 2026
klokke 16:00 uur

door

Alessio Catanzaro
geboren te Bologna
in 1995

Promotor:
Prof. dr. D. Garlaschelli

Co-promotor:
Dr. S. Patil

Promotiecommissie:
Prof. dr. ir. S. J. van der Molen
Prof. dr. K. E. Schalm
Prof. V. Latora, *Queen Mary University & Università di Catania*
Prof. M. Á. Serrano Moral, *Universitat de Barcelona & ICREA*
Dr. C. Stegehuis, *Universiteit Twente*

Contents

1 Introduction **11**
Introduction to Part I 12
Introduction to Part II 15

I Multi-Scale Network Model in the infinite renormalizability regime: Clustering and Spectrum **19**

2 Spectra of random graphs with discrete scale invariance **21**
§2.1 Introduction 22
§2.2 Preliminaries 23
§2.2.1 Realized versus expected matrix. 23
§2.3 Leading eigenmodes from the expected matrix. 25
§2.3.1 Largest eigenvalue and principal eigenvector 25
§2.3.2 Next eigenvalues 28
§2.3.3 Next Eigenvectors 35
§2.4 Bulk of the Spectrum 36
§2.5 Concluding remarks 41

3 Clustering without Geometry in an Independent Edges Network Model **43**
§3.1 Introduction: clustering versus geometry. 44
§3.2 The Multi-Scale Model. 46
§3.3 The Local Clustering Coefficient. 46
§3.3.1 Clustering Function. 48
§3.3.2 Limiting expression for local clustering function 48
§3.4 Hub and leaf behaviour. 52
§3.4.1 Asymptotics of the clustering function 53
§3.5 Average Clustering Coefficient. 56
§3.5.1 Computation of Average Clustering Coefficient 57
§3.6 Lack of self-averaging. 59
§3.6.1 Computation of $r_{0/1}$ 59
§3.7 Lack of self-averaging in the fraction of next low-degree nodes 63

§3.8 Extension to stable weights . . . 63
§3.9 Concluding Remarks. . . . 69

II Renormalization of Interacting Random Network Models 75

4 Topological Renormalization for Random Network Models 77
§4.1 Introduction . . . 78
§4.2 Marginalization framework . . . 78
§4.2.1 Marginalization Flow for Exponential Random Graphs . . . 79
§4.2.2 What is different from 'usual' Renormalization . . . 81
§4.2.3 The Generator of Missing Information . . . 81
§4.3 RG flows for Random Network Models . . . 82
§4.3.1 Approximated and Exact closure under RG . . . 84
§4.4 Towards applications to Network Inference . . . 86
§4.4.1 Temporal Network reconstruction at higher resolution . . . 86
§4.4.2 All-to-all and multiplex networks . . . 91
§4.5 Final remarks . . . 95

5 Renormalization of 1-D Interacting Random Graph 97
§5.1 Preliminary remarks . . . 98
§5.1.1 The (effective) generalized statistical mechanics of random graphs 100
§5.1.2 Universality classes of graph Hamiltonians . . . 102
§5.1.3 Outline . . . 103
§5.2 Line graphs, spin chains, and closed RG transformations . . . 104
§5.2.1 Maximum coordination number two ≡ disordered spin chain . 107
§5.3 Homogeneous RG flow . . . 107
§5.4 Random disorder and induced probability flows . . . 110
§5.4.1 Sparse disorder . . . 112
§5.4.2 Uniform disorder . . . 114
§5.5 Implications and applications . . . 117
§5.5.1 Peer, preferential, susceptibility, and reinforcement effects on random networks . . . 117
§5.5.2 Network inference with limited data . . . 122
§5.6 Concluding remarks and future directions . . . 123

6 The Harmonic Oscillator of Random Graphs 125
§6.1 Introduction . . . 126
§6.2 Solution to II order Random Network . . . 127
§6.2.1 Evaluation of the partition function . . . 128
§6.2.2 Effective Separability of the partition function . . . 131
§6.2.3 Phase structure . . . 133
§6.3 II order interacting models . . . 134

§6.3.1 RNM with Homogeneous interactions 134
§6.3.2 A square-like interaction 138
§6.3.3 Community Induced Interactions 139
§6.3.4 Node mediated interaction 145
§6.4 On the Mean Field Approximation and Phase Structure 152
§6.5 Final Remarks 155

Bibliography **157**

Summary **171**

Samenvatting **173**

Acknowledgements **175**

Curriculum Vitae **177**

CHAPTER 1
Introduction

This thesis deals with the broad topic of renormalization of complex networks, with a particular focus on the topological nature of such models of complex systems. It is divided into two main parts.

In the first part, we discuss the characterization of the spectrum and the clustering function of a recently proposed network model with independent edges. This model can be regarded as the 'fixed point' of a topology-preserving aggregation process that allows to deal with multi-scale clustered data. We investigate its properties in the regime where the node features are distributed according to an infinite-mean stable probability density function, making the model scale-invariant both in the functional form of the link probability and in the hidden variable distribution. In particular, in Chapter 2 we analyze the model spectral properties, while in Chapter 3 we discuss its clustering function.

In the second part, we elaborate on a specific renormalization scheme that is equivalent to marginalization over disordered uncertainty, in the context of random network models that might display a correlation (possibly higher-order) among links. In particular in Chapter 4 we introduce the marginalization framework, and show two classes of effectively solvable models exist. In Chapter 5 we analyze in detail the first class of models that become solvable under the marginalization flow, and elaborate on its applications to temporal networks. Finally, in Chapter 6 we tackle the second class of those models, by providing a general solution framework for II order Random Network Models.

Introduction to Part I

The first part of this thesis is concerned with the mathematical analysis of a recently introduced model of random networks, the *Multi-Scale Model* (MSM) [1, 2, 3], and in particular with two of its structural properties: the spectrum of its adjacency matrix and its clustering function behaviour. Before presenting the technical results, we offer here a wider perspective on the motivations for studying this model and on the significance of the findings collected in Chapters 2 and 3.

Why the Multi-Scale Model?

The scientific study of networks, from the Internet and social platforms to biological regulatory systems and financial markets, has produced an increasingly sophisticated catalogue of features that characterise real-world, complex networks [4, 5, 6]. Among the most robust empirical observations, two stand out for their ubiquity across completely different domains. First, real networks are *heterogeneous*: a relatively small number of highly connected nodes (hubs) coexist with a very large number of poorly connected ones, and the probability of observing a node with a given number of connections falls off as a power law with an exponent typically close to two [4, 5, 7]. Second, real networks are *clustered*: two nodes that share a common neighbour are much more likely to be directly connected to each other than two randomly chosen nodes [4, 8, 9]. This second feature, sometimes described by saying that "the friend of my friend is also my friend", is known as the clustering effect, and it is responsible for the tight-knit community structure observable in, for instance, social or collaboration networks.

Reproducing both of these properties simultaneously within a single, tractable random graph model has proven surprisingly elusive. Classical random graph models, such as the Erdős–Rényi model [10], connect nodes independently and with equal probability, thus capturing neither heterogeneity nor clustering. Hidden-variable models - also called fitness models [11, 12, 13] - address the heterogeneity problem by assigning each node a numerical weight (a fitness, or hidden variable) drawn from a heavy-tailed distribution, and connecting pairs of nodes with a probability that depends on their respective weights. This framework successfully reproduces the power-law degree distribution observed in real networks [5, 7, 11, 12]. However, models in which edges are drawn independently given the node fitnesses generically produce vanishingly small clustering as the network grows, unless nodes are additionally embedded in some underlying geometric space [14, 15, 16, 17, 18, 19]. The price of geometry is conceptual and practical: one must postulate the existence of a latent metric space, choose its dimension and curvature, and accept that many network properties are then determined by that geometric structure rather than emerging from the interaction rules alone [16, 20, 21].

The Multi-Scale Model and its renormalization origin

The model studied in this first part of the thesis takes a completely different starting point. Rather than assuming geometry, it asks: which random graph model is *self-similar under aggregation*? More concretely, if one groups nodes into "supernodes" - each supernode inheriting the combined fitnesses of its members - and declares two supernodes to be connected whenever any of their constituent nodes are connected, which model remains, in functional form, unchanged by this coarse-graining operation? This is a renormalization group question, phrased in the language of random networks, and the non-trivial answer is essentially unique for independent edges: it is the MSM [1, 2, 3], in which the connection probability between two nodes is a specific function of their fitnesses, and the fitness distribution belongs to the class of stable probability distributions with infinite mean [22].

The model can be regarded as a *fixed point* of a topology-preserving aggregation flow: iterating the coarse-graining procedure leaves the model connection probability invariant. This is a property shared by critical systems in statistical physics and it suggests that the MSM occupies a special, structurally stable position in the space of all possible random graph models.

From this perspective, understanding the MSM is not just an exercise in the analysis of a particular model; it is a step towards understanding what kinds of network structure are robust under changes of resolution or observation scale.

Spectral properties and discrete scale invariance

Chapter 2 is devoted to the spectral analysis of the MSM's adjacency matrix. The spectrum of a random graph - that is, the collection of eigenvalues and eigenvectors of its adjacency matrix - encodes an enormous amount of information about the network's structure [25]. Eigenvalues and eigenvectors are central to spectral clustering algorithms [26, 27, 28], to the study of diffusion and random walks on graphs, to the analysis of epidemic spreading, and to many methods of network inference.

For most random graph models with finite-mean fitnesses, the spectrum decomposes cleanly into two regions: a bulk, whose limiting distribution is described by results from classical random matrix theory [29, 30, 31], and a small number of isolated outlier eigenvalues that detach from the bulk and carry structural information about communities, hubs, and other macroscopic features [32, 33, 34]. The number of such outliers is typically small, independent of network size.

The MSM, by contrast, behaves in a radically different way. Because the fitness distribution has infinite mean, standard random matrix techniques break down entirely. What one finds instead is that an extensive number of eigenvalues - a number that increases (albeit slowly) with the size of the network - emerge from the bulk. These outlier eigenvalues are not scattered randomly in the complex plane: they actually lie at the intersections of the real axis with a logarithmic spiral, whose equation we give in full detail. The corresponding eigenvectors display complex-valued scaling exponents and log-periodic oscillations: their entries, when plotted as a function of the node fitness, do not follow a simple power law but instead oscillate (to preserve

orthogonality) in a pattern whose period is set by the logarithm of the network size.

This log-periodicity is the network signature of *discrete scale invariance*: the property that the system is invariant not under continuous rescaling, but only under rescaling by a specific factor (or its integer powers). Discrete scale invariance is known to arise in a variety of physical and biological systems displaying log-periodic corrections to scaling laws [35, 36, 37, 38], and its appearance here connects the network renormalization programme to a broader body of research on complex scaling phenomena.

Beyond the outlier eigenvalues, the bulk of the MSM spectrum is also unusual. Its edge is of the same order as the outlier eigenvalues - both scale as the square root of the number of nodes - unlike in models with finite-mean fitnesses, where the bulk is parametrically smaller than the largest outlier. This "bulk catching up" phenomenon makes it extremely hard to separate structural information from noise using a simple threshold on eigenvalue magnitude, and calls for new analytical strategies.

Clustering without geometry

Chapter 3 turns to the clustering properties of the MSM. As mentioned above, one of the long-standing criticisms of independent-edge network models is their inability to reproduce meaningful clustering [4, 14, 15]. The main result of Chapter 3 is a rigorous proof that the MSM, in the infinite-mean regime, is an exception to this rule: it naturally generates a non-trivial clustering function without any appeal to an underlying geometric space.

The mechanism is subtle. In the infinite-mean regime, the fitness of the nodes spans many orders of magnitude. Nodes with very high fitness - the hubs - are connected to virtually everyone else, so their neighbours form a group among which further connections are relatively rare: the hub's local clustering coefficient is close to zero. Nodes with low or moderate fitness, on the other hand, are connected almost exclusively to hubs; since leaf nodes with degree small but greater than one prevalently connect to hubs, and hubs are connected in a tight clique, the leaf local clustering coefficient approaches one. The overall result is a network that is, in a precise asymptotic sense, "almost everywhere" clustered, with only the vanishingly small fraction of nodes that are extreme hubs remaining unclustered. This is strikingly consistent with empirical observations in social and technological networks, where high-degree nodes tend to bridge otherwise weakly connected communities rather than being embedded in tightly knit cliques [8, 39, 40].

The analysis also reveals an interesting lack of self-averaging: the number of low degree nodes, when computed across different realisations of the same random graph model, exhibits significant fluctuations that do not vanish as the network grows. This is a consequence of the infinite-mean fitness distribution, which allows individual realisations to differ dramatically in the presence or absence of extreme hubs. The lack of self-averaging has practical implications for the statistical inference of network parameters from a single observed network.

Introduction to Part II

The second part of this thesis shifts the focus from the characterisation of a specific fixed-point model to the development of a general renormalization theory applicable to a broad class of random network models. The unifying thread is the idea of *marginalization*: when parts of a network are hidden, unresolved, or simply unobserved, the appropriate way to account for this ignorance is to average the model over the missing degrees of freedom, rather than to discard them or to treat them as fixed at some nominal value. This operation turns out to be precisely equivalent to a renormalization group (RG) transformation in the context of exponential random graph models. Part II develops this equivalence systematically and exploits it to derive exact and approximate solutions for several families of network models in which edges are not drawn independently.

Why correlations matter

The independent-edge assumption is a powerful simplification, and it underlies the tractability of many of the most widely used random graph models. However, real networks frequently display correlations among edges that cannot be captured by independent-edge models. In a social network, the probability that two individuals are connected is influenced not only by their individual attributes but also by the presence or absence of mutual acquaintances - a genuinely higher-order effect. In a financial network, the probability that two institutions share an exposure depends on the broader structure of counterparty relationships, creating systemic-risk profiles that independent-edge models systematically underestimate [17, 41, 42, 43]. In a biological network of gene regulation, the formation of feedback loops and motifs creates statistical dependencies among edges that are functionally significant [44, 45, 46].

Exponential random graph models (ERGMs) provide the most natural language for encoding such dependencies [47, 48]: they assign a probability to each network configuration in proportion to the exponential of a linear combination of graph statistics (edge counts, triangle counts, degree sequences, and so on), exactly as the Boltzmann weight in statistical mechanics assigns probabilities to microscopic configurations. Within this framework, independent-edge models correspond to Hamiltonians that contain only terms linear in individual edges, while correlated models involve Hamiltonians with higher-order terms coupling different edges.

The challenge with correlated models is their intractability: the normalisation constant (the partition function) is generically computationally hard to evaluate, and closed-form expressions for observables are available only in very special cases [49, 50]. The renormalization approach developed in Part II provides a systematic strategy to identify and exploit these special cases.

Marginalization as renormalization

The starting point of Chapter 4 is an observation that connects two apparently unrelated concepts. On one hand, in the theory of disordered spin systems (such as spin glasses), a standard technique for studying models on irregular or partially observed graphs is *decimation*: one integrates out a subset of the degrees of freedom, obtaining an effective model on the remaining sites [51, 52, 53]. On the other hand, in network science, the standard probabilistic approach is to write a probability distribution over the observable part of the network alone, and run the inference machinery on the reduced complete statistics.

The central insight of Chapter 4 is that these two operations are, in the context of exponential random graph models, complementary. Marginalising over unobserved network components is the same as performing a decimation step in the renormalization group sense. This equivalence is not merely formal: it allows one to import the entire conceptual machinery of the RG - fixed points, universality classes, flow equations, relevant and irrelevant directions - into the study of network models [54, 55, 56].

A crucial consequence of this equivalence is that marginalization generically introduces effective higher-order interactions among the remaining variables. If the original model has only pairwise edge couplings, integrating out a subset of edges will in general produce a model in which the probabilities of the remaining edges depend on each other in a complex way, through higher-order interaction terms. This is the network analogue of the well-known fact that coarse-graining a lattice model generates effective couplings between spins that were previously non-interacting [51]. Managing the proliferation of these higher-order terms is the central technical challenge of Part II, and Chapters 5 and 6 are devoted to two complementary strategies for dealing with it.

Exactly solvable models and the role of topology

The first strategy, developed in Chapter 5, is to identify network topologies and coupling structures for which the RG flow *closes exactly* within a finite-dimensional family of models. This means that, no matter how many marginalization steps one performs, the resulting effective model remains within the same parametric family, with only the numerical values of the coupling constants changing.

Chapter 5 provides a complete analysis of one such exactly solvable class: random network models defined on a one-dimensional chain. This is the network analogue of the Ising chain [58, 59, 60], and the exact solvability has the same origin: the locality of the interactions, which prevents the generation of long-range effective couplings under decimation. Building on this base case, Chapter 5 systematically analyses the effect of adding disorder - in the form of spatially varying coupling constants or link probabilities - to the chain model. Two types of disorder are considered: sparse disorder (affecting only a finite fraction of sites) and uniform disorder (affecting all sites in a statistically independent way). In both cases, the exact RG equations can be derived and the induced probability flows characterised analytically.

The exact solvability of the chain model is not merely a mathematical curiosity.

Chapter 5 shows that the RG equations for the chain can be mapped onto a variety of network phenomena of practical interest: peer effects (the tendency of individuals to adopt the behaviour of their network neighbours), preferential attachment (the rich-get-richer mechanism underlying the growth of many real networks), susceptibility effects (in which the formation of a connection depends on the current state of a node), and reinforcement learning dynamics. By understanding the RG flow of the chain model, one gains insight into the universality classes that govern all of these phenomena.

Second-order models and the harmonic oscillator of random graphs

The second strategy, developed in Chapter 6, takes a different approach. Rather than restricting to a specific topology for which the RG flow closes exactly, Chapter 6 studies the class of models to which most correlated network models flow under marginalization when higher-order interactions are irrelevant. In this case, the flow identifies a specific class of "second-order" models, characterised by pairwise interactions between *edges* (rather than between nodes). The name "second-order random network model" reflects this pairwise edge coupling structure, which is the simplest form of edge correlation.

The central result of Chapter 6 is that second-order random network models can be solved exactly and in full generality, for arbitrary interaction structures. The key observation is that the problem of computing the partition function of a second-order model can be recast as a problem in algebraic graph theory: specifically, as the diagonalization of the *link-adjacency matrix*, a matrix that encodes the coupling between pairs of edges rather than between nodes. The eigenmodes of this matrix determine the phase structure of the model, and the general single-link probability can be expressed as a function of these eigenmodes.

This structural insight has far-reaching consequences. It provides a unified framework for solving a large family of models that had previously been treated separately and in special cases [49, 50]. It reveals that the phase behaviour of correlated network models can be much richer than that of independent-edge models: depending on the eigenvalue structure of the link-adjacency matrix, a second-order model can exhibit multiple coexisting phases, complex phase diagrams, and transitions analogous to those found in condensed-matter physics. It also establishes a connection between network science and algebraic graph theory that has not previously been exploited in this context.

The title "harmonic oscillator of random graphs" reflects both the central role of the second-order coupling structure - analogous to the quadratic potential of a harmonic oscillator, which is the simplest exactly solvable problem in quantum mechanics - and the expectation that the methods developed in Chapter 6 will serve as a building block for the analysis of more complex, higher-order network models, much as the harmonic oscillator serves as a building block in perturbation theory.

PART I

MULTI-SCALE NETWORK MODEL IN THE INFINITE RENORMALIZABILITY REGIME: CLUSTERING AND SPECTRUM

CHAPTER 2

Spectra of random graphs with discrete scale invariance

This chapter is based on:
Alessio Catanzaro, Rajat Subhra Hazra, and Diego Garlaschelli. *"Spectra of random graphs with discrete scale invariance"*. arXiv preprint arXiv:2509.12407 (2025).

Abstract

Random graphs defined by an occurrence probability that is invariant under node aggregation have been identified recently in the context of network renormalization. The invariance property requires that edges are drawn with a specific probability that, in the annealed case, depends on a necessarily infinite-mean node fitness. The diverging mean determines many properties that are uncommon in models with independent edges, but at the same time widespread in real-world networks. Here we focus on the leading eigenvalues and eigenvectors of the adjacency matrix of the model, where the n nodes are assigned a Pareto (α)-distributed fitness with $0 < \alpha < 1$. We find that the leading eigenvalues are all of order $\sqrt{n}$, alternate in sign and are located at the intersection between the real axis and a logarithmic spiral in the complex plane, which we characterize analytically in terms of the Gamma function. We also calculate the associated eigenvectors, finding that they display complex-valued scaling exponents and log-periodicity, which are signatures of discrete scale invariance. In contrast with the typical finite-rank behaviour of random graphs with finite-mean variables, we find that a growing number of the leading eigenvalues emerges from the bulk, whose edge extends up to order $\sqrt{n}$ and therefore reaches the same scale as that of the structural eigenvalues.

§2.1 Introduction

Many models have been proposed to reproduce the heterogeneous structure ubiquitously observed in real-world networks. Among these, a wide class of random graph models assigns a weight (or fitness, or hidden variable) x_i to each vertex i ($i = 1, n$ where n is the number of nodes) from a usually fat-tailed distribution $\rho(x)$, and connects pairs of vertices i and j with a probability p_{ij} that is a certain positive function $f(x_i, x_j)$ [11, 12, 64, 65].

Recently, in the context of network renormalization [54], a specific such model that complements the heterogeneity of nodes with an invariance principle under graph coarse-graining has been identified [1, 2, 3, 66]. In this so-called (annealed) Multi-Scale Model (MSM), one may group nodes into equally sized 'supernodes', each inheriting the sum of the fitness of its members, and impose that both the connection probability $p_{ij} = f(x_i, x_j)$ and the fitness distribution $\rho(x)$ remain unchanged in their functional form, while their parameters follow exact renormalization rules. In other words, aggregating nodes into supernodes according to any homogeneous partition – and connecting two supernodes whenever there is a connection between any of their constituent nodes – defines a coarse-grained version of the graph, governed by an occurrence probability that has the same functional form as that of the original graph, up to a rescaling of the parameters [1, 3].

This renormalization scheme does not rely on any underlying geometry or distance notion between nodes. This stands in contrast to geometric renormalization approaches [20, 21], where the nodes to be merged need be spatially proximate, and a preliminary definition of metric distance is needed. Rather, when nodes are merged into supernodes of equal cardinality, the invariance requirement in the MSM is closer in spirit to the concept of discrete scale invariance [35], whose conequences for the graph properties have not been investigated so far and are one of our main focuses here.

In particular, we are interested in the spectral properties [25] of the MSM, which turn out to be quite unique due to he peculiar structure of the fitness. Indeed, the requirement that the node fitness is a positive random ('annealed') variable that is stable under addition implies that it has an α-stable distribution with infinite mean ($0 < \alpha < 1$) [1, 3], where classical random matrix results [25, 29, 30, 67, 68, 69, 31] break down. When a symmetric random matrix is decomposed into a low-rank informative part plus noise, strong signals can produce eigenvalue outliers outside the bulk spectrum, as first shown in [32]. In networks, such outliers reveal structures like communities or hubs [25, 33]. While infinite-mean regimes have been studied for Lévy matrices [70, 71, 72], random graphs add further randomness from adjacency realizations, not captured by those approaches. Our goal is to shed light on the surprising behaviour of the eigenmodes of the adjacency matrix in this infinite-mean regime, which follows from discrete node aggregation invariance.

§2.2 Preliminaries

The annealed MSM is an inhomogeneous random graph on n nodes, each node i being assigned a positive weight x_i, sampled independently from an α-stable distribution $\rho_\alpha(x)$, hence with power-law tails decaying as $x^{-1-\alpha}$ for $0 < \alpha < 1$ [1]. In this regime of α, every moment of $\rho_\alpha(x)$ diverges. Given x_i and x_j, the probability that an edge exists between i and j is

$$p_{ij} = f_n(x_i, x_j) = 1 - \exp(-\varepsilon_n x_i x_j), \quad i, j = 1, n \tag{2.1}$$

if $i \neq j$ and $p_{ii} = 0$ if $i = j$, where the parameter ε_n may be chosen to scale with n in a desired way. Under an arbitrary homogeneous partition whereby the n nodes are aggregated into supernodes (each with the same number of participating nodes), the form of both $p_{ij} = f_n(x_i, x_j)$ and $\rho_\alpha(x)$ is unchanged, up to a parameter shift in the latter – reflecting the fact that the fitness of a supernode is the sum of the fitnesses of its constituent nodes [1, 3]. Since $\rho(x)$ for positive α-stable variables is not known in closed-form (except for $\alpha = 1/2$), for analytical tractability we follow [3] and consider a pure Pareto(α) distribution with the same tail as an α-stable distribution:

$$\rho_\alpha(x) = \alpha\ x^{-1-\alpha} \quad 0 < \alpha < 1 \tag{2.2}$$

for $x \geq 1$, and $\rho_\alpha(x) = 0$ otherwise. We also consider $\varepsilon_n = n^{-1/\alpha}$ as a scaling that allows the graph to be arbitrarily large while remaining sparse, allowing precise asymptotic calculations for large n [3]. In this regime, the tail of the expected degree distribution $P(k)$ decays as k^{-2} irrespective of the value of α, the expected link density is of order $\ln n/n$, while the average local clustering coefficient remains finite [1, 3] – which are all widespread properties found in real-world networks.

§2.2.1 Realized versus expected matrix.

We denote the realized adjacency matrix of the graph as $\mathbf{A}$, where A_{ij} is a Bernoulli variable: $A_{ij} = 1$ with probability p_{ij} and 0 otherwise. Note that $\mathbf{A}$ is symmetric and $A_{ii} = 0$. By construction, the expectation of $\mathbf{A}$ (conditioning on the realized weights) is the matrix $\mathbf{P} = \mathbb{E}[\mathbf{A}|\{x_i\}]$ with entries $P_{ij} = p_{ij} = f_n(x_i, x_j)$ and captures the deterministic, structural signal imposed by the weights, while the actual $\mathbf{A}$ can be seen as a noisy realization of that signal.

Indeed, a central idea in our analysis is to decompose the adjacency matrix into the sum of an expected component and a fluctuation component, and then use Random Matrix Theory (RMT) intuition to separate the spectrum into bulk and outlier components. Let us write:

$$\mathbf{A} = \langle \mathbf{A} \rangle + (\mathbf{A} - \langle \mathbf{A} \rangle) = \mathbf{P} + \mathbf{H}, \tag{2.3}$$

where $\mathbf{H} = \mathbf{A} - \mathbf{P}$ is the zero-mean noise matrix. By construction, $\mathbf{P}$ is a dense matrix (encoding for the structure of the network) with a highly skewed pattern due to the heavy-tailed x_i, whereas $\mathbf{H}$ captures the randomness of edge realizations. If the matrix $\mathbf{P}$ were of small, finite rank and $\mathbf{H}$ were a 'typical' random matrix with

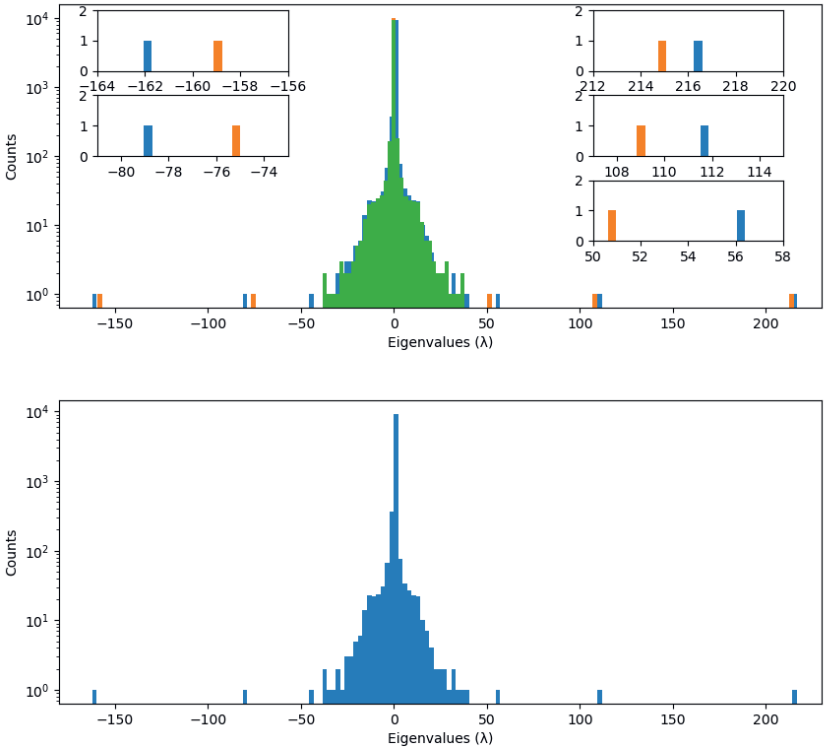


Figure 2.1: The full spectrum of $\boldsymbol{A}$*, a single realization of the network adjacency matrix from* $\boldsymbol{P}$*, with* $n = 10^4$ *and* $\alpha = 0.5$*. In the top panel the histogram (blue) of* $\boldsymbol{A}$*. We see separation into atomic delta densities and a central, continuous bulk. In the bottom panel we show the separation of the spectrum into two regimes, the outliers being those of* $\boldsymbol{A}$ *in blue, and of* $\boldsymbol{P}$ *in orange. We also superimpose in green the eigenvalue histogram of* $\boldsymbol{H}$*, that replicates the bulk behaviour of* $\boldsymbol{A}$*.*

a bounded spectral distribution, then standard results show that each eigenvalue of $\mathbf{P}$ above a certain magnitude would detach from the continuous spectrum of $\mathbf{H}$. This phenomenon, rigorously characterized in spiked random matrix theory, underlies applications like spectral clustering [26, 27, 28]. In our case, $\mathbf{P}$ is not finite-rank: as we argue in the following sections though, we find evidence of an 'effective' rank of order $\ln n$. In this sense, we are close in spirit to recent results such as Ref. [73], that are hinting at the fact that generalized BBP transitions could be happening even when the rank of the perturbation matrix is growing as $o(n)$. An illustration of the typical spectra of $\mathbf{A}$ and $\mathbf{P}$ is provided in Fig. 2.1. Our main goal is the characterization of these spectra and the associated eigenvectors, with a specific attention to the outlier modes carrying structural information.

§2.3 Leading eigenmodes from the expected matrix.

We are interested in the solution of the eigenmode equation [25]

$$\mathbf{P}v_k = \lambda_k v_k \quad k = 1, n \tag{2.4}$$

with $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$. Denoting the i-th entry of each eigenvector v_k as $v_k^{(i)}$, we can rewrite Eq.(2.4) as

$$\sum_{j=1}^{n} p_{ij} v_k^{(j)} = \lambda_k v_k^{(i)} \quad i = 1, n \quad k = 1, n \tag{2.5}$$

Note that, since $\mathbf{P}$ is real and symmetric, its eigenvalues are real. Also note that any two vertices i, j with the same realized weight $x_i = x_j$ are statistically equivalent, as their (expected) properties in the graph depend only on the weight x. Similarly, if $x_i = x_j$ then the i-th and j-th entries of each eigenvector v_k will be equal: $v_k^{(i)} = v_k^{(j)}$. This allows us to write $v_k^{(i)} = \nu_k(x_i)$ where $\nu_k(x)$ must be a continuous function of the weight x. For large enough n, Eq.(2.5) can be recast as the integral expression

$$\Pi_{\alpha,n}[\nu_k(x)] = \frac{\lambda_k}{n}\ \nu_k(x). \tag{2.6}$$

where $\nu_k(x)$ is an eigenfunction (with eigenvalue λ_k/n) of the integral operator

$$\Pi_{\alpha,n}[g(x)] \equiv \int_0^{+\infty} dy\ \rho_\alpha(y)\ (1 - e^{-\varepsilon_n xy})\ g(y), \tag{2.7}$$

corresponding to the eigenvector v_k (with eigenvalue λ_k) of the original matrix $\mathbf{P}$. A natural normalization for $\nu_k(x)$ will emerge from the calculations that follow. If an eigenfunction $\nu_k(x)$ is found, a convenient way to go back to the actual eigenvector entry $v_k^{(j)}$ is the following. We first order all vertices so that $x_1 \leq x_2 \leq \cdots \leq x_n$. Then, since a pure Pareto(α) variable can be represented as a uniform random variable, raised to the power $-1/\alpha$, in the interval $[0, 1]$, we arrive at the relation $x_j = (n/j)^{1/\alpha}$ (we will use this relation as an alternative, deterministic assignment of the fitness of nodes to produce a 'noise-free' power law distribution even for finite n). Finally,

$$v_k^{(j)} = \nu_k(x_j) = \nu_k\left(\frac{n^{1/\alpha}}{j^{1/\alpha}}\right) \quad j = 1, n \quad k = 1, n. \tag{2.8}$$

§2.3.1 Largest eigenvalue and principal eigenvector

We start with the principal component $k = 1$. From Perron-Frobenius theorem, λ_1 is be positive and v_1 has all entries of the same sign. We will set the normalization of v_1 in such a way that all its entries are positive, so that $\nu_1(x)$ is a positive-valued function. When evaluating the operator 2.7 on the eigenfunction, the multiplication of $\rho_\alpha(x)$ by $\nu_1(x)$ formally generates a modified, effective distribution of the weights.

We then look for an eigenfunction such that $\nu_1(x)\rho_\alpha(x) = \rho_\beta(x)$ is again a Pareto pdf with a new index β. This is obtained for the eigenfunction

$$\nu_1(x) = \frac{\beta}{\alpha}\, x^{\alpha-\beta}, \quad \beta \in \mathbb{R}^+. \tag{2.9}$$

Setting $g(x) = \nu_1(x)$ into Eq. (2.7), we get

$$\Pi_{\alpha,n}[\nu_1(x)] = 1 - \varphi_\beta(\varepsilon_n x) \tag{2.10}$$

where we have introduced the Laplace Transform (LT) φ_α of the Pareto-α density, defined as

$$\varphi_\alpha(t) \equiv \int_0^{+\infty} dx\ \rho_\alpha(x)\ e^{-tx} = \alpha \int_1^{+\infty} dx\ x^{-1-\alpha}\ e^{-tx}, \tag{2.11}$$

and we have exploited the normalization property

$$\int_0^{+\infty} dy\ \rho_\alpha(y)\ \nu_1(y) = \int_0^{+\infty} dy\ \rho_\beta(y) = 1. \tag{2.12}$$

The above expression sets our normalization for $\nu_1(x)$, which is the analogous of the ℓ_1 norm $\sum_{i=1}^n |v_1^{(i)}| = 1$ for the original eigenvector v_1.

We can easily derive the asymptotics properties of the Laplace transform of a Pareto distribution. Indeed

$$\begin{aligned}
\Pi_{\alpha,n}[\nu_1(x)] &= 1 - \varphi_\beta(\varepsilon_n x) \\
&= 1 - \int_1^{+\infty} dy\ \rho_\beta(y)\ e^{-\varepsilon_n xy} \\
&= 1 - \beta \int_1^{+\infty} dy\ y^{-1-\beta}\ e^{-\varepsilon_n xy} \\
&= 1 - \beta \int_{\varepsilon_n x}^{+\infty} dt\ \left(\frac{t}{\varepsilon_n x}\right)^{-1-\beta} \frac{e^{-t}}{\varepsilon_n x} \\
&= 1 - \frac{\beta}{(\varepsilon_n x)^{-\beta}} \int_{\varepsilon_n x}^{+\infty} dt\ t^{-1-\beta}\ e^{-t} \\
&= 1 - \frac{\beta}{(\varepsilon_n x)^{-\beta}}\ \Gamma(-\beta, \varepsilon_n x),
\end{aligned} \tag{2.13}$$

where we have made the change of variables $\varepsilon_n xy \to t$ and introduced the incomplete Gamma function $\Gamma(z,s)$, which has the following definition and asymptotic relation with the (complete) Gamma function $\Gamma(z) \equiv \Gamma(z,0)$ for $\Re[z] < 0$:

$$\Gamma(z,s) \equiv \int_s^{+\infty} dt\ t^{z-1}e^{-t}\ \sim\ \Gamma(z) + \frac{s^z}{z}, \qquad s \to 0^+. \tag{2.14}$$

Since in our case $\Re[z] = -\alpha/2 < 0$, the use of the above relation is justified. Plugging

Eq. (2.14) into Eq. (2.13), we obtain for $\varepsilon_n x \to 0^+$:

$$\begin{aligned} \Pi_{\alpha,n}[\nu_1(x)] &\sim 1 - \frac{\beta}{(\varepsilon_n x)^{-\beta}} \left[\Gamma(-\beta) - \frac{(\varepsilon_n x)^{-\beta}}{\beta} \right] \\ &= -\beta\ (\varepsilon_n x)^{\beta}\ \Gamma(-\beta) \\ &= (\varepsilon_n x)^{\beta}\ \Gamma(1-\beta) \end{aligned} \tag{2.15}$$

This allows us to confirm that $\nu_1(x)$ is asymptotically an eigenfunction, for a suitable choice of β. Indeed, applying Eq. (2.15) into Eq. (2.10), we can recast Eq. (2.6) as

$$\Pi_{\alpha,n}[\nu_1(x)] \sim (\varepsilon_n x)^{\beta}\ \Gamma(1-\beta) = \frac{\lambda_1}{n}\ \frac{\beta}{\alpha}\ x^{\alpha-\beta} \tag{2.16}$$

(provided we identify $\beta \equiv \alpha - \beta$, i.e. $\beta = \alpha/2$) to find

$$\lambda_1 = \frac{\alpha}{\beta}\ \Gamma\left(1-\beta\right) n\ \varepsilon_n^{\beta} = -\alpha\ \Gamma\left(-\frac{\alpha}{2}\right) n^{1/2}. \tag{2.17}$$

This square-root growth of λ_1 can be confirmed via actual computations on the matrix $\mathbf{P}$ itself (see Fig. 2.2). Using Eq. (2.8), we also obtain all the eigenvector entries as

$$v_1^{(j)} = \nu_1(x_j) = \frac{1}{2} x_j^{\alpha/2} = \frac{1}{2} \left(\frac{n}{j}\right)^{1/2} \qquad j = 1, n. \tag{2.18}$$

This is also confirmed numerically later on.

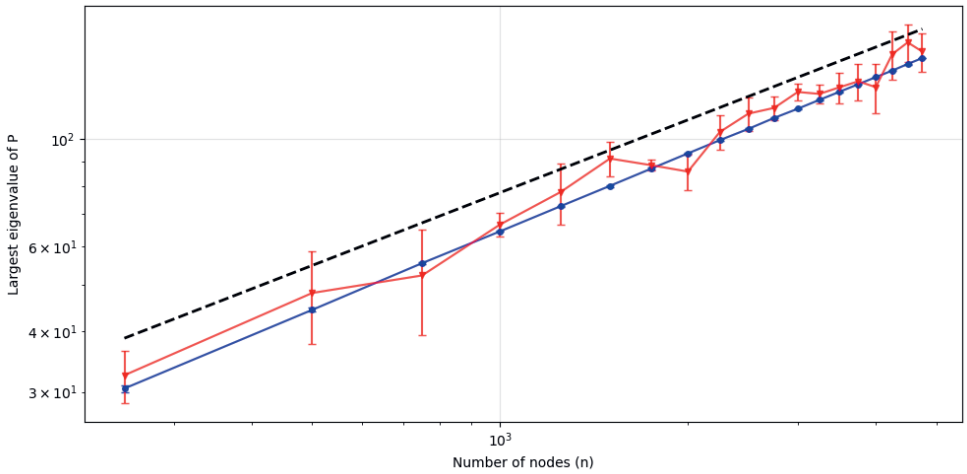


Figure 2.2: Largest eigenvalue λ_1 of $\boldsymbol{P}$ versus the number n of nodes, with weights $\{x_i\}_{i=1}^n$ either sampled i.i.d. from a Pareto distribution (in red, averaged over 5 samples) or assigned deterministically as $x_i = (n/i)^{1/\alpha}$ (in blue), compared with the analytic result in Eq. (2.17) (in black). Here $\alpha = 0.5$.

§2.3.2 Next eigenvalues

We now turn $k > 1$. Let us preliminarily introduce auxiliary complex-valued test functions $\nu_k^{\pm}(x)$ of the form still given by Eq. (2.9), with $\beta \in \mathbb{C}$:

$$\nu_k^{\pm}(x) = \frac{\beta_k^{\pm}}{\alpha}\, x^{\alpha-\beta_k^{\pm}}, \quad \beta_k^{\pm} \equiv \gamma_k \pm \mathbf{i}\,\omega_k \in \mathbb{C} \tag{2.19}$$

where $\gamma_k = \Re[\beta_k^{\pm}] \in \mathbb{R}$ and $\pm\omega_k = \Im[\beta_k^{\pm}] \in \mathbb{R}$ are the real and imaginary parts of $\beta_k^{\pm}$, respectively. Note that β_k^{+} and β_k^{-} are complex conjugate. Just like $\rho_\beta(x) \equiv \nu_1(x)\rho_\alpha(x)$ is a Pareto pdf with modified (real) index β, which has the same tail behaviour of a stable distribution with index β, $\rho_{\beta_k^{\pm}}(x) \equiv \nu_k^{\pm}(x)\rho_\alpha(x)$ is formally a generalized Pareto distribution with the same 'tail behaviour' as the recently defined stable distributions with complex-valued index $\beta_k^{\pm}$ [74, 75]. As we now show, these complex-valued scaling exponents generate log-periodicity, as typically found in systems with discrete scale invariance [35].

We will use the fact that the Gamma function of the complex conjugate of its argument is the complex conjugate of the Gamma function. The discussion in the previous section extends to complex-valued parameters $\beta_k^{\pm} = \gamma_k \pm \mathbf{i}\,\omega_k$, leading to the following generalization of Eq. (2.15):

$$\Pi_{\alpha,n}[\nu_k^{\pm}(x)] \sim (\varepsilon_n x)^{\beta_k^{\pm}}\ \Gamma(1-\beta_k^{\pm}). = (\varepsilon_n x)^{\gamma}(\varepsilon_n x)^{\pm i\omega}\ \Gamma(1-\gamma \mp i\omega_k^{\pm}) \tag{2.20}$$

We then introduce the candidate eigenfunction

$$\begin{aligned} \nu_k(x) &= \Re\left[\Gamma(1-\beta_k^{-})\ \nu_k^{+}(x)\right] && (2.21)\\ &= \frac{1}{2}\left[\Gamma(1-\beta_k^{-})\ \nu_k^{+}(x) + \Gamma(1-\beta_k^{+})\ \nu_k^{-}(x)\right] && (2.22) \end{aligned}$$

,

then Eq. (2.20) leads to the following asymptotics:

$$\begin{aligned} \Pi_{\alpha,n}[\nu_k(x)] &\sim \frac{1}{2}\,\Gamma(1-\beta_k^{+})\,\Gamma(1-\beta_k^{-})\ \left[(\varepsilon_n x)^{\beta_k^{+}} + (\varepsilon_n x)^{\beta_k^{-}}\right] && (2.23)\\ &= \left|\Gamma(1-\beta_k^{+})\right|^2\ \Re\left[(\varepsilon_n x)^{\beta_k^{+}}\right] && (2.24)\\ &= |\Gamma(1-\gamma_k \pm i\omega_k)|^2(\varepsilon_n x)^{\gamma_k}\left[(\varepsilon_n x)^{+i\omega} + (\varepsilon_n x)^{-i\omega}\right]. && (2.25) \end{aligned}$$

To enforce the eigenvector equation $\Pi_{\alpha,n}[\nu_k(x)] = \frac{\lambda_k}{n}\ \nu_k(x)$, the above expression has to equal

$$\begin{aligned} \frac{\lambda_k}{n}\ \nu_k(x) &= \frac{\lambda_k}{2n\alpha}\left[\Gamma(1-\beta_k^{-})\ \beta_k^{+}\ x^{\alpha-\beta_k^{+}} + \Gamma(1-\beta_k^{+})\beta_k^{-}\ x^{\alpha-\beta_k^{-}}\right] && (2.26)\\ &= \frac{\lambda_k}{n\alpha}\Re\left[\Gamma(1-\beta_k^{-})\ \beta_k^{+}\ x^{\alpha-\beta_k^{+}}\right]. && (2.27) \end{aligned}$$

This requires either $\beta_k^{+} \equiv \alpha - \beta_k^{+}$ and $\beta_k^{-} \equiv \alpha - \beta_k^{-}$, which would however retrieve the case $k = 1$ that we have already considered for the largest eigenvalue

($\beta_k^{\pm} = \gamma_k = \alpha/2$ and $\omega_k = 0$, since $\alpha \in \mathbb{R}$), or $\beta_k^{+} \equiv \alpha - \beta_k^{-}$ and $\beta_k^{-} \equiv \alpha - \beta_k^{+}$, which still leads to $\gamma_k = \alpha/2$ (for all k) but puts no restriction on ω_k, so that $\beta_k^{\pm} = \alpha/2 \pm \mathbf{i}\,\omega_k$. The defining restriction on ω_k is obtained by matching the other terms in Eqs. (2.25) and (2.27), yielding the two simultaneous conditions

$$\lambda_k = -\alpha\ \Gamma\left(-\frac{\alpha}{2} \mp \mathbf{i}\,\omega_k\right)\ n^{1/2\ \pm\ \mathbf{i}\,\omega_k/\alpha} = -\alpha\ \Gamma\left(-\frac{\alpha}{2} \mp \mathbf{i}\,\omega_k\right)\ n^{1/2}\ e^{\pm \mathbf{i}\,(\omega_k/\alpha)\ln n}. \tag{2.28}$$

Since the right-hand sides of the above expression are complex conjugates of each other, the two simultaneous conditions ensure that λ_k is real, as required for the eigenvalues of a symmetric matrix. This means that ω_k has to be such that $\arg[\lambda_k] = 0, \pi$ modulo 2π, i.e.

$$\arg\left[\Gamma\left(-\frac{\alpha}{2} + \mathbf{i}\,\omega_k\right)\right] = \frac{\omega_k}{\alpha}\ln n - k\pi, \quad k \in \mathbb{N}, \qquad -\pi \le \arg\left[\Gamma\left(-\frac{\alpha}{2} + \mathbf{i}\,\omega_k\right)\right] < +\pi, \tag{2.29}$$

Note that the (necessarily real-valued) eigenvalues λ_k ($k \ge 1$) can be found graphically as the intersections between the real axis and the two spirals

$$\sigma^{\pm}(\omega) \equiv -\alpha\ \Gamma\left(-\frac{\alpha}{2} \mp \mathbf{i}\,\omega\right) n^{1/2} e^{\pm \mathbf{i}\,(\omega/\alpha)\ln n} \tag{2.30}$$

defined parametrically in the complex plane by extending the ω_k in Eq.(2.28) to any real value $\omega \ge 0$. See Fig. 2.3 for an illustration of this result, for three representative values of the index $\alpha = 0.2, 0.5, 0.8$. When the solutions for ω_k that realize Eq. (2.29) (or equivalently the intersections) are inserted into Eq.(2.28), they produce all the eigenvalues as

$$\lambda_k = (-1)^k\ \alpha\left|\Gamma\left(-\frac{\alpha}{2} + \mathbf{i}\,\omega_k\right)\right|\ n^{1/2}, \tag{2.31}$$

which nicely extends the expression for λ_1 to values $k > 1$.

Graphical inspection

Unfortunately, Eq. (2.29) cannot be solved explicitly for ω_k, except for $k = 0$ (for which there is no solution for ω_0) and $k = 1$, for which $\omega_1 = 0$ is a solution, leading to the value λ_1 when inserted into Eq.(2.31): indeed, $\Gamma\left(-\frac{\alpha}{2}\right) < 0$, so that $\arg\Gamma\left(-\frac{\alpha}{2}\right) = -\pi$ solves Eq. (2.29) for $k = 1$ and $\omega_1 = 0$. The solutions for $k > 1$ can be inspected graphically, either as the aforementioned intersections between the real axis and the complex spiral, or as the intersections between the two curves defined as the l.h.s. and r.h.s. of Eq. (2.29). Specifically, the l.h.s. defines the function

$$f(\omega) \equiv \arg\Gamma\left(-\frac{\alpha}{2} + \mathbf{i}\,\omega\right), \tag{2.32}$$

while the r.h.s. defines the family of straight lines

$$g_k(\omega) \equiv \omega\frac{\ln n}{\alpha} - k\pi, \quad k \ge 0. \tag{2.33}$$

The solutions ω_k to Eq. (2.29) are the values at the intersections

$$f(\omega_k) = g_k(\omega_k), \quad k \ge 0. \tag{2.34}$$

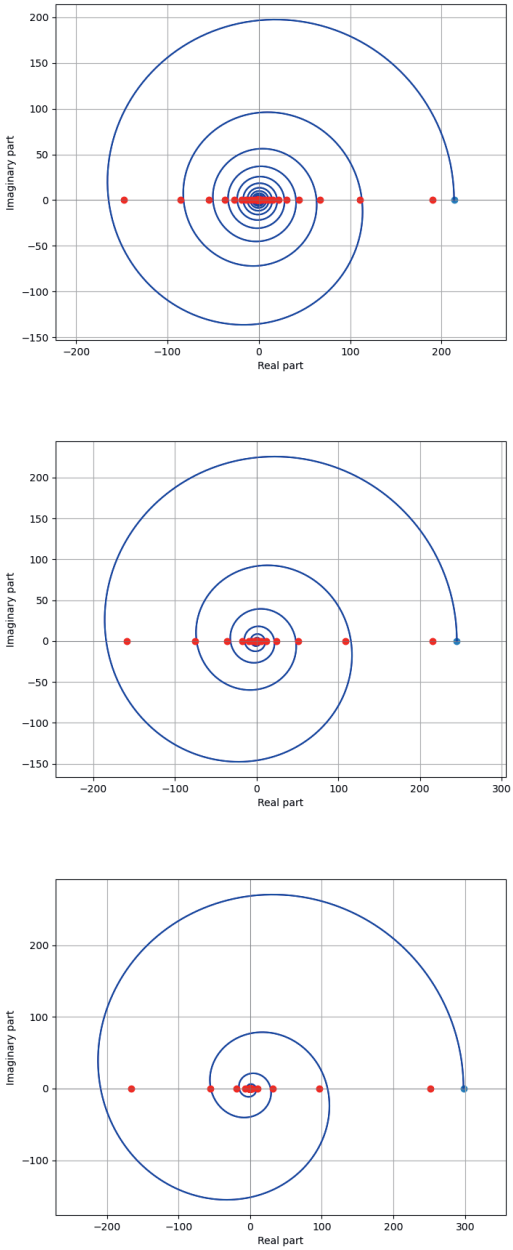


Figure 2.3: Comparison between the actual real-valued eigenvalues (red) of the expected adjacency matrix $\boldsymbol{P}$*, obtained by solving Eq.* (2.29) *numerically, and the predicted eigenvalues obtained as the intersection between the real axis and the logarithmic spiral* σ_+ *(blue) defined in Eq.* (2.30)*. Here* $n = 10^4$ *and* $\alpha = 0.2, 0.5, 0.8$ *from top to bottom.*

In Fig.2.4 we plot $f(\omega)$ and $g_k(\omega)$ for a few small values of k. The plot confirms that there is no intersection point ω_0 for $k = 0$, while the intersection for $k = 1$ is found at $\omega_1 = 0$, for which $f(\omega_1) = g_1(\omega_1) = -\pi$. What is more useful from the graphical inspection is that the next few intersections for $1 < k \leq k^*$ occur at a range of values ω_k close to some value ω_α at which $f(\omega)$ has a stationary point and varies very slowly around the corresponding value $\phi_\alpha \equiv f(\omega_\alpha)$, which is approximately a local 'plateau' for the function (the value k^* is defined such that for $k > k^*$ the value $f(\omega_k)$ is away from the plateau). We can therefore approximate the solutions ω_k for the first few values $k = 2, 3, \ldots, k^*$ as the points at which $g_k(\omega)$ intersects the constant line $f(\omega) \approx \phi_\alpha$, and replace Eq. (2.34) with

$$\phi_\alpha \approx g_k(\omega_k), \quad 1 < k \leq k^*. \tag{2.35}$$

which, using Eq. (2.33), is solved by the values

$$\omega_k \approx \alpha \frac{k\pi + \phi_\alpha}{\ln n}, \quad \phi_\alpha \equiv f(\omega_\alpha), \quad 1 < k \leq k^*. \tag{2.36}$$

To complete our characterization of these approximate solutions, we need to estimate both k^* and ω_α.

Estimation of ω_α

We start from the determination of ω_α. Since the latter is defined as a stationary point of $f(\omega)$, we look for its value by calculating the derivative $f'(\omega)$ and requiring $f'(\omega_\alpha) = 0$, or more precisely:

$$\frac{\partial}{\partial \omega} \arg\Gamma\left(-\frac{\alpha}{2} + \mathbf{i}\,\omega\right)\Big|_{\omega=\omega_\alpha} = 0. \tag{2.37}$$

We now rearrange the identity $z = |z| e^{\mathbf{i} \arg z} = \sqrt{z\,\bar{z}}\; e^{\mathbf{i} \arg z}$ (valid for every complex number z, where $\bar{z}$ is the complex conjugate of z) to express

$$\arg z = -\mathbf{i} \ln \frac{z}{|z|} = \frac{\mathbf{i}}{2} \ln \frac{\bar{z}}{z}. \tag{2.38}$$

If $z = \arg\Gamma\left(-\frac{\alpha}{2} + \mathbf{i}\,\omega\right)$, we get

$$\arg\Gamma\left(-\frac{\alpha}{2} + \mathbf{i}\,\omega\right) = \frac{\mathbf{i}}{2} \ln \frac{\Gamma\left(-\frac{\alpha}{2} - \mathbf{i}\,\omega\right)}{\Gamma\left(-\frac{\alpha}{2} + \mathbf{i}\,\omega\right)} = \frac{\mathbf{i}}{2} \ln\Gamma\left(-\frac{\alpha}{2} - \mathbf{i}\,\omega\right) - \frac{\mathbf{i}}{2} \ln\Gamma\left(-\frac{\alpha}{2} + \mathbf{i}\,\omega\right). \tag{2.39}$$

In order to be able to differentiate the above expression with respect to ω, we need a closed-form expression for $\ln\Gamma(z)$ (also called the 'log-Gamma function'), which is available as follows:

$$\ln\Gamma(z) = -\gamma z - \ln z + \sum_{m=1}^{+\infty} \left[\frac{z}{m} - \ln\left(1 + \frac{z}{m}\right)\right], \tag{2.40}$$

where

$$\gamma \equiv \lim_{m\to+\infty} \left[-\ln m + \sum_{l=1}^{m} \frac{1}{l}\right] = 0.5772\ldots \tag{2.41}$$

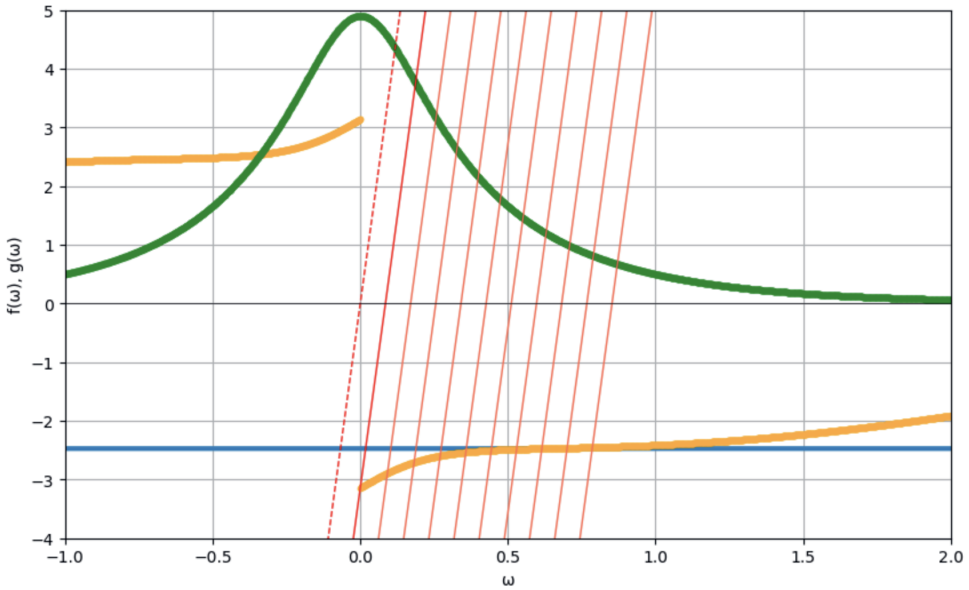


Figure 2.4: Graphical Solution of Eq. (2.29). *In orange, the function* $f(\omega)$ *defined in Eq.* (2.32), *while the straight lines* $g_k(\omega)$ *defined by Eq.* (2.33) *are in red, with* k *from* 0 *to* 10. *The dashed red line corresponds to* $g_0(\omega)$, *which has no intersection* ω_0 *with* $f(\omega)$. *The solid red line at its immediate right is* $g_1(\omega)$, *which intersects* $f(\omega)$ *at the exact value* $\omega_1 = 0$ *where* $f(\omega_1) = g_1(\omega_1) = -\pi$. *Further on the right, the next straight lines* $g_k(\omega)$ *(with* $k > 1$*) are illustrated in light red. The solution to Eq.* (2.29) *for these values of* k *can be obtained graphically where the orange and light red lines meet. In this regime,* $f(\omega)$ *can be approximated as the horizontal line* $f(\omega) \approx \phi_\alpha \equiv f(\omega_\alpha)$ *(blue line), where the value* ω_α *is the stationary point such that* $f'(\omega_\alpha) = 0$. *The horizontal line is no longer a good approximation for* $f(\omega_k)$ *for larger* k, *but that is already the regime for which the value of* $\left|\Gamma\left(-\frac{\alpha}{2} + \mathbf{i}\,\omega_k\right)\right|$ *(plotted in green) is exponentially suppressed, as one can verify from Eqs.* (2.53).

is the Euler-Mascheroni constant. Using Eq. (2.40) with $z = -\alpha/2 \pm \mathbf{i}\,\omega$ into Eq. (2.39), we arrive at

$$\arg\Gamma\left(-\frac{\alpha}{2} + \mathbf{i}\,\omega\right) = -\gamma\omega - \arg\left(-\frac{\alpha}{2} + \mathbf{i}\,\omega\right) + \sum_{m=1}^{+\infty}\left[\frac{\omega}{m} - \arg\left(m - \frac{\alpha}{2} + \mathbf{i}\,\omega\right)\right]. \quad (2.42)$$

Now, in order to differentiate with respect to ω as prescribed by Eq. (2.37), we preliminarily calculate

$$\frac{\partial}{\partial\omega}\arg\left(m - \frac{\alpha}{2} + \mathbf{i}\,\omega\right) = \frac{\mathbf{i}}{2}\frac{\partial}{\partial\omega}\ln\left(\frac{m - \frac{\alpha}{2} - \mathbf{i}\,\omega}{m - \frac{\alpha}{2} + \mathbf{i}\,\omega}\right) = \frac{m - \frac{\alpha}{2}}{\left(m - \frac{\alpha}{2}\right)^2 + \omega^2}, \quad (2.43)$$

where we have used Eq. (2.38) again. Using the above expression into Eq. (2.42), we obtain

$$\frac{\partial}{\partial\omega}\arg\Gamma\left(-\frac{\alpha}{2} + \mathbf{i}\,\omega\right) = -\gamma + \frac{\frac{\alpha}{2}}{\left(\frac{\alpha}{2}\right)^2 + \omega^2} + \sum_{m=1}^{+\infty}\left[\frac{1}{m} - \frac{m - \frac{\alpha}{2}}{\left(m - \frac{\alpha}{2}\right)^2 + \omega^2}\right]. \quad (2.44)$$

Now, imposing Eq. (2.37) translates into the following:

$$-\gamma + \frac{\frac{\alpha}{2}}{\left(\frac{\alpha}{2}\right)^2 + \omega_\alpha^2} + \sum_{m=1}^{+\infty} \left[\frac{1}{m} - \frac{m - \frac{\alpha}{2}}{\left(m - \frac{\alpha}{2}\right)^2 + \omega_\alpha^2} \right] = 0, \tag{2.45}$$

which, if the sum is truncated to any finite order for m, can be solved algebraically for ω_α. In general, the number of solutions of the above equation (i.e. stationary points of the function f) may depend on α. How many (and how closely) such solutions are recovered after truncating the sum will depend on the order of truncation. For our purposes here (i.e. identifying the first stationary point as shown in Fig. 2.4), we simply truncate to zeroth order and neglect the sum altogether, to get the condition

$$-\gamma + \frac{\frac{\alpha}{2}}{\left(\frac{\alpha}{2}\right)^2 + \omega_\alpha^2} = 0, \tag{2.46}$$

which is solved by the single (positive) value

$$\omega_\alpha \equiv \sqrt{\frac{\alpha}{2}\left(\frac{1}{\gamma} - \frac{\alpha}{2}\right)}. \tag{2.47}$$

This finally allows us to compute

$$\phi_\alpha \equiv f(\omega_\alpha) = f\left(\sqrt{\frac{\alpha}{2}\left(\frac{1}{\gamma} - \frac{\alpha}{2}\right)}\right), \tag{2.48}$$

which, when inserted into Eq. (2.36), produces our approximate solutions for ω_k to Eq. (2.29) with $1 < k \leq k^*$. A comparison between the actual solutions ω_k of Eq. (2.29) and the approximate solutions given in (2.36) is shown in Fig. 2.5, confirming a very good agreement, especially for smaller values of α. Note that our only use of these approximate solutions is the approximate proportionality between ω_k and k, which is used below to prove the exponential decay of the prefactor of the leading eigenvalues with k. This linearity only requires that $f(\omega)$ is sufficiently flat around some value ϕ_α before $\left|\Gamma\left(-\frac{\alpha}{2} + \mathbf{i}\,\omega\right)\right|$ is exponentially suppressed (see Fig. 2.4). A more accurate estimate of ϕ_α would improve the estimate of ϕ_α used in Eq. (2.36) (hence improving the slope of the predicted lines in Fig. 2.5), but not modify the approximate proportionality between ω_k and k, which is all we need in what follows.

Exponential decay of the prefactor of the leading eigenvalues

We now turn to the estimation of the order of k^*. Let us consider the Stirling approximation [159, (5.11.9)] for the absolute value of the Gamma function:

$$|\Gamma(a + \mathbf{i}\,b)| \sim \sqrt{2\pi}|b|^{a-\frac{1}{2}} e^{-\frac{\pi|b|}{2}}, \tag{2.49}$$

where $a, b \in \mathbb{R}$. Choosing $a = -\alpha/2$ and $b = \omega_k \geq 0$, this implies

$$\left|\Gamma\left(-\frac{\alpha}{2} + \mathbf{i}\,\omega_k\right)\right| \sim \sqrt{2\pi}\ \omega_k^{-\frac{\alpha+1}{2}} e^{-\frac{\pi\omega_k}{2}}, \tag{2.50}$$

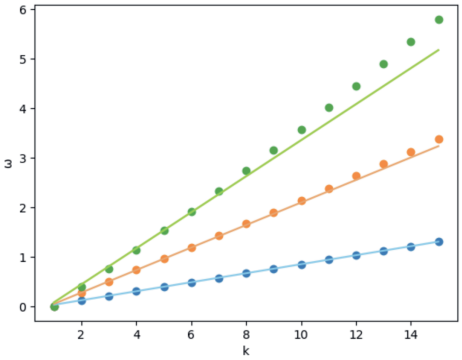


Figure 2.5: Values of ω_k as a function of k. The isolated points are the actual, numerical solutions of Eq. (2.29), for $\alpha = 0.2, 0.5, 0.8$ (in blue, orange and green respectively). The solid lines are the approximate linear solutions defined by Eq. (2.36) with $k > 1$ (plus the exact solution $\omega_1 = 0$) for the same values of $\alpha = 0.2, 0.5, 0.8$ (and same colors).

showing an exponential decay with increasing ω_k. This exponential decay is visible in Fig. 2.4. If we further use the expression we derived in Eq. (2.36) for the approximate solutions ω_k for $1 < k \leq k^*$, we see that the above behaviour translates into an exponential decay as a function of $k/\ln n$:

$$\left|\Gamma\left(-\frac{\alpha}{2} + \mathbf{i}\,\omega_k\right)\right| \sim \sqrt{2\pi}\left(\alpha\frac{k\pi+\phi_\alpha}{\ln n}\right)^{-\frac{\alpha+1}{2}} e^{-\frac{\alpha\pi}{2}\frac{k\pi+\phi_\alpha}{\ln n}}, \quad 1 < k \leq k^*. \tag{2.51}$$

Looking at Fig. 2.4, we see that when $k \gtrsim k^*$ (i.e. when the intersections between $f(\omega)$ and $g_k(\omega)$ depart from the approximate plateau of constant height ϕ_α), the value of $\left|\Gamma\left(-\frac{\alpha}{2} + \mathbf{i}\,\omega\right)\right|$ is already exponentially suppressed. Therefore Eq. (2.51) indicates that k^* is of order $\ln n$. Obviously, the magnitude of the corresponding eigenvalue λ_k of $\mathbf{P}$, which is given by Eq. (2.31), is also exponentially (in $k/\ln n$) suppressed:

$$|\lambda_k| = \alpha\left|\Gamma\left(-\frac{\alpha}{2} + \mathbf{i}\,\omega_k\right)\right|\sqrt{n} \tag{2.52}$$

$$\sim \alpha\sqrt{2\pi n}\left(\alpha\frac{k\pi+\phi_\alpha}{\ln n}\right)^{-\frac{\alpha+1}{2}} e^{-\frac{\alpha\pi}{2}\frac{k\pi+\phi_\alpha}{\ln n}}. \tag{2.53}$$

This means that, for $k = O(\ln n)$, the approximation in Eq. (2.36) breaks down, but on the other hand the corresponding values of λ_k become exponentially small. This is also where, moving from the expected matrix $\mathbf{P}$ to a realized random matrix $\mathbf{A}$, the randomness will create noisy eigenvalues 'eating up' the eigenvalues of $\mathbf{P}$.

Indeed, generalized BBP-type results typically assume that the expected matrix $\mathbf{P}$ has a fixed, finite rank in order to guarantee a clean separation between informative (structural) eigenvalues and the random bulk. In our model this assumption fails literally, since $\mathbf{P}$ is not finite-rank. Nevertheless, both our numerics and the above arguments indicate that the k-th largest (in absolute value) eigenvalue λ_k of $\mathbf{P}$ decays rapidly with $k/\ln n$, so that the rank of the matrix is *effectively* of order $\ln n$, and

hence $o(n)$. This aligns with recent literature suggesting that a BBP-like transition can persist to order $o(n)$, hence also beyond strict finite rank [73, 76]. Numerics also indicate that the edge of the random spectrum contributed by $\mathbf{H}$ scales as $C(\alpha)\sqrt{n}$ with $C(\alpha)$ independent of n. Conequently, only those signal eigenvalues that are not exponentially suppressed can rise above the bulk, i.e., indices up to $k^* \sim \ln n$. In this sense, the signal behaves as if it had *effective rank* $\Theta(\ln n)$: the spectrum of $\mathbf{A}$ consists of order $\ln n$ outliers (delta spikes from $\mathbf{P}$) on top of the bulk determined by $\mathbf{H}$.

§2.3.3 Next Eigenvectors

We now turn to the eigenvectors for $k > 1$. Plugging ω_k into Eq. (2.22), we get

$$\nu_k(x) = \frac{x^{\alpha/2}}{\alpha} \Re\left[\left(\frac{\alpha}{2} - \mathbf{i}\,\omega_k\right)\Gamma\left(1 - \frac{\alpha}{2} - \mathbf{i}\,\omega_k\right) x^{\mathbf{i}\,\omega_k}\right]. \tag{2.54}$$

Using Eq. (2.8), we re-express the eigenvector entries as

$$v_k^{(j)} = \sqrt{\frac{n}{j\alpha^2}}\,\Re\left[\left(\frac{\alpha}{2} - \mathbf{i}\,\omega_k\right)\Gamma\left(1 - \frac{\alpha}{2} - \mathbf{i}\,\omega_k\right)\left(\frac{n}{j}\right)^{\frac{\mathbf{i}\,\omega_k}{\alpha}}\right] \tag{2.55}$$

$$= -\sqrt{\frac{n}{j}}\left(\frac{\alpha}{4} + \frac{\omega_k^2}{\alpha}\right)\Re\left[\Gamma\left(-\frac{\alpha}{2} - \mathbf{i}\,\omega_k\right)\left(\frac{n}{j}\right)^{\frac{\mathbf{i}\,\omega_k}{\alpha}}\right] \tag{2.56}$$

Since $x^{\mathbf{i}\,\omega_k} = e^{\mathbf{i}\,\omega_k \ln x}$ and $j^{\mathbf{i}\,\omega_k/\alpha} = e^{\mathbf{i}\,(\omega_k/\alpha)\ln j}$, the structure of the above eigenfunctions and eigenvectors is oscillating, log-periodic in their arguments (x and j respectively), and modulated by a power-law profile ($x^{\alpha/2}$ and $j^{-1/2}$ respectively). In Figs. 2.6 2.7 2.8 we show that our analytically derived log-periodic formula tracks the actual eigenvectors of both $\mathbf{P}$ and $\mathbf{A}$ remarkably well for small k.

These eigenvectors do not display traditional delta-like localization, as they are broadly supported on a wide range of weight percentiles with largest amplitude around the hub node ($j = 1$), followed by power-law decay $j^{-1/2}$. As k grows large enough (below we provide evidence that this happens for k growing with n but of order at most $\ln n$), we find a departure of the eigenvectors of $\mathbf{A}$ from those of $\mathbf{P}$, due to the entrance of the associated eigenvalues into the random bulk where eigenvectors are expected to be more uniformly spread and less structured.

We also note that, when combined with Eq. (2.31), the above expression allows to rewrite the eigenvector entries in terms of the corresponding eigenvalue λ_k as

$$v_k^{(j)} = \frac{(-1)^{k+1}\,\lambda_k}{\sqrt{j}}\left(\frac{\alpha}{4} + \frac{\omega_k^2}{\alpha}\right)\Re\left[\,j^{\frac{\mathbf{i}\,\omega_k}{\alpha}}\right] = \frac{(-1)^{k+1}\,\lambda_k}{\sqrt{j}}\left(\frac{\alpha}{4} + \frac{\omega_k^2}{\alpha}\right)\cos\left(\frac{\mathbf{i}\,\omega_k}{\alpha}\ln j\right). \tag{2.57}$$

In particular, the first entry $v_k^{(1)}$ of each eigenvector v_k (which is the entry that localizes around the hub node $j = 1$) is

$$v_k^{(1)} = (-1)^{k+1}\,\lambda_k\left(\frac{\alpha}{4} + \frac{\omega_k^2}{\alpha}\right), \tag{2.58}$$

so that, for $j = 2, n$,

$$v_k^{(j)} = v_k^{(1)} \frac{\cos\left(\frac{\mathbf{i}\,\omega_k}{\alpha} \ln j\right)}{\sqrt{j}}. \tag{2.59}$$

The above expressions illustrate how each (leading) eigenvector v_k has a structure where the first entry $v_k^{(1)}$ (located at the hub node $j = 1$) has an amplitude governed by the corresponding eigenvalue λ_k, and the next entries $j > 1$ follow a power-law decay $\sim j^{-1/2}$ with log-periodic oscillations.

§2.4 Bulk of the Spectrum

As we have seen in Fig. 2.1 for the eigenvalues and from Figs. 2.6 2.7 2.8 for the eigenvectors, there is a value of k beyond which the spectral properties of $\mathbf{P}$ are no longer able to describe those of $\mathbf{A}$. We refer to this as the 'entrance' into the random bulk, i.e. the noisy part of the spectrum generated primarily by the fluctuation of $\mathbf{H} = \mathbf{A} - \mathbf{P}$. The analysis of this part of the spectrum is highly non-trivial, and eludes the purposes of this chapter. What we are mainly interested in here is the estimation of where the signal 'meets' the bulk. In particular, we look for an upper bound for the spectral norm of $\mathbf{H}$, whose elements $H_{ij} = A_{ij} - P_{ij}$ are independent but not identically distributed random variables in the range $[-1, 1]$, the variance of H_{ij} being $p_{ij}(1 - p_{ij})$. The study of such so-called random matrices with a variance profile, or structured random matrices, is of great interest in the mathematical ecology community, and has fostered a lively discussion in the mathematical literature [77, 78, 79, 80, 81, 82, 83, 84]. In particular, we are interested in the bound derived in [83], that can be extended to matrices with entries having a different subgaussian distributions, like the centered Bernoulli of our case.

Indeed, let us condition on the latent variables so that the edge probabilities (p_{ij}) are deterministic and $\{A_{ij}\}_{i<j}$ are independent. Let

$$v_{ij} := p_{ij}(1 - p_{ij}), \tag{2.60}$$

$$\sigma := \max_{1 \le i \le n} \Big(\sum_{j \neq i} v_{ij} \Big)^{1/2}, \tag{2.61}$$

$$\sigma_* := \max_{i \neq j} \sqrt{v_{ij}} \le \tfrac{1}{2}. \tag{2.62}$$

Then there exist absolute constants $C, c > 0$ such that

$$\mathbb{E}\,\|\mathbf{H}\| \le C\Big(\sigma + \sigma_* \sqrt{\log n}\Big). \tag{2.63}$$

Moreover, for any $0 < \varepsilon \le \frac{1}{2}$ and any $t \ge 0$,

$$\mathbb{P}\left(\|\mathbf{H}\| \ge (1 + \varepsilon)\, 2\sqrt{2}\,\sigma + t\right) \le n \exp\left(-\frac{t^2}{c_\varepsilon\, \sigma_*^2}\right), \tag{2.64}$$

where $c_\varepsilon > 0$ is universal.

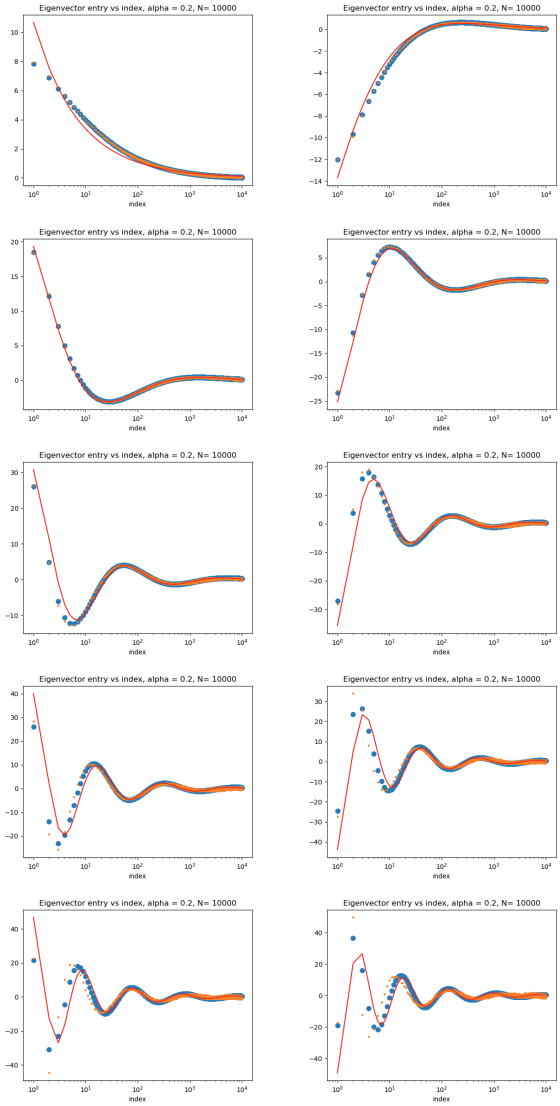


Figure 2.6: Plot of eigenvector entries $v_k^{(j)}$ versus j, for the first 10 eigenvectors $k = 1, \ldots, 10$ (from left to right and top to bottom) and for all entries $j = 1, \ldots, n$. Blue dots correspond to the actual eigenvectors entries of $\boldsymbol{P}$ obtained by solving Eq. (2.29) explicitly, while red solid lines represent the corresponding predicted entries obtained from Eq. (2.56). Orange dots are the entries of the eigenvectors of $\boldsymbol{A}$, and we can see that in the regime outside the bulk they are in quite a good accordance with those of $\boldsymbol{P}$ and hence the ones we obtained analytically. Here $\alpha = 0.2$ and $n = 10^4$.

We can show this by writing $H_{ij} = \sqrt{v_{ij}}\ \xi_{ij}$ with $\xi_{ij} = \big(A_{ij} - p_{ij}\big)/\sqrt{v_{ij}}$. Conditionally on (p_{ij}), the ξ_{ij} are independent, centered, and uniformly sub-Gaussian (bounded support), so the sub-Gaussian extension of the main bound applies with

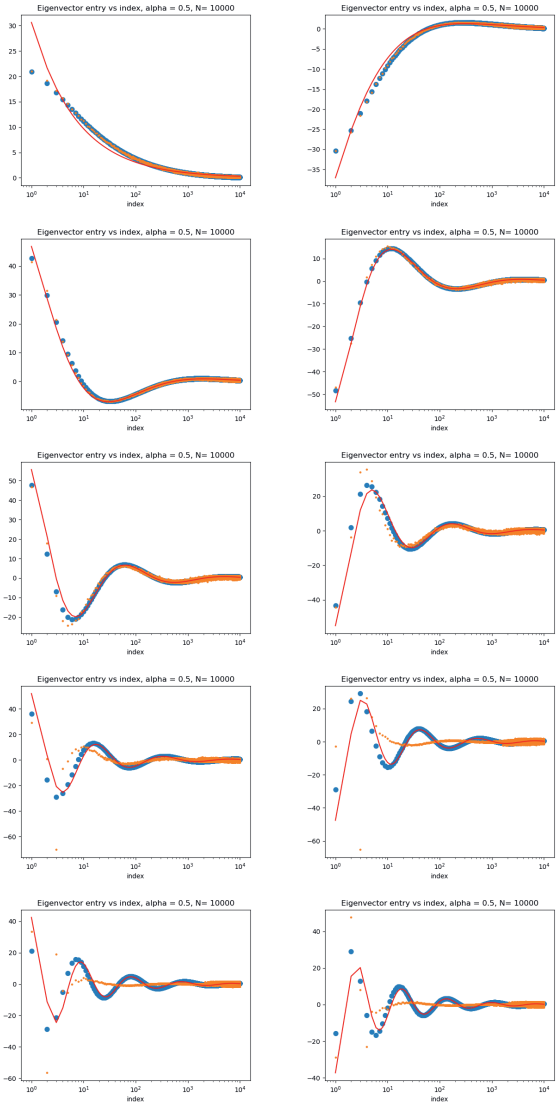


Figure 2.7: Plot of eigenvector entries $v_k^{(j)}$ versus j, for the first 10 eigenvectors $k = 1, \ldots, 10$ (from left to right and top to bottom) and for all entries $j = 1, \ldots, n$. Blue dots correspond to the actual eigenvectors entries of $\boldsymbol{P}$ obtained by solving Eq. (2.29) explicitly, while red solid lines represent the corresponding predicted entries obtained from Eq. (2.56). Orange dots are the entries of the eigenvectors of $\boldsymbol{A}$, and we can see that in the regime outside the bulk they are in quite a good accordance with those of $\boldsymbol{P}$ and hence the ones we obtained analytically. Here $\alpha = 0.5$ and $n = 10^4$.

scale matrix $b_{ij} = \sqrt{v_{ij}}$. This yields (2.63) from the *Gaussian* Theorem 1.1 together with its sub-Gaussian transfer [83](Cor. 3.3):

$$\mathbb{E}\|\mathbf{H}\| \lesssim \sigma + \sigma_* \sqrt{\log n}. \tag{2.65}$$

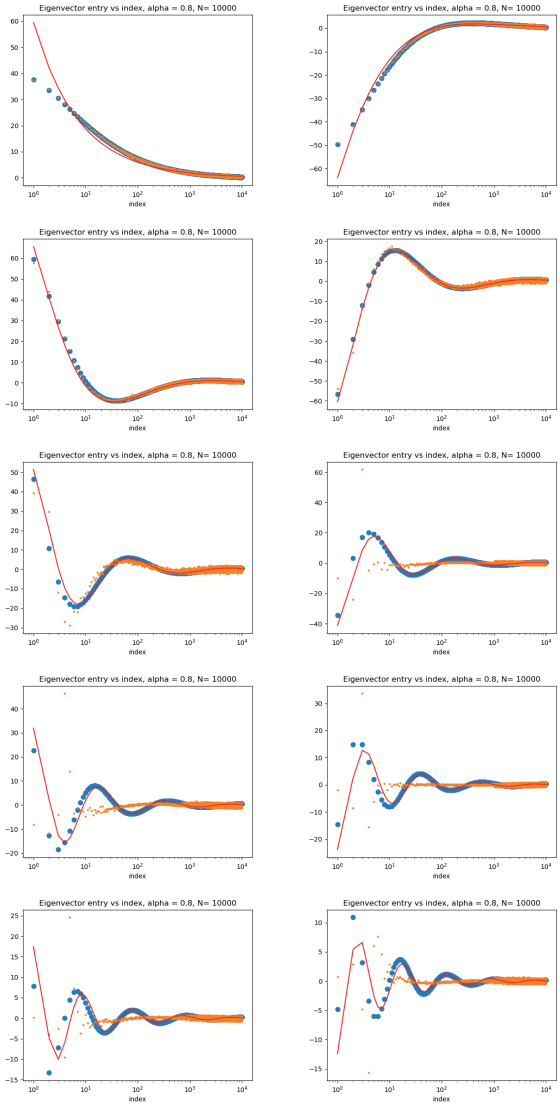


Figure 2.8: Plot of eigenvector entries $v_k^{(j)}$ versus j, for the first 10 eigenvectors $k = 1, \ldots, 10$ (from left to right and top to bottom) and for all entries $j = 1, \ldots, n$. Blue dots correspond to the actual eigenvectors entries of $\boldsymbol{P}$ obtained by solving Eq. (2.29) explicitly, while red solid lines represent the corresponding predicted entries obtained from Eq. (2.56). Orange dots are the entries of the eigenvectors of $\boldsymbol{A}$, and we can see that in the regime outside the bulk they are in quite a good accordance with those of $\boldsymbol{P}$ and hence the ones we obtained analytically. Here $\alpha = 0.8$ and $n = 10^4$.

For tails, note that $|H_{ij}| \leq 1$ and $\mathbb{E}H_{ij} = 0$. Now, [83] (Corollary 3.12) gives, for *symmetric* bounded entries, a variance-sensitive inequality of the form

$$\mathbb{P}(\|\mathbf{H}\| \geq (1+\varepsilon)2\sigma + t) \leq n \exp(-t^2/\tilde{c}_\varepsilon \sigma_*^2). \tag{2.66}$$

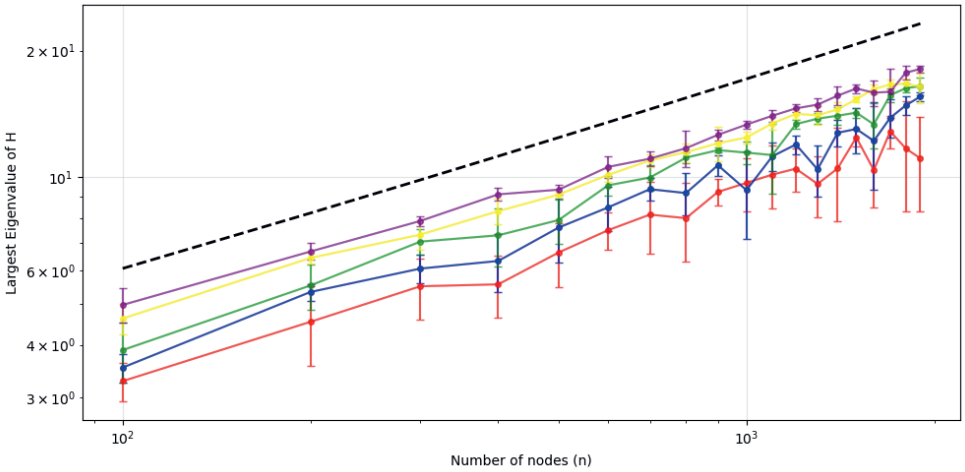


Figure 2.9: Edge of the bulk of the spectrum versus the number n of nodes. Points are the actual edge eigenvalues (largest eigenvalue of $\boldsymbol{H}$, averaged over 10 realizations), for various values of $\alpha = 0.3$ (red), $\alpha = 0.4$ (blue), $\alpha = 0.5$ (green), $\alpha = 0.6$ (yellow) and $\alpha = 0.7$ (purple). The dashed line is the upper bound $\sqrt{n}/2$.

Our entries are not symmetric, but [83](Remark 3.13) shows that symmetrization inflates only the leading constant by a factor $\sqrt{2}$, which yields (2.64). As $v_{ij} \le p_{ij}$, we have $\sigma^2 \le d_{\max}$ where $d_{\max} := \max_i \sum_{j \ne i} p_{ij}$ is the maximal expected degree. Hence

$$\mathbb{E}\|\mathbf{H}\| \le C\big(\sqrt{d_{\max}} + \sqrt{\log n}\big), \tag{2.67}$$

and

$$\|\mathbf{H}\| \le C'\Big(\sqrt{d_{\max}} + \sqrt{\log n}\Big) \quad \text{w.h.p.} \tag{2.68}$$

As a crude but universal bound, since $v_{ij} \le \frac{1}{4}$, $\sigma \le \sqrt{n}/2$, we get

$$\|\mathbf{H}\| \lesssim \tfrac{\sqrt{n}}{2} + O(\sqrt{\log n}) \tag{2.69}$$

In Fig. 2.9 we show that the square root growth (that is commonplace in Wigner matrices) seems to be the right one, and that the coefficient coming from the maximization of the row/column sum of the variance profile matrix is already a quite good envelope for all values of α. Nonetheless this crude bound needs refinement, at least to include the dependence on α. This calls for the full characterization of the bulk spectrum, which we leave to future work. What is important for our purposes here is noticing that the edge of the bulk lives on the same scale $\sqrt{n}$ as the leading eigenvalues given by Eq. (2.31). Since the prefactor of $\sqrt{n}$ for those outliers decays exponentially with ω_k (and, via Eq. (2.36), with $k/\ln n$), we see that the bulk will necessarily 'eat out' the leading eigenvalues for some value of k of order at most $\ln n$.

§2.5 Concluding remarks

In this chapter, we provided a characterization of the spectral properties of the Multi-Scale Model (MSM) in the infinite-mean regime ($0 < \alpha < 1$), which has been recently identified as an annealed random graph model that is exactly invariant under node aggregation [1, 3]. This invariance is a key concept in the field of network renormalization [54], but its connection with discrete scale invariance and the associated idea of complex scaling exponents [35, 85] had not been investigated so far. We showed that the leading eigenvalues of the expected adjacency matrix $\mathbf{P}$ scale as $\sqrt{n}$ and accurately proxy the outlier eigenvalues (whose number is growing and of order at most $\ln n$) of the realized adjacency matrix $\mathbf{A}$. Remarkably, these outliers have an approximately constant spacing on a logarithmic scale of their index k and can be identified as the intersection between the real axis and a *spira mirabilis* in the complex plane, a self-similar object naturally emerging in presence of complex scaling exponents and stable laws with complex stability index [74, 75]. We also derived a closed-form expression for the outlier eigenvectors, showing they are delocalized but biased toward high-weight nodes. The leading eigenvectors beyond the principal (Perron–Frobenius) one exhibit log-periodic oscillations, a feature naturally produced by discrete scale invariance [35, 85, 86] which is here present in the property that the same random graph model describes exactly (up to a parameter rescaling) any coarse-grained version of the original graph, where sets of nodes of the same cardinality are aggregated into supernodes [1, 3]. Our results qualitatively confirm the findings of [33, 34], while partially extending them to the regime of infinite-rank perturbations of random matrices as in more recent attempts [73, 76]. In summary, the MSM illustrates how infinite-mean heterogeneity generates spectral phenomena far from classical paradigms. Our analysis suggests parallels with other models such as Chung–Lu, where broad degree distributions induce hub-driven outliers. Open directions include testing the universality of the $\sqrt{n}$ scaling and log-periodicity across other networks and models. More broadly, our results provide a framework for understanding spectra in multi-scale and heavy-tailed systems, and lay the ground for future work in these fields.

CHAPTER 3

Clustering without Geometry in an Independent Edges Network Model

This chapter is based on:
Alessio Catanzaro, Remco van der Hofstad, and Diego Garlaschelli. *"Clustering without geometry in sparse networks with independent edges"*. arXiv preprint arXiv:2603.13159 (2026).

Abstract

The coexistence of sparsity and clustering (non-vanishing average fraction of triangles per node) is one of the few structural features that, irrespective of finer details, are ubiquitously observed across large real-world networks – motivating the search for generic models producing sparse clustered graphs. Earlier results suggested that sparse random graphs with independent edges fail to reproduce clustering, unless edge probabilities are assumed to depend on underlying metric distances that, thanks to the triangular inequality, naturally favour triadic closure. This observation has opened a debate on whether clustering implies (latent) geometry in real-world networks. Alternatively, recent models of higher-order networks can replicate clustering by abandoning edge independence. Here we prove mathematically, and confirm numerically, that a sparse random graph with independent edges, recently identified in the context of network renormalization as an invariant model under node aggregation, produces finite clustering without any geometric or higher-order constraint. The underlying mechanism is the presence of an infinite-mean node fitness which also implies a power-law degree distribution and, as a novel phenomenon that we characterize rigorously, the breakdown of self-averaging of various network properties. Therefore, as an alternative to geometry or higher-order dependencies, node aggregation invariance emerges as a basic route to realistic network properties.

§3.1 Introduction: clustering versus geometry.

While real-world networks differ in various structural details, the vast majority of them are characterized by virtually ubiquitous features, the most popular of which are the broad (i.e. infinite-variance) distribution of node degrees (numbers of links per node), the *(ultra)small-world* property (i.e. average shortest path length growing (at most) logarithmically with the number of nodes), and the coexistence of *sparsity* (vanishing link density) and *local clustering* (finite average density of triangles around nodes) in the limit of infinitely large networks [4, 6]. While several models based on different generic mechanisms can successfully replicate broad degree distributions and short path lengths [4], the coexistence of sparsity and local clustering is much more difficult to replicate – and hence more informative, as it requires specific mechanisms that can discriminate models at a finer level.

The prototypical example of this difficulty is illustrated by *random graph models with independent edges*, where different pairs of nodes are independently connected. In the Erdős-Rényi (ER) model [10], where edges are independent and identically distributed (i.i.d.) with the same connection probability p, it is easy to show that the link density (expected fraction of realized links) and the local clustering coefficient (defined as the fraction of triangles around a node, averaged over all nodes) have the same expected value p for networks with a large number n of nodes. This automatically implies that, as n diverges, one cannot simultaneously produce finite local clustering and vanishing link density. To overcome this limitation of the ER model, two main routes have been explored.

One route introduces dependencies between edges. This can happen via explicit triadic interactions [9, 87], which however still lead to comparable levels of density and clustering, or via a monopartite projection[1] of a bipartite network [40], which produces a network of overlapping cliques (the edges of which are all mutually dependent) where local clustering can remain very large even for vanishing density. More recent attempts introduce higher-order networks such as simplicial complexes and hypergraphs [88, 89], where edges belonging to the same simplex or hyperedge are mutually dependent. While these more complicated settings can realize the coexistence of sparsity and clustering at the expense of introducing higher-order dependencies, whether that coexistence can be achieved for more parsimonious models with independent edges remains an open question.

The second route explores that question explicitly by preserving edge independence while introducing heterogeneity. More precisely, one can turn the *unconditional* link independence of the ER model into a *conditional* one, by replacing, for each specific pair of nodes i and j ($i, j = 1, \ldots, n$), the constant connection probability p with a probability p_{ij} that is a function $f(w_i, w_j)$ of the realized values w_i, w_j of a certain node-specific variable w (called *fitness* or weight, hidden variable, latent variable, Lagrange multiplier, etc.) that is preliminarly assigned to vertices via an i.i.d.

[1] In monopartite projections of bipartite networks, the original nodes are first initially connected (possibly independently of each other) to 'auxiliary' nodes of a different type (representing for instance the possibile affiliations or memberships of the original nodes), and the auxiliary nodes are then eliminated while connecting the original nodes to each other, if they share connections to the same auxiliary nodes.

draw from a certain probability density function (pdf) $\rho(w)$ [11, 13, 64, 90]. Typical choices of $f(w_i, w_j)$ are monotonically increasing in both arguments (and necessarily symmetric for undirected graphs) – implementing the idea that, the larger the fitness, the larger the probability that a node connects to other nodes[2].

In the studies carried out so far, it turned out that, while several combinations of f and ρ can generate sparse small-world networks with a broad degree distribution [11, 13, 64, 90], they systematically fail to generate finite clustering, casting doubts on whether models with independent edges can replicate local clustering. The only exception that has become popular is that of *random geometric graphs*, where the fitness w acquires the role of a *node coordinate* (which, more in general, can be a D-dimensional vector) and the connection function f depends on w_i and w_j through some *metric distance* $d_{ij} = d(w_i, w_j)$ (in which case, f generally loses the property of being monotonic in w_i[3]). A remarkable example of this success is the *hyperbolic model* [14, 15, 16, 17, 18, 19, 21] which assumes that nodes are sprinkled uniformly at random in a D-dimensional hyperbolic space and links are formed with a probability that is a certain decreasing function of the geodesic distance between nodes. In this setting, the triangular inequality ($d_{ij} \leq d_{ik} + d_{jk}\ \forall i, j, k$) guarantees that pairs of nodes i, j that successfully connect to a third node k are very likely to connect to each other (*triadic closure*), provably leading to finite clustering [91, 92]. Additionally, hyperbolicity ensures that other desirable properties such as power-law degree distributions emerge [16]. These investigations have convincingly concluded that, in random graphs with conditionally independent edges and distance-dependent connection probabilities, *geometry implies clustering*.

On the other hand, the potential reverse (non-obvious) implication that *clustering implies geometry*, i.e. that the presence of finite clustering could suffice to indicate that the network is (with high probability) the realization of a random geometric graph, is highly debated [14, 15, 16, 17, 18, 19, 21]. Whether clustering implies geometry is crucial for network theory (because, if true, it would demand for explanations and interpretations of the hidden geometry of real-world networks [21, 91]), for data science (for instance because techniques such as graph embedding and topological data analysis could then be optimized for real-world network geometries) and for the control of processes on networks (because geometry could aid network navigability [16, 93, 94] and the understanding of how clustering impacts epidemic and other processes [9]).

In this chapter, we prove that finite clustering can emerge independently of geometry or higher-order constraints in random graphs with independent edges conditional on infinite-mean fitness [1, 2]. While the finite-mean case has been systematically found to produce vanishing clustering, we show that in the infinite-mean case clustering is finite and strong. Notably, we find that the infinite-mean fitness also implies the breakdown of self-averaging in certain properties, including clustering itself.

[2]Some popular examples are $f(w_i, w_j) = zw_iw_j$, $f(w_i, w_j) = \Theta(w_i + w_j - z)$, $f(w_i, w_j) = zw_iw_j/(1 + zw_iw_j)$, and $f(w_i, w_j) = 1 - e^{-zw_iw_j}$, where $z > 0$ is a global parameter controlling for the overall link density and $w_i \geq 0\ \forall i$.

[3]For instance, if w_i and w_j are scalar (unidimensional) coordinates and their distance is defined as $d_{ij} = |w_i - w_j|$, then w_i and w_j can simultaneously increase while keeping d_{ij} (and hence p_{ij}) unchanged.

§3.2 The Multi-Scale Model.

We consider the inhomogeneous random graph model introduced in [1] and studied in [2, 3, 95]. Each node i is assigned a positive weight (fitness) w_i $(i = 1, n)$, sampled i.i.d. from a certain pdf $\rho(w)$. Given w_i, w_j, nodes i, j connect with probability

$$p_{ij} = 1 - e^{-\delta_n w_i w_j}, \tag{3.1}$$

where the scaling δ_n can be chosen in such a way that the link density vanishes as $n \to \infty$. The connection probability (3.1) is also known in the literature as the Norros–Reittu model [96] and has been used for various purposes [97, 98], by considering a finite-mean $\rho(w)$. In this work we are however interested in the specific 'annealed' regime considered within the framework of network renormalization [54], where the model emerges as a fixed point of a renormalization flow in the space of models, induced by coarse-graining the network by aggregating nodes into super-nodes and recalculating the induced connection probability between super-nodes [1]. The specific connection rule in Eq. (3.1) makes the model invariant, provided the weight of each block is defined as (a random variable equal to) the sum of the weights of the constituent nodes [1]. If one further demands the invariance of the weight distribution, i.e. that the coarse-grained weights are distributed according to the same pdf as the constituent weights, then $\rho(w)$ should be a positively-supported, one-sided stable law $\tilde{\rho}_\alpha(w)$ with stability parameter $0 < \alpha < 1$ [99, 22] and tails decaying as $w^{-1-\alpha}$ for large w, hence with diverging mean and higher moments [1, 3]. In this way, one gets a Multi-Scale Model (MSM) interpretable as describing connections between arbitrarily nested (blocks of) nodes, thereby describing the graph at multiple scales of resolution (coarse-graining) simultaneously [1].

Since a stable law $\tilde{\rho}_\alpha(w)$ with $0 < \alpha < 1$ is known in closed form only for $\alpha = 1/2$ (Lévy distribution), we first follow [3] and replace it with a pure Pareto distribution

$$\rho_\alpha(w) = \alpha w^{-1-\alpha}, \quad w \geq 1,\ 0 < \alpha < 1 \tag{3.2}$$

(which has exactly the same tail behaviour as an α-stable distribution and serves as an analytically tractable counterpart [3, 99]) and then discuss the extension to actual α-stable laws $\tilde{\rho}_\alpha(w)$ via both analytical and numerical arguments. Given (3.2), one can show that choosing $\delta_n = n^{-1/\alpha}$ implies that the link density vanishes as $\log n/n$, i.e. the average degree grows only as $\log n$, and the degree distribution is power-law [3].

We stress that the MSM can be enriched with additional dyadic features [1, 2] which allow the connection probabilities to depend, for instance, on geometric distances between coordinates assigned to nodes. Nonetheless, to investigate whether clustering can be produced *without* geometry, here we obviously consider the simplest, distance-independent version in Eq. (3.1).

§3.3 The Local Clustering Coefficient.

Abstractly, the clustering coefficient quantifies the tendency of a graph to form closed triangles. Globally, it can be defined as the ratio between the total number of closed

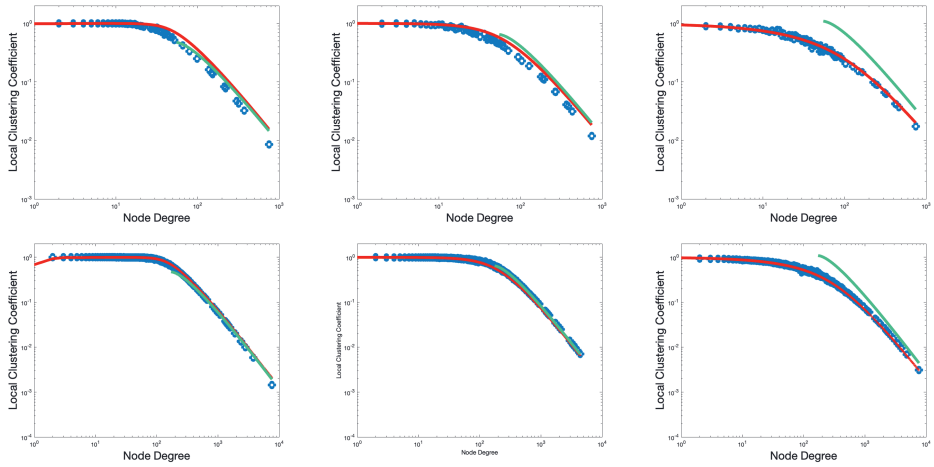


Figure 3.1: ***C****lustering functions for different values of* n *and* α *for the MSM model* (3.1) *with Pareto weights* (3.2)*. Blue circles: empirical clustering function* (3.4) *versus the degree* k *computed on actual realized graphs (obtained by sampling the weights once, and sampling the graph once conditionally on the realized weights). Red curves: our analytical expression* (3.6) *for the annealed clustering function evaluated on the actual degree* $k = a\sqrt{n}$*. Green curves: our asymptotic calculation* (3.44) *valid for diverging reduced degrees. We evaluate and plot all functions only for degrees larger than* 1*, to avoid ambiguities in the definition of the clustering coefficient for* $k < 2$*. Top:* $n = 10^3$*, bottom:* $n = 10^4$*. From left to right:* $\alpha = 0.3, 0.5, 0.7$*.*

triangles and the number of potential (either open or closed) triangles (also called cherries, or wedges). In our regime with $0 < \alpha < 1$, this ratio vanishes for the MSM model [3], in line with what observed in sparse real-world networks. Our interest here is the (more distinctive) *local* clustering coefficient C_v, which quantifies the tendency of the neighbours of the specific node v to form closed triangles around v. For a node v with degree D_v, C_v is defined as the ratio between the number Δ_v of triangles containing v and the number of wedges centered at v. In terms of the adjacency matrix $\mathbf{A} = (A_{ij})_{i,j=1}^n$, this reads

$$C_v = \frac{\Delta_v}{\binom{D_v}{2}} = \frac{\sum_{i \neq v} \sum_{j \neq v,i} A_{vi} A_{ij} A_{jv}}{D_v(D_v - 1)}, \quad D_v > 1. \tag{3.3}$$

The behaviour of C_v versus D_v, as well as the average value C of C_v over nodes, can effectively characterize real-world networks and discriminate among different models – especially in determining whether a finite (expected) value of C can emerge in large sparse graphs, as observed in real-world networks. Note that C_v in Eq. (3.3) is defined only for nodes with $D_v > 1$, as only nodes with at least two neighbours can host a triangle. For nodes with $D_v = 0, 1$, one can either leave C_v undefined (and exclude it from the calculation of C) or set conventionally $C_v \equiv 0$ (and include it in the calculation of C). We will consider both cases and discuss their implications.

§3.3.1 Clustering Function.

For a given graph, we consider the average clustering coefficient as a function of the degree k of each node, i.e., the *empirical clustering function*

$$C(k) = \frac{1}{N_k} \sum_{v:\, D_v=k} \frac{\Delta_v}{\binom{k}{2}}, \tag{3.4}$$

where $N_k = \sum_{v:\, D_v=k} \delta_{D_v,k}$ is the number of nodes with degree k. In terms of expectations ($\mathbb{E}$) over the model, we introduce the *annealed clustering function* as

$$\bar{C}(k) = \frac{1}{\mathbb{E}[N_k]}\, \mathbb{E}\left[\sum_{v:\, D_v=k} \frac{\Delta_v}{\binom{k}{2}} \right]. \tag{3.5}$$

One of our main results is a rigorous calculation of $\bar{C}(k)$ in the asymptotic (large n) limit, valid for both small and large k, which shows excellent agreement with the empirical clustering function $C(k)$ [4]. In this setting, we redefine $\bar{C}(\cdot)$ as a function of the *reduced degree* $a = k/\sqrt{n}$, as this rescaling allows us to prove that, defining $g(x) \equiv 1 - e^{-x}$, the limiting form of $\bar{C}(a)$ as $n \to \infty$ is

$$\bar{C}(a) = \frac{\alpha^2}{a^2} \int_0^\infty \mathrm{d}x \int_0^\infty \mathrm{d}y\, \frac{g(a^{1/\alpha}\tau_\alpha x)}{x^{1+\alpha}} \frac{g(a^{1/\alpha}\tau_\alpha y)}{y^{1+\alpha}} g(xy), \tag{3.6}$$

where $\tau_\alpha \equiv \Gamma(1-\alpha)^{-1/\alpha}$, and $\Gamma(\cdot)$ is the Euler Gamma function. The proof is in section 3.3.2. In Figure 3.1, we compare the numerical evaluation of the integral (3.6) with the direct computation of the empirical clustering function (3.4) on the actual realizations of the random graph. The agreement is remarkably good, already for moderate values of n.

§3.3.2 Limiting expression for local clustering function

In this section we prove that the clustering function of the MSM indeed converges in the large n limit to the expression (3.6).

Theorem 3.3.1.
Let the vertex space of the network be $[n] = \{1, \dots, n\}$. Let $(W_v)_{v\in[n]}$ be i.i.d. Pareto random variables, so that their density equals

$$\rho_\alpha(w) = \frac{\alpha}{w^{\alpha+1}}, \qquad w \geq 1, \tag{3.7}$$

with $\alpha \in (0,1)$, so that $\mathbb{E}[W] = \infty$. Conditionally on $(W_v)_{v\in[n]}$, edges are present independently, with the edge between $u, v \in [n]$ being present with probability

$$p_{uv} = 1 - e^{-W_u W_v / n^{1/\alpha}}. \tag{3.8}$$

[4] We also believe that, through a second-moment analysis, it could be proven that in the large n limit, with high probability, the two clustering functions become the same. We will however not prove this claim in the present work.

Throughout, we let $(I_{u,v})_{1\le u<v\le n}$ *denote the edge indicator variables (or the elements of the adjacency matrix), so that*

$$\mathbb{P}\big(I_{u,v}=1 \mid (W_v)_{v\in[n]}\big)=1-e^{-W_uW_v/n^{1/\alpha}}, \tag{3.9}$$

and $(I_{u,v})_{1\le u<v\le n}$ *are conditionally independent given* $(W_v)_{v\in[n]}$*. Let the Annealed Clustering Function be defined as*

$$\bar{C}(k)=\frac{1}{\mathbb{E}\left[N_k\right]}\,\mathbb{E}\left[\sum_{v/D_v=k}\frac{\Delta_v}{\binom{k}{2}}\right] \tag{3.10}$$

Then, as $n\to\infty$*, for* $k=a\sqrt{n}$*,*

$$\bar{C}_n(a\sqrt{n})\to\bar{C}(a)=\frac{\alpha^2}{a^2}\int_0^\infty\int_0^\infty[1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}][1-e^{-xy}](xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y, \tag{3.11}$$

where $\tau_\alpha=\Gamma(1-\alpha)^{-1/\alpha}$*.*

Proof. We compute

$$\mathbb{E}[N_k]=n\,\mathbb{E}\left[\mathbb{P}(D_1=k\mid W_1)\right], \tag{3.12}$$

where we note that, conditionally on $W_1=w$, $D_1\sim\mathrm{Bin}\left(n-1,q_w\right)$, where

$$q_w=\mathbb{E}\left[1-e^{-wW/n^{1/\alpha}}\right]. \tag{3.13}$$

We also compute the numerator

$$\mathbb{E}\Big[\sum_{v\in[n]}\Delta_v\mathbb{1}_{D_v=k}\Big]=n\,\mathbb{E}\Big[\mathbb{E}\Big[\Delta_1\mathbb{1}_{D_1=k}\mid W_1\Big]\Big] \tag{3.14}$$

$$=n\binom{n-1}{2}\,\mathbb{E}\Big[\mathbb{P}(D_1'=k-2\mid W_1)\mathbb{P}(I_{1,2}I_{1,3}I_{2,3}=1\mid W_1)\Big], \tag{3.15}$$

where $D_1'\sim\mathrm{Bin}(n-3,q_w)$, and we have used conditional independence. Here, we can identify

$$D_1'=\sum_{u\in[n]\setminus[3]}I_{1,u}. \tag{3.16}$$

We see that all these expectations involve binomial random variables, which we continue to investigate, starting with the conditional success probability q_w.

Analysis of success probabilities We compute

$$q_w=\mathbb{E}\left[1-e^{-wW/n^{1/\alpha}}\right] \tag{3.17}$$

$$=\int_1^\infty[1-e^{-ws/n^{1/\alpha}}]\frac{\alpha}{s^{\alpha+1}}\mathrm{d}s \tag{3.18}$$

$$=(\frac{w}{n^{1/\alpha}})^\alpha\int_{wn^{-1/\alpha}}^\infty[1-e^{-s}]\frac{\alpha}{s^{\alpha+1}}\mathrm{d}s$$

$$=\frac{w^\alpha}{n}\int_0^\infty[1-e^{-s}]\frac{\alpha}{s^{\alpha+1}}\mathrm{d}s-\frac{w^\alpha}{n}\int_0^{wn^{-1/\alpha}}[1-e^{-s}]\frac{\alpha}{s^{\alpha+1}}\mathrm{d}s.$$

We next use that

$$\int_0^\infty [1-e^{-s}]\frac{\alpha}{s^{\alpha+1}}\mathrm{d}s = \int_0^\infty \int_0^s e^{-y}\mathrm{d}y\frac{\alpha}{s^{\alpha+1}}\mathrm{d}s \tag{3.19}$$

$$= \int_0^\infty e^{-y}\int_y^\infty \frac{\alpha}{s^{\alpha+1}}\mathrm{d}s\mathrm{d}y \tag{3.20}$$

$$= \int_0^\infty e^{-y}y^{-\alpha}\mathrm{d}y = \Gamma(1-\alpha), \tag{3.21}$$

and

$$\int_0^{wn^{-1/\alpha}} [1-e^{-s}]\frac{\alpha}{s^{\alpha+1}}\mathrm{d}s \le \int_0^{wn^{-1/\alpha}} \frac{\alpha}{s^{\alpha}}\mathrm{d}s \tag{3.22}$$

$$= \frac{\alpha}{\alpha-1}\big(wn^{-1/\alpha}\big)^{1-\alpha} \tag{3.23}$$

$$= \frac{\alpha}{\alpha-1}w^{1-\alpha}n^{1-1/\alpha}. \tag{3.24}$$

This gives excellent control over q_w for all $w = o(n^{1/\alpha})$.

Binomial local limit approximations Let $S_m \sim \mathrm{Bin}(m,p)$. Then, a local central limit approximation shows that

$$\mathbb{P}(S_m = k) = \frac{1}{\sqrt{2\pi mp(1-p)}}\mathrm{e}^{-(k-mp)^2/[2mp(1-p)]}(1+o(1)), \tag{3.25}$$

when $|k-mp| = o((mp)^{2/3})$, and uniformly in k and p. Such approximations have a long history, we refer to the books by Bollobás [100] and Janson, Łuczak and Ruciński [101]. In particular, the upper bound in (3.25) follows from [100, Theorem 1.2], the lower bound from [100, Theorem 1.5], and an analysis of the error terms present there, showing that they are $o(1)$ when $|n-mp| = o((mp)^{2/3})$.

In our case, we have that $m = n-1$ or $m = n-2$, $p = q_w$ with $w = W_1$, so that

$$\mathbb{P}(D_1 = k \mid W_1 = w) = \frac{1}{\sqrt{2\pi mq_w(1-q_w)}}\mathrm{e}^{-(k-mq_w)^2/[2mq_w(1-q_w)]}(1+o(1)), \tag{3.26}$$

when $|k-mq_w| = o((mq_w)^{2/3})$, and uniformly in k and w. Since, for us, k is fixed, while w is being integrated out over, the above means that the main contribution for $W_1 = w$ comes from w for which $mq_w \approx k$, with values of w for which

$$|k-mq_w| \ge C\sqrt{k}\log k \tag{3.27}$$

receiving probability at most

$$\mathrm{e}^{-C\log k} = k^{-C}, \tag{3.28}$$

which is negligible. Thus, we can safely think of $mq_w = k$, which means

$$mq_w = w^\alpha\Gamma(1-\alpha)(1+o(1)) = k, \tag{3.29}$$

so that

$$w = k^{1/\alpha}\Gamma(1-\alpha)^{-1/\alpha}(1+o(1)). \tag{3.30}$$

This implies that, with $m = n-1$,

$$\begin{aligned}\mathbb{E}[N_k] &= n\,\mathbb{E}\left[\mathbb{P}(D_1 = k \mid W_1)\right] \qquad (3.31)\\ &= (1+o(1))n\int_1^\infty \frac{1}{\sqrt{2\pi m q_w(1-q_w)}}\mathrm{e}^{-(k-mq_w)^2/[2mq_w(1-q_w)]}\frac{\alpha}{w^{\alpha+1}}\mathrm{d}w\\ &= (1+o(1))\int_1^\infty \frac{1}{\sqrt{2\pi n q_w}}\mathrm{e}^{-(k-nq_w)^2/[2nq_w]}\frac{\alpha}{w^{\alpha+1}}\mathrm{d}w,\end{aligned}$$

as long as $nq_w \gg 1$, and thus, since $k \approx nq_w$, also $k \gg 1$. Since the above integral is dominated by nq_w that are close to k, we may replace nq_w by k everywhere, except in the $(k-nq_w)^2$ term. This gives

$$\mathbb{E}[N_k] = (1+o(1))\frac{n}{\sqrt{2\pi k}}\int_1^\infty \mathrm{e}^{-(k-nq_w)^2/[2k]}\frac{\alpha}{w^{\alpha+1}}\mathrm{d}w. \tag{3.32}$$

We next substitute $nq_w \approx w^\alpha\Gamma(1-\alpha)$ to arrive at

$$\begin{aligned}\mathbb{E}[N_k] &= (1+o(1))\frac{n}{\sqrt{2\pi k}}\int_1^\infty \mathrm{e}^{-(k-w^\alpha\Gamma(1-\alpha))^2/[2k]}\frac{\alpha}{w^{\alpha+1}}\mathrm{d}w \qquad (3.33)\\ &= (1+o(1))\frac{n\alpha}{w_k^{\alpha+1}\sqrt{2\pi k}}\int_1^\infty \mathrm{e}^{-(k-w^\alpha\Gamma(1-\alpha))^2/[2k]}\mathrm{d}w,\end{aligned}$$

where w_k solves $w^\alpha\Gamma(1-\alpha) = k$, and thus

$$w_k = k^{1/\alpha}\Gamma(1-\alpha)^{-1/\alpha}. \tag{3.34}$$

We then let $z = w^\alpha\Gamma(1-\alpha)$, so that $\mathrm{d}z = \alpha w^{\alpha-1}\Gamma(1-\alpha)\mathrm{d}w$, so that

$$\mathrm{d}w = \Gamma(1-\alpha)^{-1}(z/\Gamma(1-\alpha))^{(1-\alpha)/\alpha}\mathrm{d}z = z^{(1-\alpha)/\alpha}\Gamma(1-\alpha)^{-1/\alpha}\mathrm{d}z, \tag{3.35}$$

so that

$$\begin{aligned}\mathbb{E}[N_k] &= (1+o(1))\frac{n\alpha}{w_k^{\alpha+1}\sqrt{2\pi k}}\int_1^\infty \mathrm{e}^{-(k-z)^2/[2k]}z^{(1-\alpha)/\alpha}\Gamma(1-\alpha)^{-1/\alpha}\mathrm{d}z \qquad (3.36)\\ &= (1+o(1))\frac{n\alpha}{w_k^{\alpha+1}\sqrt{2\pi k}}\Gamma(1-\alpha)^{-1/\alpha}k^{(1-\alpha)/\alpha}\int_1^\infty \mathrm{e}^{-(k-z)^2/[2k]}\mathrm{d}z\\ &= (1+o(1))\frac{n\alpha}{w_k^{\alpha+1}\sqrt{2\pi k}}k^{(1-\alpha)/\alpha}\Gamma(1-\alpha)^{-1/\alpha}\sqrt{2\pi k}\\ &= n(1+o(1))\alpha k^{-(\alpha+1)/\alpha}\Gamma(1-\alpha)^{(\alpha+1)/\alpha}k^{(1-\alpha)/\alpha}\Gamma(1-\alpha)^{-1/\alpha}\\ &= n(1+o(1))\frac{\alpha\Gamma(1-\alpha)}{k^2}.\end{aligned}$$

This confirms that $\mathbb{E}[N_k] = \Theta(1)$ when $k = \Theta(\sqrt{n})$, so that one cannot expect that $N_k/\mathbb{E}[N_k] \xrightarrow{p} 1$ for such values of k.

Since (3.25) depends on k, for k large, in a relatively weak way, this means that, for values of w close to w_k,

$$\mathbb{P}(D_1' = k-2 \mid W_1 = w) \approx \mathbb{P}(D_1 = k \mid W_1 = w). \tag{3.37}$$

Thus, since such factors appear both in $\mathbb{E}[N_k]$ and $\mathbb{E}\left[\sum_{v\in[n]} \Delta_v \mathbb{1}_{D_v=k}\right]$, they cancel out asymptotically, so that

$$\frac{\mathbb{E}\left[\sum_{v\in[n]} \Delta_v \mathbb{1}_{D_v=k}\right]}{\mathbb{E}[N_k]} \approx n\binom{n-1}{2}\mathbb{P}(I_{1,2}I_{1,3}I_{2,3} = 1 \mid W_1 = w_k). \tag{3.38}$$

Recalling (3.10), this then shows that

$$\bar{C}_n(k) = (1+o(1))\frac{\binom{n-1}{2}}{\binom{k}{2}}\mathbb{P}(I_{1,2}I_{1,3}I_{2,3} = 1 \mid W_1 = w_k). \tag{3.39}$$

We now take $k = a\sqrt{n}$, and obtain

$$\bar{C}_n(a\sqrt{n}) = (1+o(1))\frac{n\alpha^2}{a^2}\int_1^\infty\int_1^\infty [1-e^{-xw_k/n^{1/\alpha}}][1-e^{-yw_k/n^{1/\alpha}}][1-e^{-xy/n^{1/\alpha}}](xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y. \tag{3.40}$$

We first note that, for $k = a\sqrt{n}$,

$$w_{a\sqrt{n}} = n^{1/(2\alpha)}a^{1/\alpha}\Gamma(1-\alpha)^{-1/\alpha} \equiv n^{1/(2\alpha)}a^{1/\alpha}\tau_\alpha. \tag{3.41}$$

We rescale x and y by $n^{-1/2\alpha}$ to obtain

$$\bar{C}_n(a\sqrt{n}) = (1+o(1))\frac{\alpha^2}{a^2}\int_{n^{-1/2\alpha}}^\infty\int_{n^{-1/2\alpha}}^\infty [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}][1-e^{-xy}](xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y \tag{3.42}$$

$$= (1+o(1))\frac{\alpha^2}{a^2}\int_0^\infty\int_0^\infty [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}][1-e^{-xy}](xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y,$$

as predicted in (3.11).

§3.4 Hub and leaf behaviour.

To refine the characterization of clustering in the model, we investigate the limiting behavior of the annealed clustering function for nodes in two distinct regimes of the degree spectrum: *leaf* and *hub* nodes, corresponding to poorly and highly connected ones, respectively. For leaf (L) nodes, defined as nodes with vanishing reduced degree $a \to 0$, we find that the function in Eq. (3.6) has the asymptotic plateau

$$\bar{C}_L \equiv \lim_{a\to 0}\bar{C}(a) = 1, \tag{3.43}$$

which implies that in the low-degree limit the local neighborhood of a node is almost fully clustered (excluding of course nodes with $k = 0, 1$). This behaviour can be

intuitively explained in terms of an interconnected core of hubs coexisting with a majority of low-degree nodes, whose few links are largely connected with the hubs. Therefore, low-degree nodes have a high chance to close triangles between neighbours because the hubs are densely connected among themselves. For hub (H) nodes, defined as nodes with diverging reduced degree $a \to \infty$, we find (see SI) that the function in Eq. (3.6) decays as

$$\bar{C}(a) \sim \bar{C}_H(a) \equiv 2\Gamma(1-\alpha)\frac{\log a}{a^2}, \quad a \to \infty. \tag{3.44}$$

Figure 3.1 shows that the asymptotics (3.43) and (3.44) are in excellent agreement both with the empirical clustering function (3.4) computed on actual realizations of the random graph and with the full analytical expression (3.6) for the annealed clustering function (3.5). We now prove these claims.

§3.4.1 Asymptotics of the clustering function

Lemma 3.4.1 ($\bar{C}(a)$ is bounded by 1).
For all $a > 0$, $\bar{C}(a) \leq 1$.

Proof. Estimating $1 - e^{-xy} \leq 1$ bounds $\bar{C}(a)$ by

$$\bar{C}(a) \leq \frac{\alpha^2}{a^2}\Big(\int_0^\infty [1 - e^{-xa^{1/\alpha}\tau_\alpha}]x^{-(\alpha+1)}\mathrm{d}x\Big)^2, \tag{3.45}$$

for which we note that

$$\begin{aligned}
\int_0^\infty [1 - e^{-xa^{1/\alpha}\tau_\alpha}]x^{-(\alpha+1)}\mathrm{d}x &= \int_0^\infty \int_0^{xa^{1/\alpha}\tau_\alpha} e^{-y}\mathrm{d}y x^{-(\alpha+1)}\mathrm{d}x && (3.46)\\
&= \int_0^\infty e^{-y}\int_{ya^{-1/\alpha}/\tau_\alpha}^\infty x^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y && (3.47)\\
&= \frac{1}{\alpha}\int_0^\infty e^{-y}(ya^{-1/\alpha}/\tau_\alpha)^{-\alpha}\mathrm{d}y\\
&= \frac{a\tau_\alpha^\alpha}{\alpha}\int_0^\infty e^{-y}y^{-\alpha}\mathrm{d}y = \frac{a\tau_\alpha^\alpha}{\alpha}\Gamma(1-\alpha)\\
&= \frac{a}{\alpha},
\end{aligned}$$

using that $\tau_\alpha = \Gamma(1-\alpha)^{-1/\alpha}$. Thus, indeed, $\bar{C}(a) \leq 1$.

We next investigate $\bar{C}(a)$ for $a \ll 1$:

Lemma 3.4.2 ($\bar{C}(a)$ is close to 1 for a small).
As $a \searrow 0$,

$$\bar{C}(a) = 1 + o(1). \tag{3.48}$$

Proof. For $a \ll 1$, we write

$$\bar{C}(a) = \frac{\alpha^2}{a^2}\int_0^\infty\int_0^\infty [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}](xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y \tag{3.49}$$
$$-\frac{\alpha^2}{a^2}\int_0^\infty\int_0^\infty [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}]e^{-xy}(xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y.$$

The first integral equals 1 by Lemma 3.4.1, so that

$$\bar{C}(a) = 1-\bar{C}_1(a), \tag{3.50}$$

where

$$\bar{C}_1(a) = \frac{\alpha^2}{a^2}\int_0^\infty\int_0^\infty [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}]e^{-xy}(xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y. \tag{3.51}$$

We first bound the integrals where either x or y is in $[0, a^{-\varepsilon}]$, for some $\varepsilon > 0$ sufficiently small, as

$$\begin{aligned}\frac{2\alpha^2}{a^2}&\int_0^\infty\int_0^{a^{-\varepsilon}} [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}]e^{-xy}(xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y \\ &\le \frac{2\alpha^2}{a^2}\int_0^\infty [1-e^{-xa^{1/\alpha}\tau_\alpha}]x^{-(\alpha+1)}\mathrm{d}x a^{1/\alpha}\int_0^{a^{-\varepsilon}} y^{-\alpha}\mathrm{d}x \\ &\le \frac{2\alpha}{a(\alpha-1)}a^{1/\alpha-\varepsilon(1-\alpha)} = o(1),\end{aligned} \tag{3.52}$$

using (3.46) and the fact that $1/\alpha - \varepsilon(1-\alpha) > 1$ for $\varepsilon > 0$ sufficiently small since $\alpha \in (0,1)$. Thus,

$$\bar{C}_1(a) = o(1) + \frac{\alpha^2}{a^2}\int_{a^{-\varepsilon}}^\infty\int_{a^{-\varepsilon}}^\infty [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}]e^{-xy}(xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y, \tag{3.53}$$

which is bounded by $e^{-a^{-2\varepsilon}} = o(1)$. This shows that $\bar{C}_1(a) = o(1)$ for $a \searrow 0$, and thus proves that $\bar{C}(a) = 1 + o(1)$.

For $a \to \infty$, the integral should be dominated by small x, y, so that one might consider to expand $1 - e^{-xy} = xy$, to obtain

$$\begin{aligned}\bar{C}(a) &= (1+o(1))\frac{\alpha^2}{a^2}\int_0^\infty\int_0^\infty [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}](xy)^{-\alpha}\mathrm{d}x\mathrm{d}y \\ &= (1+o(1))\frac{\alpha^2}{a^2}\big(a^{1/\alpha}\tau_\alpha\big)^{2\alpha-2}\Big(\int_0^\infty [1-e^{-x}]x^{-\alpha}\mathrm{d}x\Big)^2.\end{aligned} \tag{3.54}$$

However, the integral that is squared is infinite, so that we need to proceed alternatively. In fact, this analysis is quite intricate, and a careful split is necessary. The main result is as follows:

Lemma 3.4.3 (Large a asymptotics of $\bar{C}(a)$).
As $a \nearrow \infty$,

$$\bar{C}(a) = 2\Gamma(1-\alpha)\frac{\log a}{a^2} + O(a^{-2}). \tag{3.55}$$

Proof. We again split the integral conveniently. We split the integral depending on whether x or y is in $[0, a^{-1/\alpha}/\tau_\alpha]$ or not. This gives

$$\bar{C}(a) = \bar{C}_1(a) + \bar{C}_2(a) + \bar{C}_3(a), \tag{3.56}$$

where

$$\bar{C}_1(a) = \frac{\alpha^2}{a^2} \int_{a^{-1/\alpha}/\tau_\alpha}^{\infty} \int_{a^{-1/\alpha}/\tau_\alpha}^{\infty} [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}][1-e^{-xy}](xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y, \tag{3.57}$$

$$\bar{C}_2(a) = \frac{\alpha^2}{a^2} \int_{0}^{a^{-1/\alpha}/\tau_\alpha} \int_{0}^{a^{-1/\alpha}/\tau_\alpha} [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}][1-e^{-xy}](xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y. \tag{3.58}$$

We bound, again using (3.46) and $1 - e^{-xy} \le xy$,

$$\begin{aligned}\bar{C}_2(a) &\le \frac{\alpha^2}{a^2} \int_{0}^{a^{-1/\alpha}/\tau_\alpha} \int_{0}^{a^{-1/\alpha}/\tau_\alpha} [1 - e^{-xa^{1/\alpha}\tau_\alpha}][1 - e^{-ya^{1/\alpha}\tau_\alpha}](xy)^{-\alpha}\mathrm{d}x\mathrm{d}y \qquad (3.59)\\ &= \frac{\alpha^2}{a^2}\Big(\int_{0}^{a^{-1/\alpha}/\tau_\alpha} y^{-\alpha}\mathrm{d}y\Big)^2 \\ &= \Theta(1)a^{-2}(a^{-1/\alpha})^{2(1-\alpha)} = o(a^{-2}).\end{aligned}$$

We continue with $\bar{C}_3(a)$, which is the most involved. We bound $1 - e^{-ya^{1/\alpha}\tau_\alpha} \le ya^{1/\alpha}$ to obtain

$$\bar{C}_3(a) \le \frac{2\alpha^2 a^{1/\alpha}}{a^2} \int_{0}^{a^{-1/\alpha}/\tau_\alpha} \int_{a^{-1/\alpha}/\tau_\alpha}^{\infty} (xy)^{-(\alpha+1)} y[1 - e^{-xy}]\mathrm{d}x\mathrm{d}y. \tag{3.60}$$

We bound $1 - e^{-xy} \le (xy) \wedge 1$. The integral where $x \le 1/y$ is bounded by

$$\Theta(1)a^{-2+1/\alpha} \int_{0}^{a^{-1/\alpha}/\tau_\alpha} y^{-\alpha+1} \int_{0}^{1/y} x^{-\alpha}\mathrm{d}x\mathrm{d}y \le \Theta(1)a^{-2+1/\alpha} \int_{0}^{a^{-1/\alpha}/\tau_\alpha} \mathrm{d}y \le \Theta(1)a^{-2}. \tag{3.61}$$

Similarly, the integral where $x > 1/y$ is bounded by

$$\Theta(1)a^{-2+1/\alpha} \int_{0}^{a^{-1/\alpha}/\tau_\alpha} y^{-\alpha} \int_{1/y}^{\infty} x^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y \le \Theta(1)a^{-2+1/\alpha} \int_{0}^{a^{-1/\alpha}/\tau_\alpha} \mathrm{d}y \le \Theta(1)a^{-2}. \tag{3.62}$$

This shows that $\bar{C}_3(a) = O(a^{-2})$.

We finally compute $\bar{C}_1(a)$, which we rewrite as

$$\begin{aligned}\bar{C}_1(a) &= \alpha^2\tau_\alpha^{2\alpha} \int_{1}^{\infty} \int_{1}^{\infty} [1 - e^{-x}][1 - e^{-y}][1 - e^{-xya^{-2/\alpha}/\tau_\alpha^2}](xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y \qquad (3.63)\\ &= \tau_\alpha^{2\alpha}\,\mathbb{E}[1 - e^{-W_1W_2a^{-2/\alpha}/\tau_\alpha^2}] + \bar{e}_1(a),\end{aligned}$$

where we let $\bar{e}_1(a)$ denote the contribution due to e^{-x} and e^{-y}. We use [3, (4.6)], which states that

$$\mathbb{E}[1 - e^{-\delta W_1 W_2}] = (1 + o(1))\alpha\Gamma(1-\alpha)\delta^\alpha \log(1/\delta). \tag{3.64}$$

This gives

$$\bar{C}_1(a) = \tau_\alpha^{2\alpha}\alpha\tau_\alpha^{-2\alpha}\frac{2}{\alpha}\Gamma(1-\alpha)\frac{\log a}{a^2} + \bar{e}_1(a) = 2\Gamma(1-\alpha)\frac{\log a}{a^2} + \bar{e}_1(a). \tag{3.65}$$

Finally,

$$\begin{aligned}\bar{e}_1(a) &\le 2\alpha^2\tau_\alpha^{2\alpha}\int_1^\infty\int_1^\infty e^{-x}[1 - e^{-xya^{-2/\alpha}/\tau_\alpha^2}](xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y \\ &= 2\tau_\alpha^{2\alpha}\,\mathbb{E}\left[e^{-W_1}(1 - e^{-W_1W_2a^{-2/\alpha}/\tau_\alpha^2})\right] = O(a^{-2}).\end{aligned} \tag{3.66}$$

This can be seen by letting X be defined by

$$\mathbb{P}(X \in [a,b]) = \frac{1}{\mathbb{E}\left[e^{-W_1}\right]}\,\mathbb{E}\left[e^{-W_1}\mathbb{1}_{\{X\in[a,b]\}}\right], \tag{3.67}$$

so that

$$\bar{e}_1(a) = \Theta(1)\,\mathbb{E}\left[1 - e^{-XWa^{-2/\alpha}/\tau_\alpha^2}\right]. \tag{3.68}$$

Since X has thin tails (it has an exponential moment), the random variable XW is regularly varying with exponent α, so that

$$\mathbb{E}\left[1 - e^{-XWa^{-2/\alpha}/\tau_\alpha^2}\right] = \Theta(a^{-2}), \tag{3.69}$$

as claimed. The above follows from the fact that XW is again heavy-tailed with tail exponent α, as proved in [161, Proposition 5.2].

§3.5 Average Clustering Coefficient.

The previous results finally allow us to estimate the overall expected value $\bar{C}$ of the average local clustering coefficient C by averaging the annealed clustering function $\bar{C}(a)$ over the distribution $P(a)$ of the reduced degree a as

$$\bar{C} \equiv \int da\ \bar{C}(a)P(a) \simeq \int_{a\in L} \mathrm{d}a\ \bar{C}_L(a)P(a) + \int_{a\in H} \mathrm{d}a\ \bar{C}_H(a)P(a)$$

where we have formally split the integral into two contributions, produced by leaf ($a \in L$) and hub ($a \in H$) nodes respectively, consistently with the two regimes characterized asymptotically in (3.43) and (3.44). A rigorous definition of this split is provided in next section 3.5.1, where we also show that, in order to compute $\bar{C}$ exactly, it is not necessary to identify precisely the crossover value(s) of a separating the two regions. Notably, it turns out that the integral over $a \in H$ in (3.70) vanishes, implying that hubs no not contribute to the overall clustering. The remaining contribution

of the leaf nodes $a \in L$ is always finite, but its precise value depends on whether (following our initial discussion on the two possible definitions of C) nodes with degree $k < 2$ are considered as having $C_{k<2} = 0$ and hence included in the interval L (in which case $\bar{C}_L(a) \equiv 0$ for $0 \leq a \leq 1/\sqrt{n}$) or not (in which case $C_{k<2}$ is undefined and not included in the integration). In next section 3.5.1 we find that, depending on the choice,

$$\lim_{n\to\infty} \bar{C} = \begin{cases} 1 & \text{if } C_{k<2} \text{ undefined} \\ 1 - r_{0/1} & \text{if } C_{k<2} = 0 \end{cases}, \tag{3.70}$$

where $r_{0/1}$ is the expected fraction of nodes with degree 0 or 1. We plot the two versions of the average local clustering coefficient in Fig. 3.2, confirming the agreement with the respective theoretical result. This result is quite remarkable, as it means that the network is fully clustered, in the sense that for almost all nodes that can form triangles, their neighbours are connected and actually close the triangle.

§3.5.1 Computation of Average Clustering Coefficient

The limiting asymptotics for hubs and leaves allows us to derive an estimate of the average clustering coefficient of the network $\bar{C}$. Indeed, one needs to integrate the annealed clustering function $\bar{C}(a)$ over the distribution $P(a)$ of the reduced degree a, and from figure 3.1 it is clear that one can split the integral into the two leaf (L) and hub (H) regimes as

$$\bar{C} = \int_{a_{min}}^{\infty} \mathrm{d}a \bar{C}(a) P(a) = \int_{a_{min}}^{a^*} \mathrm{d}a \bar{C}_L(a) P(a) + \int_{a^*}^{\infty} \mathrm{d}a \bar{C}_H(a) P(a), \tag{3.71}$$

where $C_H(a) \sim 2\Gamma(1-\alpha)\frac{\log a}{a^2}$ (for $a > a^*$) is the asymptotic clustering function of hubs, while $C_L(a) \sim 1$ (for $a_{min} < a < a^*$) is the asymptotic plateau for leaves, a_{min} is the minimal reduced degree above which the clustering function starts to have meaning, and a^* is an appropriate crossover value between the two regimes. This crossover is visible in Figure 3.1, although its value is not easily identified. Luckily, as we will show in the following computations, the actual value of a^* is not necessary for the computation of the clustering to go through, and we can proceed without specifying it.

The main issue in the evaluation of the integral in Eq. (3.71) is that the degree distribution is not known everywhere but only in the tail [3], where it has the asymptotic behaviour $\Gamma(1-\alpha)k^{-2}$. Transforming to the distribution $P(a)$ of the reduced degree $a = \frac{k}{\sqrt{n}}$,

$$P(a) \sim \frac{\Gamma(1-\alpha)}{\sqrt{n}} a^{-2}. \tag{3.72}$$

Nonetheless, the first integral can be evaluated by realizing that by definition it evaluates to the fraction of leaves nodes or, equivalently

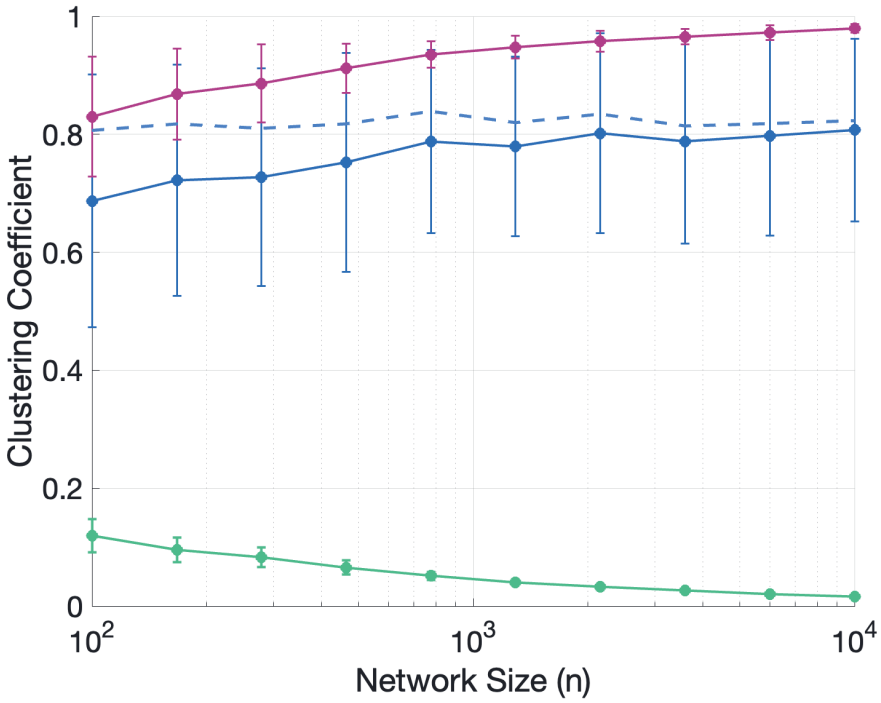


Figure 3.2: Average local clustering coefficient C versus network size n, for $\alpha = 0.5$. For each sampling of the n node weights, a single adjacency matrix is realized, and then weights are redrawn (up to 10^3 times for each n). In purple the clustering coefficient defined excluding the nodes with degree 0 and 1. In blue, the solid line is the average clustering coefficient evaluated by setting the clustering function to zero for $k = 0, 1$. We highlight that the error bars over this quantity do not shrink as n increases. The dashed blue line is 1 minus the average of $r_{0/1}$ over realizations. Finally, in green the difference between $1 - r_{0/1}$ and $\bar{C}$ computed factoring in nodes with $k < 2$, and we notice that in this case the error shrinks as desired.

$$\int_{a_{min}}^{a^*} \mathrm{d}a \bar{C}_L(a) P(a) = \int_{a_{min}}^{a^*} \mathrm{d}a P(a) = \begin{cases} 1 - \int_{a^*}^{\infty} da P(a) & \text{if } C_{k<2} \text{ undefined} \\ 1 - r_{0/1} - \int_{a^*}^{\infty} da P(a) & \text{if } C_{k<2} = 0 \end{cases} \tag{3.73}$$

and the last term can be recomprised in the second integral as it is evaluated on the same interval, so that the average clustering coefficient has the expression

$$\bar{C} = \begin{cases} 1 + \int_{a^*}^{\infty} P(a)(\bar{C}_H(a) - 1) & \text{if } C_{k<2} \text{ undefined} \\ 1 - r_{0/1} + \int_{a^*}^{\infty} da P(a)(\bar{C}_H(a) - 1) & \text{if } C_{k<2} = 0 \end{cases} \tag{3.74}$$

We can evaluate this last integral as

$$\begin{aligned}\int_{a^*}^{\infty} da P(a)(\bar{C}_H(a)-1) &= \int_{a^*}^{\infty} da\left[2\Gamma(1-\alpha)\frac{\log a}{a^2}-1\right]\frac{\Gamma(1-\alpha)}{a^2\sqrt{n}} \qquad (3.75)\\ &= \frac{\Gamma(1-\alpha)}{\sqrt{n}}\frac{6\Gamma(1-\alpha)\log a^* + 2\Gamma(1-\alpha) - 9(a^*)^2}{9(a^*)^3}\\ &= o(1),\end{aligned}$$

and see that it is suppressed by a square root of n term, whatever the value of a^*.

§3.6 Lack of self-averaging.

The result in Eq.(3.70) is calculated conditionally on the realized weights $\{w_i\}_{i=1}^n$. It shows that, if nodes with $k < 2$ are excluded, $\bar{C}$ converges to the (deterministic) maximum value 1, with vanishing fluctuations. By contrast, as we show now, for finite but large n the fractions of nodes with degree zero and one concentrate around functions of the rescaled total weight $S_n = \delta_n \sum_{j=1}^n w_j$, which in turn converges in distribution to an α-stable random variable and therefore remains a fluctuating quantity. If one looks instead at isolated and pendant nodes, the fractions concentrate around functions of the rescaled total weight $S_n = \delta_n \sum_{j=1}^n w_j$, which in turn converges in distribution to an α-stable random variable. This remains a fluctuating quantity even in the large n limit, as proven in the followings and confirmed numerically in Figure 3.2. This means that, for given n and α, different realizations of the weights produce different values of $r_{0/1}$ and of $\bar{C}$, if nodes with $k < 2$ are included. Therefore, under this second definition, $\bar{C}$ converges in distribution to the random variable $1 - r_{0/1}$, remaining macroscopically fluctuating even in the large graph limit. This behaviour follows from the infinite-mean nature of the weights and is a notable manifestation of the (rather unique) lack of self-averaging in macroscopic network properties (in particular the fractions of nodes with degree $k < 2$), which is not typically observed in models with independent edges and finite-mean weights. Notably, the lack of self-averaging extends to further 'low degree' nodes. This means that any macroscopic property that is a function of the full degree distribution, and not just its tail, will in general keep memory of the fluctuations of the fractions of low-degree nodes. Figure 3.3 confirms that other macroscopic properties (spectral gap, largest connected component size) besides the clustering coefficient keep fluctuating in the thermodynamic limit. We now prove these claims, and show these results are confirmed numerically in Figs. 3.2 3.3.

§3.6.1 Computation of $r_{0/1}$

We start by computing the fraction of nodes that are disconnected from the graph as

$$r_0 \equiv \frac{1}{n}\sum_{i=1}^{n}\prod_{j:\, j\neq i}(1-p_{ij}) = \frac{1}{n}\sum_{i=1}^{n} e^{-\delta_n w_i \sum_{j\neq i} w_j}, \qquad (3.76)$$

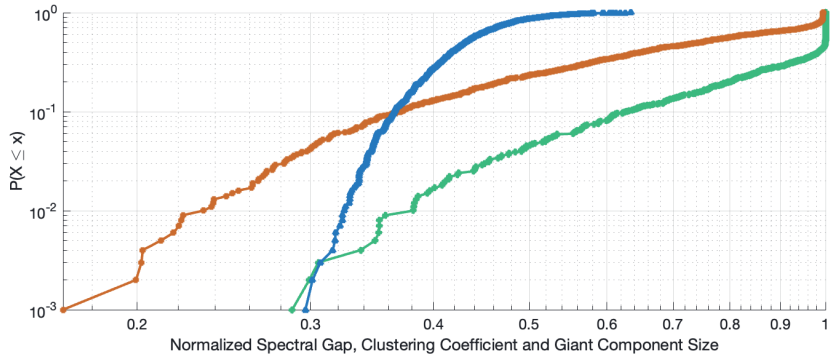


Figure 3.3: ***Cumulative distributions of network properties across*** 10^3 ***realizations with resampled weights*** ($n = 10^4$, $\alpha = 0.5$)***.*** *Orange: average local clustering coefficient including isolated and pendant nodes. Green: fraction of nodes in the largest connected component. Blue: normalized spectral gap (difference between the two largest eigenvalues of the adjacency matrix, divided by the largest one).*

where we recall that $\delta_n = n^{-1/\alpha}$. Since the $j = i$ term in the sum in the exponential makes little difference, we can approximate

$$r_0 \approx \frac{1}{n}\sum_{i=1}^{n} e^{-\delta_n w_i \sum_j w_j}. \tag{3.77}$$

We next use that most of the weights have normal size, and they contribute most to the sum over $i \in \{1, \dots, n\}$. Thus,

$$r_0 \approx \mathbb{E}\left[e^{-WS_n} \mid S_n\right], \tag{3.78}$$

where we take the expectation over the Pareto variable W in the above, we abbreviate

$$S_n = \delta_n \sum_{j=1}^{n} w_j, \tag{3.79}$$

and condition on this total rescaled weight. We note that the random variable S_n converges in distribution to an α-stable distribution, since the random variables $(w_i)_{i=1}^n$ are Pareto variables with infinite mean. Thus,

$$r_0 \approx \mathbb{E}\left[e^{-W\mathcal{S}} \mid \mathcal{S}\right], \tag{3.80}$$

where $\mathcal{S}$ is an α-stable random variable. Equation (3.80) requires some interpretation. What it means is that the proportion of vertices with r_0 degree 0, given the vertex weights $(w_i)_{i=1}^n$, is close to the rhs of (3.78), i.e., it holds that $r_0 \approx \mathbb{E}\left[e^{-WS_n} \mid S_n\right]$, where the rescaled total weight S_n is defined in (3.79). In particular, this means that, when simulating the network several times (and thus redrawing the vertex weights), the proportion of vertices with degree 0 *fluctuates macroscopically*, even in the large

graph limit. This is a type of *non-self-averaging* phenomenon that is rather unusual in network science. The simulations shown in Figures 3.6 3.7 3.8 confirm that (3.80) is an excellent approximation of r_0, and that it remains truly random in the large-graph limit.

We can also use (3.80) us to relate our approximation to [3, Proposition 4], in which $\mathbb{E}[r_0]$ is shown to be approximated by

$$\mathbb{E}[r_0] \approx \mathbb{E}\left[e^{-\Gamma(1-\alpha)W^\alpha}\right], \tag{3.81}$$

which indeed equals $\mathbb{E}\left[e^{-W\mathcal{S}}\right]$, since, for every $a \geq 0$,

$$\mathbb{E}\left[e^{-a\mathcal{S}}\right] = e^{-\Gamma(1-\alpha)a^\alpha}, \tag{3.82}$$

which follows from [3, Theorem 5]. For a practical computation of the approximation of r_0, we use the fact, again for $a \geq 0$,

$$\mathbb{E}\left[e^{-aW}\right] = \alpha a^\alpha \Gamma(-\alpha, a), \tag{3.83}$$

where we recall that $\Gamma(r, a)$ is the incomplete gamma function. Applying this to (3.78), we are led to

$$r_0 \approx \alpha S_n^\alpha \Gamma(-\alpha, S_n). \tag{3.84}$$

.

We show in Figures 3.6 3.7 3.8 that (3.84) is in very good accordance with the actual fraction of disconnected nodes (central panels).

We extend the above analysis to the proportion r_1 of vertices of degree 1. We start from

$$\begin{aligned} r_1 &\equiv \frac{1}{n}\sum_{i,k:\, i\neq k} p_{ik} \prod_{j:\, j\neq i,k} (1-p_{ij}) \\ &= \frac{1}{n}\sum_{i,k:\, i\neq k} (1-e^{-\delta_n w_i w_k}) e^{-\delta_n w_i \sum_{j:\, j\neq i,k} w_j} \\ &\approx \sum_k \mathbb{E}\left[(1-e^{-W\delta_n w_k}) e^{-W\delta_n \sum_{j:\, j\neq k} w_j}\right], \end{aligned} \tag{3.85}$$

where the asymptotics follows from the fact that the main contribution in the sum over i arises from normal weight vertices, so that the average over i can be replaced by an expectation with respect to W.

Note that the dominant contributions to the sum over k in (3.85) correspond to those k for which w_k is of the order $1/\delta_n = n^{1/\alpha}$, which means that these are the maximal-weight vertices. This will be the guiding principle to derive the asymptotics of the above formula.

For simplicity, and without loss of generality, we will assume that the weights are *ordered* from large to small, so that $w_1 \geq w_2 \cdots \geq w_n$. We can rewrite

$$\delta_n \sum_{j:\, j\neq k} w_j = S_n - \delta_n w_k = S_n - y_k^{(n)}, \tag{3.86}$$

where we recall S_n from (3.79), and we define the rescaled vertex weights $(y_k^{(n)})_{k\geq 1}$ as

$$y_k^{(n)} = \delta_n w_k. \tag{3.87}$$

In terms of these definitions, we obtain the finite-size approximation

$$r_1 \approx \sum_k \mathbb{E}\Big[(1-e^{-Wy_k^{(n)}})e^{-W(S_n-y_k^{(n)})} \mid (y_k^{(n)})_{k\geq 1}\Big]. \tag{3.88}$$

This approximation is inspired by the asymptotics properties of r_1. Indeed, as $n\to\infty$, we have the convergence in distribution

$$\Big(S_n, (y_k^{(n)})_{k\geq 1}\Big)\Big(\sum_{j\geq 1}\Gamma_j^{-1/\alpha}, (\Gamma_k^{-1/\alpha})_{k\geq 1}\Big), \tag{3.89}$$

where denotes convergence in distribution in the product topology, $(\Gamma_k)_{k\geq 1}$ are gamma random variables, i.e.,

$$\Gamma_k = \sum_{i=1}^k E_i, \tag{3.90}$$

where $(E_i)_{i\geq 1}$ are independent and identically distributed exponential random variables with mean 1. Because of this, we see that the random variables in (3.88) converge in distribution in the large graph limit, so that

$$r_1 \sum_k \mathbb{E}\Big[(1-e^{-Wy_k})e^{-W(\mathcal{S}-y_k)} \mid (y_k)_{k\geq 1}\Big], \tag{3.91}$$

where

$$\Big(\mathcal{S}, (y_k)_{k\geq 1}\Big) \equiv \Big(\sum_{j\geq 1}\Gamma_j^{-1/\alpha}, (\Gamma_k^{-1/\alpha})_{k\geq 1}\Big). \tag{3.92}$$

Thus, as for r_0, also r_1 converges in *distribution*, but it remains to fluctuate even in the large graph limit.

For a practical computation of the approximation of r_0, we again use (3.83) on the representation

$$r_1 \approx \sum_{k=1}^n \mathbb{E}\Big[e^{-W(S_n-y_k^{(n)})} - e^{-WS_n} \mid (y_k^{(n)})_{k\geq 1}\Big], \tag{3.93}$$

to arrive at

$$r_1 \approx \alpha \sum_{k=1}^n \Big[(S_n-y_k^{(n)})^\alpha \Gamma(-\alpha, S_n-y_k^{(n)}) - S_n^\alpha \Gamma(-\alpha, S_n)\Big], \tag{3.94}$$

which has the advantage that it no longer contains the expectation with respect to W. We show in Figures 3.6 3.7 3.8 that (3.94) is in very good accordance with the actual fraction of disconnected nodes (bottom panels).

§3.7 Lack of self-averaging in the fraction of next low-degree nodes

The fact that both r_0 and r_1 do not converge to a limiting deterministic value, but rather *in distribution* to a random variable whose variance is non-vanishing even when n tends to infinity, is a peculiar aspect of the MSM. This property applies not only to isolated or pendant nodes, but also to all the next 'low degree' nodes. Indeed, one can estimate r_k, the fraction of nodes whose degree is k, generalizing (3.76) and (3.85) as follows:

$$\begin{aligned} r_k &\equiv \frac{1}{n} \sum_{\substack{i_0,i_1,\dots,i_k \\ (i_r \neq i_s \,\forall r,s)}} p_{i_0 i_1} p_{i_0 i_2} \cdots p_{i_0 i_k} \prod_{j\,:\, j\neq i_0,i_1,\dots i_k} (1-p_{i_0 j}) \qquad (3.9 \\ &= \frac{1}{n} \sum_{\substack{i_0,i_1,\dots,i_k \\ (i_r \neq i_s \,\forall r,s)}} \left(1-e^{-\delta_n w_{i_0} w_{i_1}}\right)\left(1-e^{-\delta_n w_{i_0} w_{i_2}}\right)\cdots\left(1-e^{-\delta_n w_{i_0} w_{i_k}}\right) e^{-\delta_n w_{i_0} \sum_{j\,:\, j\neq i_0,i_1,\dots i_k} w} \end{aligned}$$

The key point now is that if one wants to proceed as in the last step of (3.85) and replace the sum over i_0 divided by n with the expectation over the whole sample, one should ensure that the sample without k weights is still representative of a Pareto distribution. This happens if k is small compared to n. In this case, we can proceed as in (3.85) with the approximation

$$\begin{aligned} r_k &\approx \sum_{\substack{i_1,i_2,\dots,i_k \\ (i_r \neq i_s \,\forall r,s)}} \mathbb{E}\left[(1-e^{-W\delta_n w_{i_1}})(1-e^{-W\delta_n w_{i_2}})\cdots(1-e^{-W\delta_n w_{i_k}})e^{-W\delta_n \sum_{j\,:\, j\neq i_0,i_1,\dots i_k} w_j}\right] \\ &= \sum_{\substack{i_1,i_2,\dots,i_k \\ (i_r \neq i_s \,\forall r,s)}} \mathbb{E}\left[(1-e^{-W y_{i_1}^{(n)}})(1-e^{-W y_{i_2}^{(n)}})\cdots(1-e^{-W y_{i_k}^{(n)}})e^{-W\left(S_n - \sum_{j=i_1,i_2,\dots i_k} y_j^{(n)}\right)}\right]. \end{aligned}$$

conditional on the realized values of the various $y_j^{(n)}$. Again, as $n\to\infty$, those random variables will converge *in distribution* to the Gamma variables (3.89), and also r_k will be converging *in distribution* to the quantity

$$r_k \sum_{\substack{i_1,i_2,\dots i_k \\ (i_r \neq i_s \,\forall r,s)}} \mathbb{E}\left[(1-e^{-W y_{i_1}})(1-e^{-W y_{i_2}})\dots(1-e^{-W y_{i_k}})e^{-W\left(\mathcal{S} - \sum_{j=i_1,i_2,\dots i_k} y_j\right)}\right], \tag{3.96}$$

thus still fluctuating macroscopically even in the large n limit, as we confirm in Figures 3.4 and 3.5. Indeed the figures show that the dispersion of r_k is significant, with the fluctuations shrinking as k (and not n) increases, suggesting that, for larger values of the degree, self-averaging is eventually restored. We leave the thorough study of this intriguing phenomenon to future works on the MSM.

§3.8 Extension to stable weights

In this section we briefly discuss how our results can be extended to the case in which the weights are sampled from a genuine α-stable distribution $\tilde{\rho}_\alpha(w)$ with positive

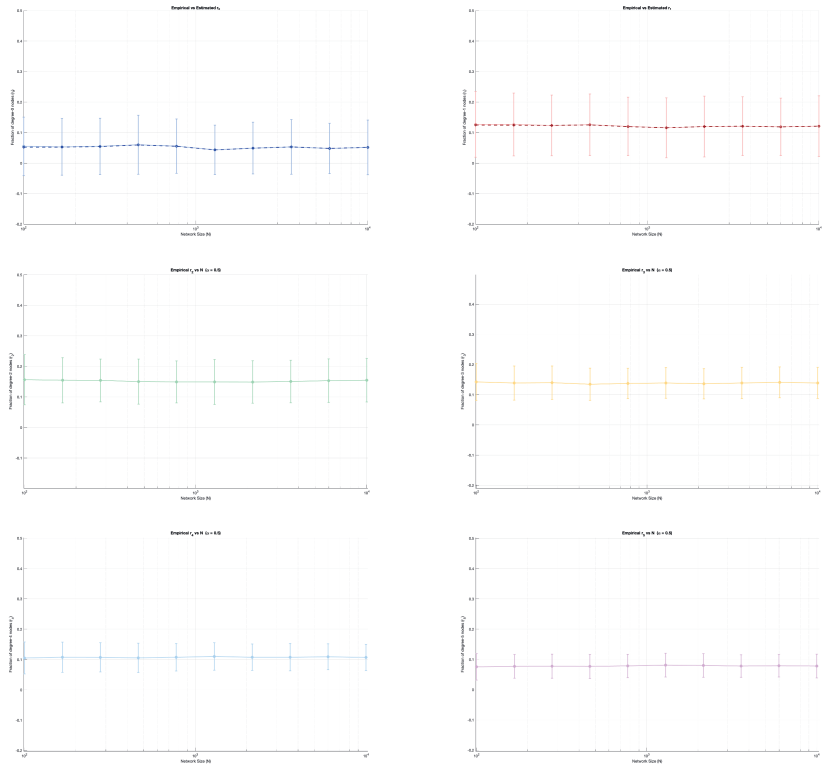


Figure 3.4: ***Fractions of low-degree (up to*** $k = 5$***) nodes versus network size, with*** $\alpha = 0.5$**.** *From left to right and top to bottom, we plot the fractions of nodes with a specific low degree (respectively* $k = 0, 1, 2, 3, 4, 5$*) across a sample of* 1000 *different realizations for each value of* n*. For the first two figures we include the theoretical estimate we explicitly computed in* (3.84) (3.88)*. We notice that the average value across different values of* n *is roughly constant, while the dispersion of values around that average does not shrink as* n *increases, as one would expect by assuming the self averaging of the quantities. We also notice that the error bars shrink as* k *increases, suggesting that, for larger-degree nodes, self averaging eventually occurs.*

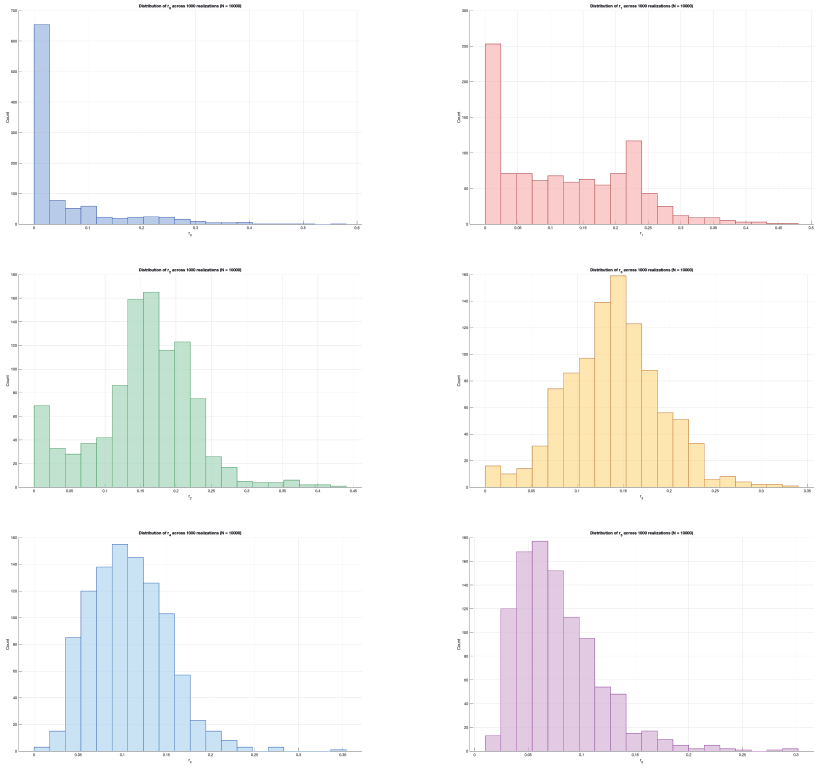


Figure 3.5: ***Distribution of the fraction of low-degree (up to*** $k = 5$***) nodes, for*** $\alpha = 0.5$***,*** $n = 1000$ ***(over a sample of*** 1000 ***realizations).*** *From left to right and top to bottom, we plot the distribution of the fractions of nodes with low degree (respectively* $k = 0, 1, 2, 3, 4, 5$*), across a sample of* 1000 *different realizations for* $n = 10^4$*. The sample variance* σ_k *is respectively* $\sigma_0 = 0.09$*,* $\sigma_1 = 0.1$*,* $\sigma_2 = 0.07$*,* $\sigma_3 = 0.05$*,* $\sigma_4 = 0.04$*,* $\sigma_5 = 0.03$*, indicating that the variance shrinks for larger values of the degree* k*.*

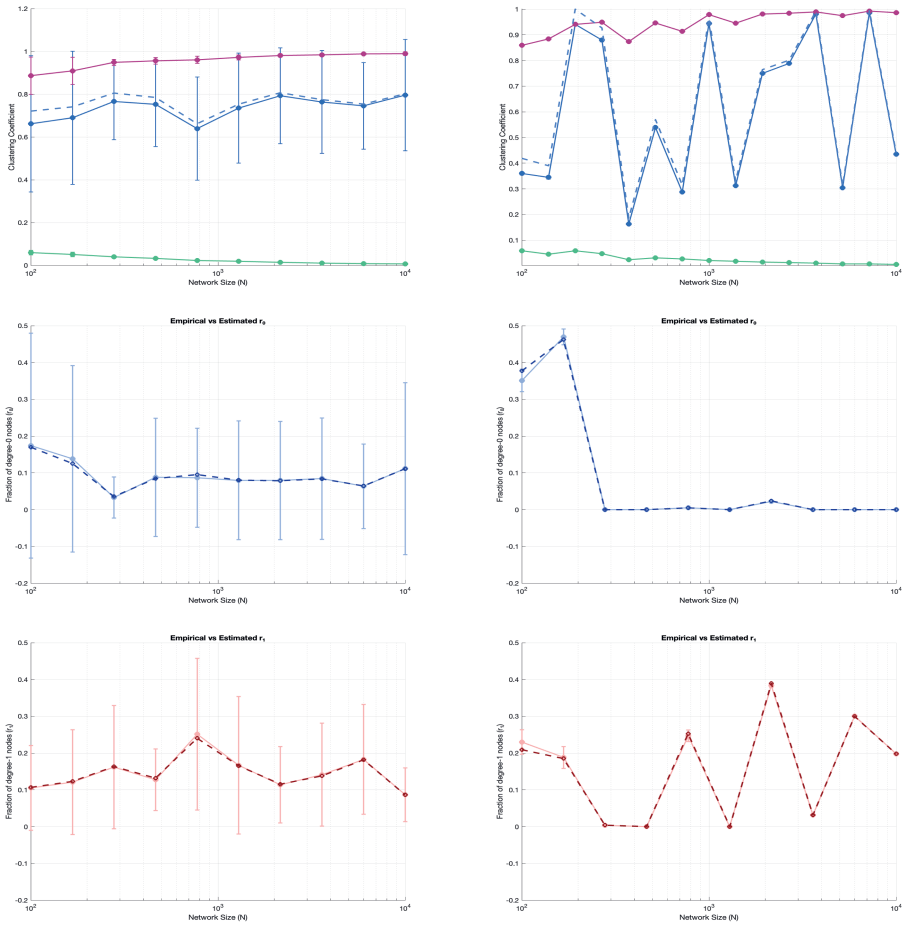


Figure 3.6: Numerics for $\alpha = 0.3$ (Pareto). On the left, simulations are done by resampling weights every time an actual network is realized (sample of 10), while on the left for each n weights are extracted only once, and 10 different adjacency matrices are realized from the same weights. On top, the same analysis performed in figure 3.2: in purple the clustering coefficient computed by excluding nodes with $k < 2$; in blue the clustering coefficient including those nodes and $1 - r_{0/1}$, respectively the solid and dashed lines; in green with error bars the difference between $1 - r_{0/1}$ and $\bar{C}$ computed factoring in nodes with $k < 2$, . In the middle, the light blue line with error bars is the actual value of r_0, while the dashed darker one is the approximation we derive in eq (3.84). At the bottom, the light red line with error bars is the actual value of r_1, while the dashed darker one is the approximation we derive in eq (3.88)

support and diverging mean. Throughout, we follow the formalism introduced in [22], and in particular we adopt the first parametrization defined there, corresponding to $k = 0$.

For stable laws, our arguments can be adapted by replacing the exact expression for the Pareto weight density, $\rho_\alpha(w) = \alpha w^{-(\alpha+1)}$ for $w \geq 1$ in (3.7), with an asymptotic

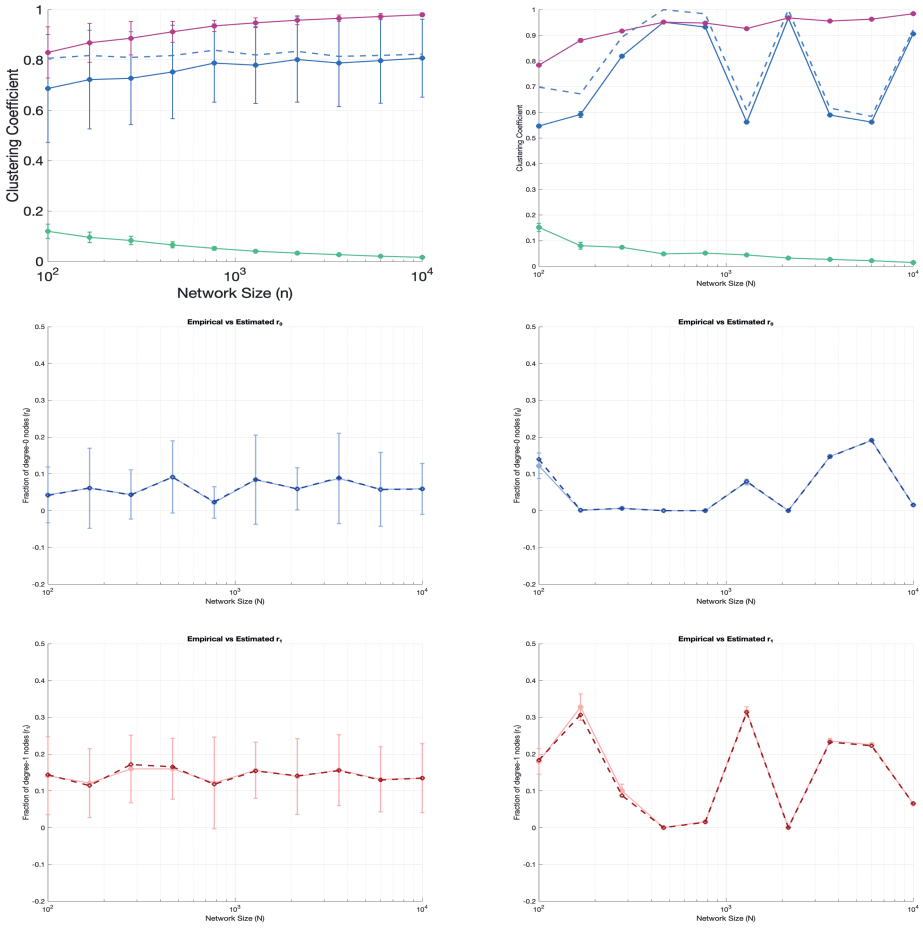


Figure 3.7: Numerics for $\alpha = 0.5$ (Pareto). On the left, simulations are done by resampling weights every time an actual network is realized (sample of 10), while on the left for each n weights are extracted only once, and 10 different adjacency matrices are realized from the same weights. On top, the same analysis performed in figure 3.2: in purple the clustering coefficient computed by excluding nodes with $k < 2$; in blue the clustering coefficient including those nodes and $1 - r_{0/1}$, respectively the solid and dashed lines; in green with error bars the difference between $1 - r_{0/1}$ and $\bar{C}$ computed factoring in nodes with $k < 2$, . In the middle, the light blue line with error bars is the actual value of r_0, while the dashed darker one is the approximation we derive in eq (3.84). At the bottom, the light red line with error bars is the actual value of r_1, while the dashed darker one is the approximation we derive in eq (3.88)

relation. Specifically, Theorem 1.2 of [22] shows that the density of an α-stable law satisfies

$$\tilde{\rho}_\alpha(w) = (1 + o(1))\, \frac{\alpha \gamma^\alpha (1+\beta) c_\alpha}{w^{\alpha+1}}, \qquad w \geq 1, \tag{3.97}$$

where the tail parameter satisfies $\alpha \in (0,1)$, $c_\alpha = \frac{\Gamma(\alpha)}{\pi} \sin\left(\frac{\pi\alpha}{2}\right)$, γ denotes the scale parameter, and β is the skewness parameter of the stable distribution.

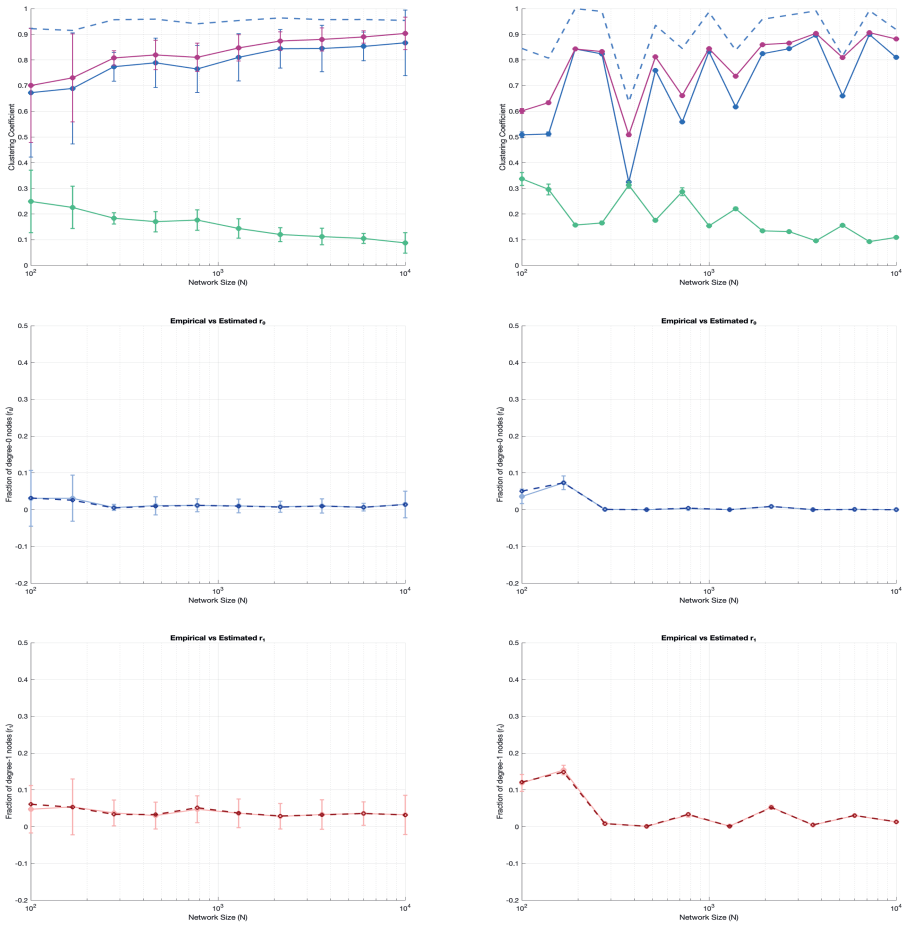


Figure 3.8: Numerics for $\alpha = 0.7$ (Pareto). On the left, simulations are done by resampling weights every time an actual network is realized (sample of 10), while on the left for each n weights are extracted only once, and 10 different adjacency matrices are realized from the same weights. On top, the same analysis performed in figure 3.2: in purple the clustering coefficient computed by excluding nodes with $k < 2$; in blue the clustering coefficient including those nodes and $1 - r_{0/1}$, respectively the solid and dashed lines; in green with error bars the difference between $1 - r_{0/1}$ and $\bar{C}$ computed factoring in nodes with $k < 2$, . In the middle, the light blue line with error bars is the actual value of r_0, while the dashed darker one is the approximation we derive in eq (3.84). At the bottom, the light red line with error bars is the actual value of r_1, while the dashed darker one is the approximation we derive in eq (3.88)

In order to work with a distribution supported on the positive real axis, we choose $\beta = 1$ and set the infimum of the support to zero. This choice fixes the location parameter to be $\delta = \gamma \tan\left(\frac{\pi\alpha}{2}\right)$. The only remaining free parameter is therefore the scale γ, which we determine by requiring that the tails of the stable and Pareto

distributions coincide asymptotically. Imposing this matching condition yields

$$\gamma = \left[\frac{\pi}{2\,\Gamma(\alpha)\sin\left(\frac{\pi\alpha}{2}\right)} \right]^{1/\alpha}. \tag{3.98}$$

With this choice of parameters, all the results we derived for Pareto-distributed weights extend directly to the case of stable sampling. In particular, since the integrals appearing in Section 3.3.2 are dominated by contributions from large values of w, the asymptotic estimates derived in Section 3.5.1 remain valid. Similarly, the arguments used in Section 3.6.1 to compute the fraction of disconnected nodes and nodes of degree one carry over almost verbatim to the stable case. The only modification is the appearance of a constant c in (3.89), which is now directly related to the scale parameter of the underlying stable law.

Indeed, in the convergence statement (3.89), the partial sum S_n has a strictly α-stable distribution for every n. Interpreting S_n as the value at time 1 of a stable Lévy process, the sum $c\sum_{j\geq 1}\Gamma_j^{-1/\alpha}$ can be viewed as the collection of jump sizes of the process, ordered by decreasing magnitude. Correspondingly, the rescaled weights $n^{-1/\alpha}W_i$ represent the increments of the process over the time intervals $[(i-1)/n, i/n]$, and the variables $(y_k^{(n)})_{k=1}^n$ in (3.87) can be interpreted as the ordered increments over these intervals. Since, with high probability and for large n, each interval contains at most one large jump, it follows that (3.89) continues to hold in this setting, with the first coordinate (corresponding to the sum) having exactly the same distribution as in the Pareto case.

Finally, in Figures 3.10, 3.11, 3.12, and 3.13, we provide numerical evidence confirming that the results presented in the main text remain valid when the weights are drawn from an α-stable distribution, in agreement with the arguments given above.

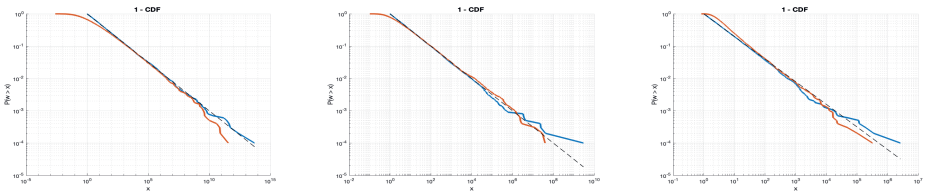


Figure 3.9: ***Comparison between pure Pareto and Stable distribution in the tail behaviour****. Using the parameter of* (3.98)*, we can match the tails of the two distributions. From left to right* $\alpha = 0.3, 0.5, 0.7$*, sample of* 10^4*.*

§3.9 Concluding Remarks.

This work contributes to the open debate about whether, in models with conditionally independent edges, the coexistence of sparsity and local clustering (along with other desirable features such as a broad degree distribution and the small-world property) can be attained *without geometry*, i.e. without assuming that the connection probability p_{ij} depends on the node attributes w_i and w_j necessarily through a function

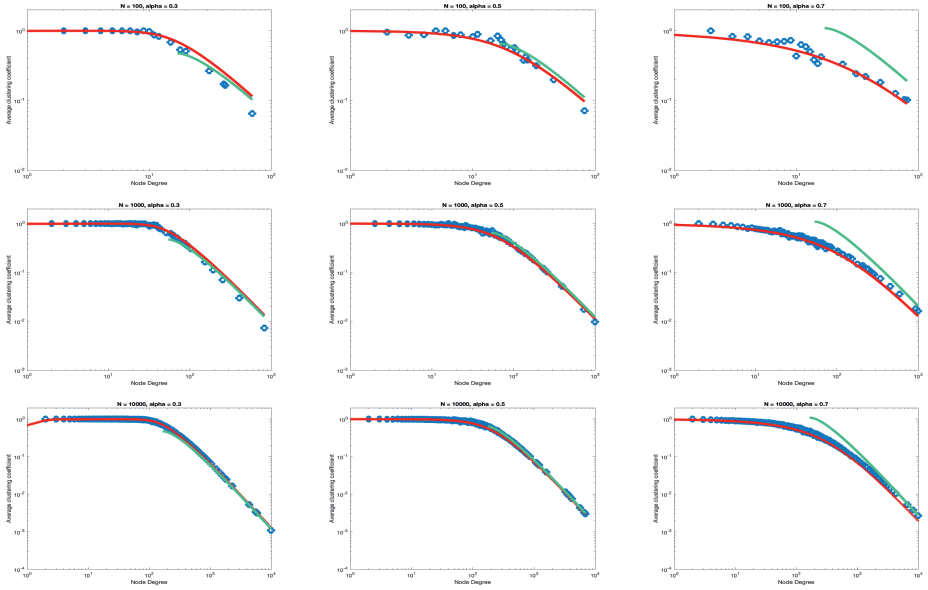


Figure 3.10: ***C****lustering functions for different values of* n *and* α *for the MSM model* (3.1) *with Stable weights* (3.2)*. Blue circles: empirical clustering function* (3.4) *(versus the reduced degree* $a = k/\sqrt{n}$*) computed on actual realized graphs sampled from the model (obtained by sampling the weights once, and sampling the graph once conditionally on the realized weights). Red curves: our analytical expression* (3.6) *for the annealed clustering function. Green curves: our asymptotic calculation* (3.44) *valid for diverging reduced degrees. We evaluate and plot all functions only for degrees larger than* 1*, to avoid ambiguities in the definition of the clustering coefficient for* $k < 2$*. From top to bottom:* $n = 10^2, 10^3, 10^4$*. From left to right:* $\alpha = 0.3, 0.5, 0.7$*.*

of some metric distance $d(w_i, w_j)$ [14, 15, 16, 17, 19, 18, 21, 87, 91]. We considered a geometry-free model with independent edges and infinite-mean fitness [1, 2], recently introduced in the context of network renormalization [54], and calculated its full clustering function – rigorously proving the resulting *finite* local clustering in the large n limit and highlighting that clustering does *not* imply geometry. The infinite-mean nature of the fitness arises from a requirement of model invariance under node aggregation [1] and generates other realistic properties including sparsity, a power-law degree distribution [2], and the breakdown of self-averaging of various network properties, which is a novel result characterized here. The model then provides, as an alternative to geometry, the hypothesis of node aggregation invariance as a basic mechanism producing realistic network properties.

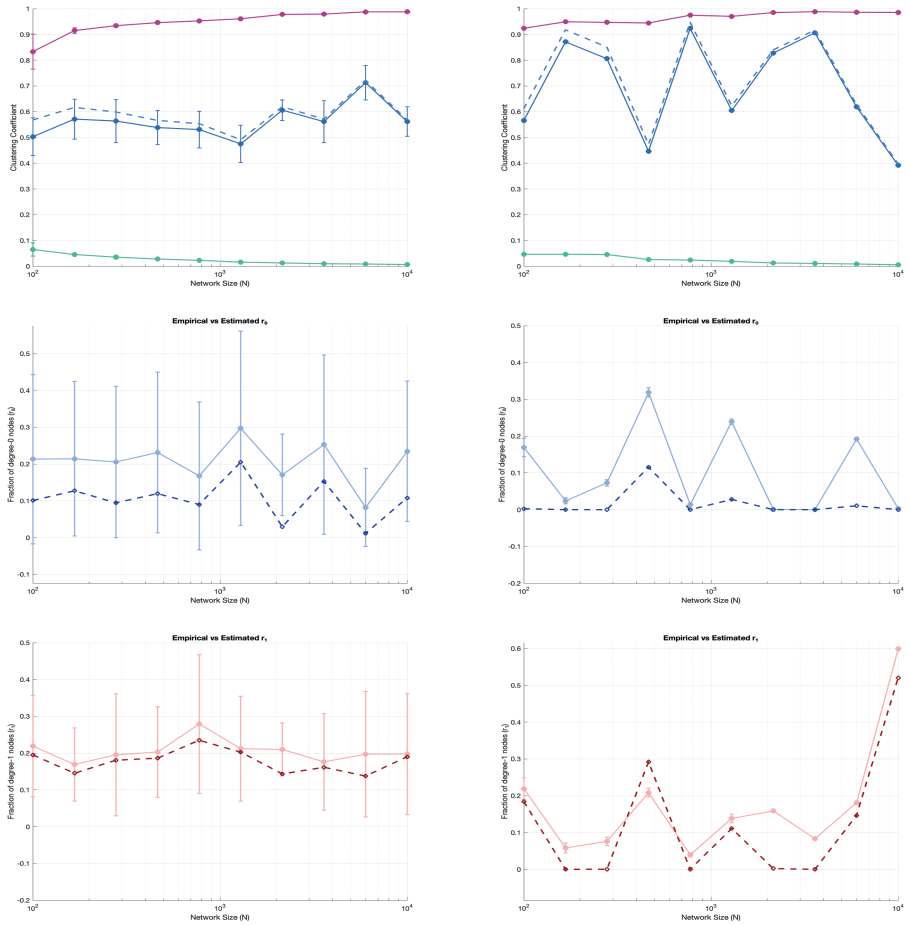


Figure 3.11: Numerics for $\alpha = 0.3$ (stable). On the left, simulations are done by resampling weights every time an actual network is realized (sample of 10), while on the left for each n weights are extracted only once, and 10 different adjacency matrices are realized from the same weights. On top, the same analysis performed in figure 3.2: in purple the clustering coefficient computed by excluding nodes with $k < 2$; in blue the clustering coefficient including those nodes and $1 - r_{0/1}$, respectively the solid and dashed lines; in green with error bars the difference between $1 - r_{0/1}$ and $\bar{C}$ computed factoring in nodes with $k < 2$, . In the middle, the light blue line with error bars is the actual value of r_0, while the dashed darker one is the approximation we derive in eq (3.84). At the bottom, the light red line with error bars is the actual value of r_1, while the dashed darker one is the approximation we derive in eq (3.88)

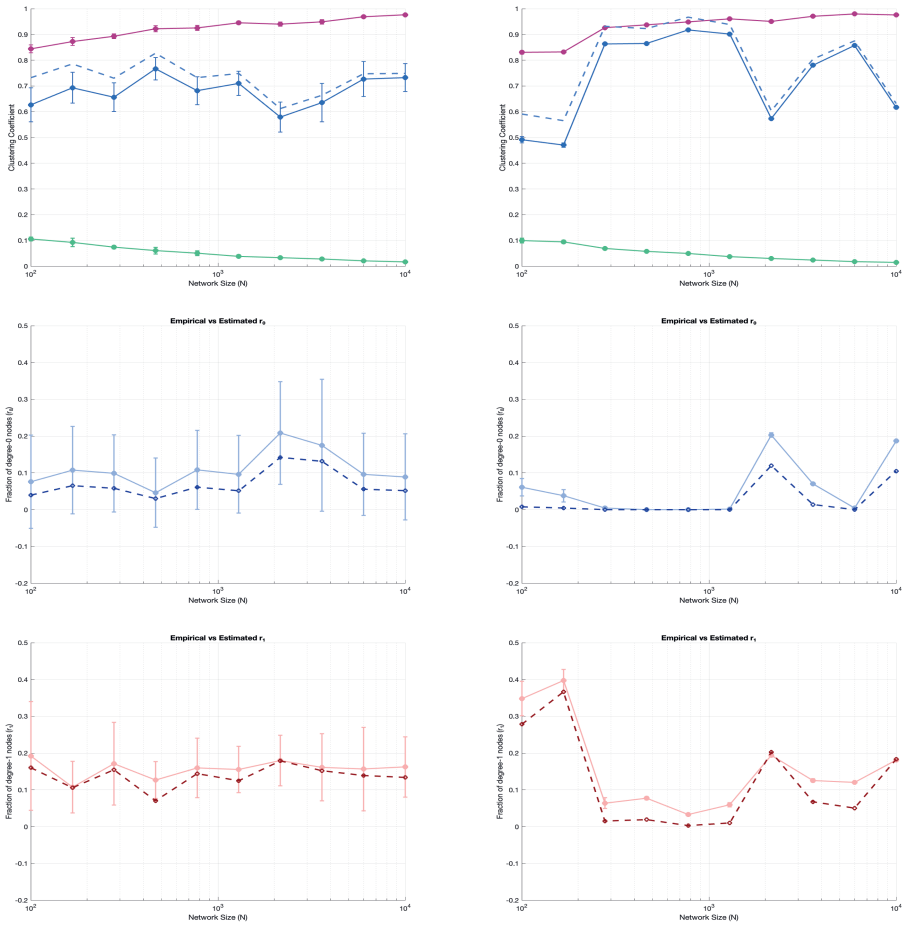


Figure 3.12: Numerics for $\alpha = 0.5$ (stable). On the left, simulations are done by resampling weights every time an actual network is realized (sample of 10), while on the left for each n weights are extracted only once, and 10 different adjacency matrices are realized from the same weights. On top, the same analysis performed in figure 3.2: in purple the clustering coefficient computed by excluding nodes with $k < 2$; in blue the clustering coefficient including those nodes and $1 - r_{0/1}$, respectively the solid and dashed lines; in green with error bars the difference between $1 - r_{0/1}$ and $\bar{C}$ computed factoring in nodes with $k < 2$, . In the middle, the light blue line with error bars is the actual value of r_0, while the dashed darker one is the approximation we derive in eq (3.84). At the bottom, the light red line with error bars is the actual value of r_1, while the dashed darker one is the approximation we derive in eq (3.88)

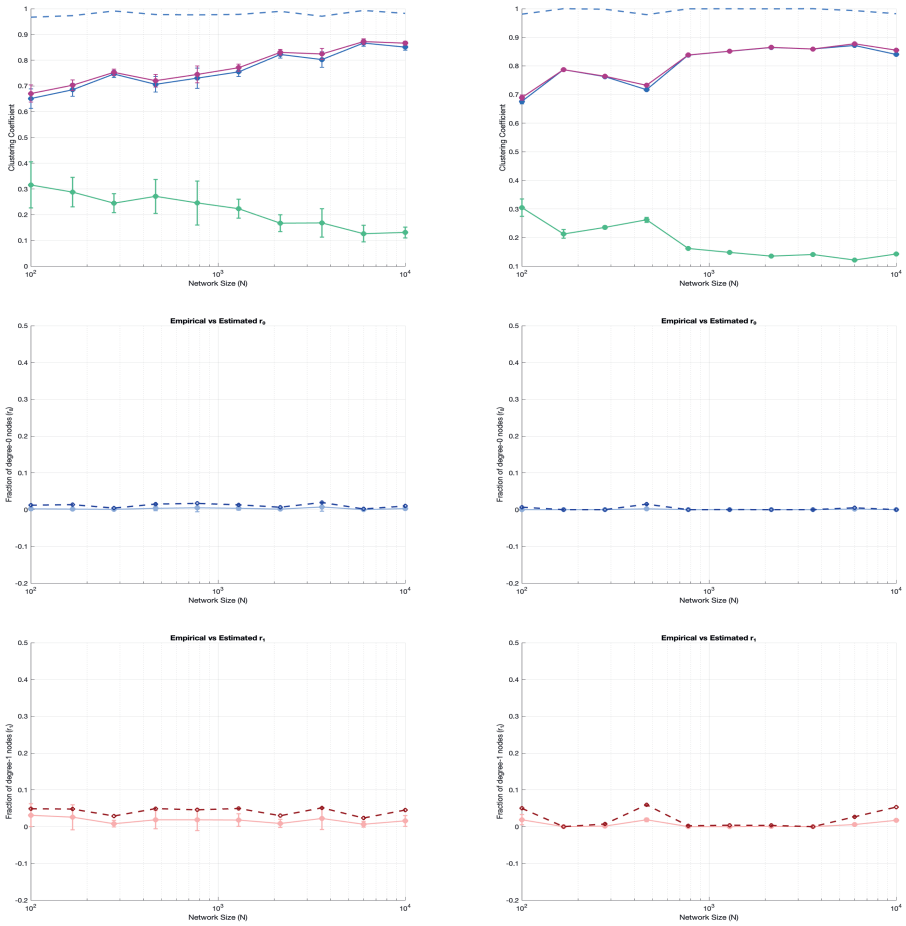


Figure 3.13: Numerics for $\alpha = 0.7$ (stable). On the left, simulations are done by resampling weights every time an actual network is realized (sample of 10), while on the left for each n weights are extracted only once, and 10 different adjacency matrices are realized from the same weights. On top, the same analysis performed in figure 3.2: in purple the clustering coefficient computed by excluding nodes with $k < 2$; in blue the clustering coefficient including those nodes and $1 - r_{0/1}$, respectively the solid and dashed lines; in green with error bars the difference between $1 - r_{0/1}$ and $\bar{C}$ computed factoring in nodes with $k < 2$, . In the middle, the light blue line with error bars is the actual value of r_0, while the dashed darker one is the approximation we derive in eq (3.84). At the bottom, the light red line with error bars is the actual value of r_1, while the dashed darker one is the approximation we derive in eq (3.88)

PART II

RENORMALIZATION OF INTERACTING RANDOM NETWORK MODELS

CHAPTER 4

Topological Renormalization for Random Network Models

Abstract

In this chapter we develop a renormalization framework for probabilistic network models based on marginalization of unobserved degrees of freedom. Starting from a general Exponential Random Graph formulation, we derive exact relations describing how effective Hamiltonians transform when links are hidden or unobserved. Unlike traditional renormalization group approaches, which rely on geometric coarse-graining and scale invariance, the present framework is driven by uncertainty: degrees of freedom are removed due to incomplete observations rather than spatial scale separation.

We show that single-link marginalization acts as a generator of a commutative semigroup of transformations on network ensembles, providing a natural notion of "uncertainty flow" in model space. The resulting renormalization equations reveal that marginalization generically induces higher-order effective interactions, leading to a lack of closure of finite-order random network models. We analyze conditions under which closure can still occur, highlighting special cases such as bounded coordination structures, including the coordination-number-two scenario where interaction propagation is constrained.

Finally, we discuss how the non-closure of the renormalization flow can be leveraged for inference. By fitting effective models in the observable (reduced) space and exploiting the derived renormalization relations, one can formulate inverse problems aimed at reconstructing microscopic Hamiltonian parameters from partial observations. This establishes a bridge between renormalization concepts and network reconstruction, suggesting new tools for inference in incomplete or uncertain network data.

§4.1 Introduction

Incomplete or partially observed networks arise naturally in many domains, ranging from social and biological systems to technological and financial networks. In most real-world settings, the full interaction structure is not directly accessible due to measurement limitations, sampling biases, or intrinsic uncertainty. This raises a fundamental question: how does partial observability affect the statistical description of network ensembles?

A natural probabilistic framework for describing network ensembles is provided by Exponential Random Graph Models (ERGMs), which encode structural constraints through an effective Hamiltonian defined over the space of adjacency matrices [4, 48, 47]. These models are widely used because of their flexibility and their principled connection to maximum-entropy inference. However, ERGMs are typically formulated assuming that the full configuration space is observable.

In this chapter we address a complementary problem: rather than performing inference in the full space, we investigate how probabilistic network models transform under *marginalization of degrees of freedom*. In particular, we ask how the effective Hamiltonian changes when some links are unobserved. This perspective is closely related in spirit to Renormalization Group (RG) ideas, where one studies how effective theories evolve when microscopic degrees of freedom are integrated out [102, 103].

The key difference from traditional RG approaches is conceptual: instead of coarse-graining based on geometric scale separation, we consider *uncertainty-driven marginalization*. That is, degrees of freedom are removed because they are unobserved, not because they belong to a particular spatial scale. This shift in viewpoint leads to a renormalization framework that is intrinsically combinatorial and tailored to network ensembles.

Our goals are threefold. First, we derive exact relations between the parameters of ERG-like Hamiltonians before and after marginalization. Second, we interpret single-link marginalization as the generator of an "uncertainty flow" acting on network models. Finally, we explore the implications of these flows for model closure and inference, showing that marginalization generically induces higher-order interactions in the effective theory, and how this can be used to infer the parameters of unobserved degrees of freedom.

§4.2 Marginalization framework

We start with very general assumptions, *i.e.* that there is a probability function $P(\cdot)$ that evaluates the probability that a certain network of size N, described by its adjacency matrix $\mathbf{A}$, is realized. We work with entries A_{ij} to be assuming values either 0 or 1 for the sake of simplicity in the computations and clarity in the exposition, but we believe this framework could be extended to weighted networks, directed graphs and more general structures.

The space of possible configurations (which we call underlying, complete or starting space) this system can have is

$$\{\mathbf{A}\} = \{0,1\}^V \tag{4.1}$$

sometimes also referred as the Boolean Hypercube. We now ask ourselves what happens if we are not able to see the whole of the configuration space, but just a subset of it.

This uncertainty compels us to reduce the available configuration space to the one defined by the adjacency matrix elements we are actually able to observe. Let us begin by introducing the tiniest bit of localized uncertainty in the system: a single link $(\bar{ij})$ cannot be observed, leading to the new configuration space

$$\{\tilde{\mathbf{A}}\} = \{\mathbf{A}\}\backslash\{A_{\bar{ij}}\} \tag{4.2}$$

which we call reduced, renormalized, or marginalized space. Being unsure whether $A_{\bar{ij}}$ is either 0 or 1, we have to marginalize the probability function of the complete space over all the possible values of the unobservable link $A_{\bar{ij}}$, so to get a probability function just for the variables we are actually observing. This reads as

$$\tilde{P}(\tilde{\mathbf{A}}) = \sum_{A_{\bar{ij}}=0,1} P(\mathbf{A}) \tag{4.3}$$

and constitutes the key formula for the renormalization procedure we will set up.

§4.2.1 Marginalization Flow for Exponential Random Graphs

The marginalization equation (4.3) we just wrote is greatly enhanced by the fact that we can express the probability function over the complete space in the exponential form [47, 48]

$$P(\mathbf{A}) = \frac{e^{-H(\mathbf{A})}}{Z} \qquad Z = \sum_{\{\mathbf{A}\}} e^{-H(\mathbf{A})} \tag{4.4}$$

where the Hamiltonian is a multilinear function of the adjacency matrix elements

$$H(A) = -h^{(0)} - \sum_{ij} h^{(1)}_{ij} A_{ij} - \sum_{ij\neq kl} h^{(2)}_{ijkl} A_{ij} A_{kl} - \sum_{\substack{ij\neq kl,mn \\ kl\neq mn}} h^{(3)}_{ijklmn} A_{ij} A_{kl} A_{mn} - \dots \tag{4.5}$$

We note that the full Exponential Random Graph formalism would require to add the constraint equations so that all the parameters of the Hamiltonian (5.2) assume their most probable value according to the observated data. However, in our framework we are precisely excluding to have available data for the underlying, complete space of all link variables. We let these parameters to be free, as we want to see how they transform under marginalization. The inference procedure is then deferred over to the reduced space, in which observables are well-defined.

In order to derive the equations connecting the parameters in the underlying space to those in the observable space, we also write the reduced space probability function in the same exponential form

$$\tilde{P}(\tilde{\mathbf{A}}) \equiv \frac{e^{-\tilde{H}(\tilde{\mathbf{A}})}}{\tilde{Z}} \qquad \tilde{Z} = \sum_{\{\tilde{\mathbf{A}}\}} e^{-\tilde{H}(\tilde{\mathbf{A}})} \tag{4.6}$$

with different coupling parameters that will be encoding the very fact that interactions among links that were involving the inaccessible one might be changing, hence leading to potentially different couplings. The Hamiltonian then has the same multilinear structure over the same (except the unseen $A_{\bar{ij}}$) variables as

$$\tilde{H}(\tilde{\mathbf{A}}) = -\tilde{h}^{(0)} - \sum_{ij \neq \bar{ij}} \tilde{h}^{(1)}_{ij} A_{ij} - \sum_{\substack{ij \neq \bar{ij}, kl \\ kl \neq \bar{ij}}} \tilde{h}^{(2)}_{ijkl} A_{ij} A_{kl} - \dots \tag{4.7}$$

The Marginalization equation (4.3), when combined with this specific form of the probability of equations (4.4) and(4.6), implicitly defines relations between the parameters of (4.7) and (4.5), in the form

$$\frac{e^{-\tilde{H}(\tilde{\mathbf{A}})}}{\tilde{Z}} = \sum_{A_{\bar{ij}}=0,1} \frac{e^{-H(\mathbf{A})}}{Z}. \tag{4.8}$$

We call the relations between the parameters of $\tilde{H}$ and H the Renormalization Group (RG) equations of the system.

The power of equation (4.8) is that it has to be true in all the possible configurations of the reduced space. This will allow for the actual recovery of parameters both at the renormalized level and (when possible) at the complete one. In particular, the fact that it has to be true in the void configuration allows us to further simplify it. As a matter of facts, when all $\{\tilde{\mathbf{A}}\} = \{0\}^N$ we have from (4.8)

$$\tilde{Z} = \frac{Z}{1 + e^{h^{(1)}_{\bar{ij}}}} \tag{4.9}$$

the can be inserted again in (4.8) so to get

$$e^{-\tilde{H}(\tilde{\mathbf{A}})} = \sum_{A_{\bar{ij}}=0,1} \frac{e^{-H(\mathbf{A})}}{1 + e^{h^{(1)}_{\bar{ij}}}} \tag{4.10}$$

which constitutes the key relation that we will be using to derive the relations between the parameters of the Hamiltonian. Before doing so, it is perhaps useful to understand in which sense the relations between the parameters of $\tilde{H}$ and H actually constitute a Renormalization Group flow. We do so in next section.

§4.2.2 What is different from 'usual' Renormalization

Although the marginalization procedure we introduced shares similarities with Renormalization Group (RG) ideas, there are important conceptual differences with standard formulations in statistical physics, which we now briefly discuss.

In conventional RG, one is typically interested in identifying fixed points of the flow, corresponding to scale-invariant theories and critical phenomena [103]. This is often enforced by requiring invariance of the partition function under coarse-graining,

$$\tilde{Z} = Z, \tag{4.11}$$

possibly up to a rescaling of thermodynamic potentials. In our framework, we do not impose such a constraint. The goal is not to identify scale-invariant models, but rather to characterize how effective network Hamiltonians change under loss of information. In general, marginalization does not preserve the partition function, and therefore fixed points are not expected.

A second key difference concerns the nature of the coarse-graining operation. Real-space RG relies heavily on an underlying geometry: spins are grouped into blocks defined by spatial proximity, and scale transformations are defined with respect to metric structure [102]. Networks, however, are fundamentally combinatorial objects. While some networks admit geometric embeddings [], many do not possess a natural notion of spatial scale.

In this work we therefore abandon the assumption of geometric coarse-graining in its entirety. Instead, degrees of freedom are removed because they are *unobserved*. The elementary transformation is the marginalization of a single link, as in Eq. (4.3). More complex uncertainty patterns can be constructed by composing such elementary steps.

This shift in perspective leads to a different interpretation of RG transformations. Rather than describing scale transformations in physical space, our flow describes how effective theories evolve as information is progressively lost. The relevant control parameter is not length scale, but observational completeness.

As we show below, single-link marginalization acts as a generator of a semigroup of transformations on probability distributions over graphs. This motivates interpreting the resulting flow as an "uncertainty RG", whose structure is determined purely by combinatorial marginalization rather than geometric scaling.

§4.2.3 The Generator of Missing Information

Studying the marginalization of a single link might seem limited at first sight, since realistic uncertainty patterns typically involve multiple missing links or nodes. However, the single-link transformation plays a fundamental role: it acts as a generator of all possible uncertainty transformations.

Let $\mathcal{M}_{\bar{ij}}$ denote the marginalization operator that removes the link $(\bar{ij})$, acting on probability distributions as

$$\left(\mathcal{M}_{\bar{ij}} P\right)(\tilde{\mathbf{A}}) = \sum_{A_{\bar{ij}}=0,1} P(\mathbf{A}). \tag{4.12}$$

If we marginalize two links $(\bar{ij})$ and $(\bar{kl})$, we obtain

$$\mathcal{M}_{\bar{kl}}\mathcal{M}_{\bar{ij}}P = \sum_{A_{\bar{kl}}=0,1}\sum_{A_{\bar{ij}}=0,1} P(\mathbf{A}). \tag{4.13}$$

Because finite sums are associative and commutative, these operators satisfy

$$\mathcal{M}_{\bar{ij}}\mathcal{M}_{\bar{kl}} = \mathcal{M}_{\bar{kl}}\mathcal{M}_{\bar{ij}}, \tag{4.14}$$

and therefore generate a commutative semigroup under composition.

Importantly, this structure is not a group: the transformations are not invertible. Multiple distributions in the complete space can map to the same marginalized distribution, reflecting the intrinsic loss of information under marginalization. In this sense, the flow is irreversible, similarly to coarse-graining transformations in standard RG.

This observation provides a structural interpretation of our framework. Any complex uncertainty pattern, which in general is a mixture of missing links, missing nodes, or combinations thereof (see Fig 4.1), can be constructed by composing elementary marginalization operators. Consequently, the single-link transformation acts as a generator of the full uncertainty semigroup acting on network ensembles.

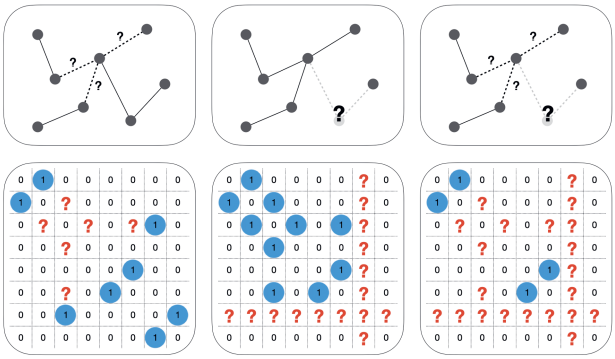


Figure 4.1: The possible ways the structure of a network is unknown, both in the structure and in the Adjacency Matrix. From left to right: individual links are missing; individual nodes are either unobserved or it is totally unknown how they are connected to the rest of the network; and a mixture of both when nodes and individual links are missing.

§4.3 RG flows for Random Network Models

In this section we show how find RG equations in general for random network models. The Hamiltonian is the one of equation (4.5), and we introduce exponential variables encoding for the correlations between links as

$$\alpha_{ij} = e^{h^{(1)}_{ij}} \quad \beta_{ijkl} = e^{h^{(2)}_{ijkl}} \quad \gamma_{ijklmn} = e^{h^{(3)}_{ijklmn}} \quad \cdots \tag{4.15}$$

and their respective partners in the reduced space

$$\tilde{\alpha}_{ij} = e^{\tilde{h}^{(1)}_{ij}} \quad \tilde{\beta}_{ijkl} = e^{\tilde{h}^{(2)}_{ijkl}} \quad \tilde{\gamma}_{ijklmn} = e^{\tilde{h}^{(3)}_{ijklmn}} \quad \cdots \tag{4.16}$$

With this identification we write equation (4.10) in the product form

$$\begin{aligned}
&\prod_{ij\neq\bar{i}\bar{j}} \tilde{\alpha}_{ij}^{A_{ij}} \prod_{\substack{ij\neq kl,\bar{i}\bar{j}\\ kl\neq\bar{i}\bar{j}}} \tilde{\beta}_{ijkl}^{A_{ij}A_{kl}} \prod_{\substack{ij\neq kl,mn,\bar{i}\bar{j}\\ kl\neq mn,\bar{i}\bar{j}\\ mn\neq\bar{i}\bar{j}}} \tilde{\gamma}_{ijklmn}^{A_{ij}A_{kl}A_{mn}} \cdots \\
&= \frac{1+\alpha_{\bar{i}\bar{j}} \prod_{ij\neq\bar{i}\bar{j}} \beta_{\bar{i}\bar{j}ij}^{A_{ij}} \prod_{\substack{ij\neq kl,\bar{i}\bar{j}\\ kl\neq\bar{i}\bar{j}}} \gamma_{\bar{i}\bar{j}ijkl}^{A_{ij}A_{kl}} \cdots}{1+\alpha_{\bar{i}\bar{j}}} \prod_{ij\neq\bar{i}\bar{j}} \alpha_{ij}^{A_{ij}} \prod_{\substack{ij\neq kl,\bar{i}\bar{j}\\ kl\neq\bar{i}\bar{j}}} \beta_{ijkl}^{A_{ij}A_{kl}} \prod_{\substack{ij\neq kl,mn,\bar{i}\bar{j}\\ kl\neq mn,\bar{i}\bar{j}\\ mn\neq\bar{i}\bar{j}}} \gamma_{ijklmn}^{A_{ij}A_{kl}A_{mn}} \cdots
\end{aligned} \tag{4.17}$$

In which the first term corresponds to a general Hamiltonian in the reduced space (explicitly written up to the third order, but higher order terms will keep the very same structure), and the second term corresponds to the marginalization of a general Hamiltonian (still up to third order) in the complete, underlying space. The key to finding the relations between parameters is again that Equation (4.17) must be true in any possible configuration. This means that we can switch on and off links in a meaningful way to derive how the correlation terms among them behave under the marginalization. For example, let us imagine that only one link (i, j), different from the unseable one, has the value one, while all other are zero. In this case equation (4.17) reduces to

$$\tilde{\alpha}_{ij} = \alpha_{ij} \frac{1+\alpha_{\bar{i}\bar{j}}\beta_{\bar{i}\bar{j}ij}}{1+\alpha_{\bar{i}\bar{j}}}, \tag{4.18}$$

which tells how the local fields are changing over the marginalization flow, and they change taking into account the binary interaction that the link (i, j) might have had with the unobservable link.

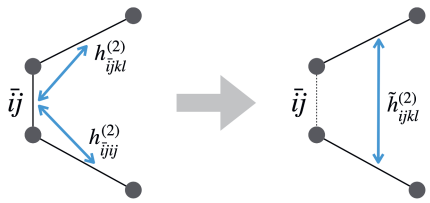


Figure 4.2: How being unable to observe a link can induce effective interactions between links that were not coupled before, for the binary interaction term.

If we activate two links: (i, j) and (k, l) Equation (4.17) reduces to

$$\tilde{\alpha}_{ij}\tilde{\alpha}_{kl}\tilde{\beta}_{ijkl} = \frac{1+\alpha_{\bar{ij}}\beta_{\bar{ij}ij}\beta_{\bar{ij}kl}\gamma_{\bar{ij}ijkl}}{1+\alpha_{\bar{ij}}}\alpha_{ij}\alpha_{kl}\beta_{ijkl} \tag{4.19}$$

and by inserting in it the relations (4.18) for the local fields for both (i,j) and (k,l) we get that the correlation term between two links changes as

$$\tilde{\beta}_{ijkl} = \beta_{ijkl}\frac{\left(1+\alpha_{\bar{ij}}\beta_{\bar{ij}ij}\beta_{\bar{ij}kl}\gamma_{\bar{ij}ijkl}\right)\left(1+\alpha_{\bar{ij}}\right)}{\left(1+\alpha_{\bar{ij}}\beta_{\bar{ij}ij}\right)\left(1+\alpha_{\bar{ij}}\beta_{\bar{ij}kl}\right)}. \tag{4.20}$$

in which we see again that it is mindful of the fact that there might have been a third order correlation among (i,j), (k,l) and the missing one, and hence the binary coupling between (i,j) and (k,l) is enhanced by such a factor.

The third order interaction parameter can be estimated as well by choosing three links (i,j), (k,l) and (m,n) to be on while all the others are off, and then simplifying the first and second order coupling with equations and (4.20), so to get

$$\tilde{\gamma}_{ijklmn} = \gamma_{ijklmn}\frac{\mathcal{N}_{ijklmn}}{\mathcal{D}_{ijklmn}} \tag{4.21}$$

where

$$\begin{aligned}\mathcal{N}_{ijklmn} = &\left(1+\alpha_{\bar{ij}}\beta_{\bar{ij}ij}\right)\left(1+\alpha_{\bar{ij}}\beta_{\bar{ij}kl}\right)\left(1+\alpha_{\bar{ij}}\beta_{\bar{ij}mn}\right)\\ &\times\left(1+\alpha_{\bar{ij}}\beta_{\bar{ij}ij}\beta_{\bar{ij}kl}\beta_{\bar{ij}mn}\gamma_{\bar{ij}ijkl}\gamma_{\bar{ij}klmn}\gamma_{\bar{ij}mnij}\delta_{\bar{ij}ijklmn}\right)\end{aligned} \tag{4.22}$$

and

$$\begin{aligned}\mathcal{D}_{ijklmn} = &\left(1+\alpha_{\bar{ij}}\right)\left(1+\alpha_{\bar{ij}}\beta_{\bar{ij}ij}\beta_{\bar{ij}kl}\gamma_{\bar{ij}ijkl}\right)\\ &\times\left(1+\alpha_{\bar{ij}}\beta_{\bar{ij}kl}\beta_{\bar{ij}mn}\gamma_{\bar{ij}klmn}\right)\left(1+\alpha_{\bar{ij}}\beta_{\bar{ij}mn}\beta_{\bar{ij}ij}\gamma_{\bar{ij}mnij}\right)\end{aligned} \tag{4.23}$$

Again, we notice that this term keeps track of the fact that the three links (i,j), (k,l) and (m,n) might have been in a fourth order interaction with the missing link, and this has to be factored in their reduced space third order term. Higher order terms can be derived in the very same fashion, but we stop at third order as they will be frankly too cumbersome.

§4.3.1 Approximated and Exact closure under RG

The RG relations derived above allow us to address a fundamental question: given a k^{th}-order Random Network Model (RNM) in the complete space, what is the order of the effective model in the reduced space?

In general, with the exception of I order models, marginalization does not preserve model order. Even if the original Hamiltonian contains interactions up to order k, integrating out degrees of freedom generates higher-order effective interactions. The

simplest example is provided by second-order models: Eq. (4.21) shows that marginalization generically induces third-order couplings, even when no such terms are present in the original Hamiltonian. Explicitly, for the second order Hamiltonian

$$H(A) = -\sum_{ij} h^{(1)}_{ij} A_{ij} - \sum_{ij \neq kl} h^{(2)}_{ijkl} A_{ij} A_{kl} - \tag{4.24}$$

the marginalization equation predicts a new term in the reduced space that will be encoding for a third order correlation among links, as

$$\tilde{h}^{(3)}_{ijklmn} = \log \frac{\mathcal{P}_{ijklmn}}{\mathcal{Q}_{ijklmn}} \tag{4.25}$$

where the numerator is

$$\mathcal{P}_{ijklmn} = \left(1 + e^{h^{(1)}_{\bar{ij}} + h^{(2)}_{\bar{ij}ij}}\right)\left(1 + e^{h^{(1)}_{\bar{ij}} + h^{(2)}_{\bar{ij}kl}}\right)\left(1 + e^{h^{(1)}_{\bar{ij}} + h^{(2)}_{\bar{ij}mn}}\right) \times \left(1 + e^{h^{(1)}_{\bar{ij}} + h^{(2)}_{\bar{ij}ij} + h^{(2)}_{\bar{ij}kl} + h^{(2)}_{\bar{ij}mn}}\right) \tag{4.26}$$

and the denominator

$$\mathcal{Q}_{ijklmn} = \left(1 + e^{h^{(1)}_{\bar{ij}}}\right)\left(1 + e^{h^{(1)}_{\bar{ij}} + h^{(2)}_{\bar{ij}ij} + h^{(2)}_{\bar{ij}kl}}\right)\left(1 + e^{h^{(1)}_{\bar{ij}} + h^{(2)}_{\bar{ij}kl} + h^{(2)}_{\bar{ij}mn}}\right) \times \left(1 + e^{h^{(1)}_{\bar{ij}} + h^{(2)}_{\bar{ij}ij} + h^{(2)}_{\bar{ij}mn}}\right). \tag{4.27}$$

This lack of closure is analogous to what happens in field theory, where integrating out microscopic degrees of freedom generates an infinite tower of effective operators. In the present context, the only models that remain closed under marginalization are highly specific cases, such as first-order models or specific low-coordination structures where interactions cannot propagate beyond bounded neighborhoods.

There is one notable exception to the lack of closure of RNM with order higher than one, and is the case in which each link interacts binarily at most with other two links. We study extensively this model and its possible twists in Chapter 5. In Section 4.4.1, we also show how we can map this model into a simple yet powerful method to perform inference on temporal network that might have been undersampled.

Rather than viewing non-closure as a limitation, we can reinterpret it as an opportunity for inference. Suppose that only the reduced space is observable. One can perform standard ERG inference in this space, estimating an effective Hamiltonian $\tilde{H}$ and validating it against suitable null models. This determines how confidently the observed network can be described by a given reduced-space model.

Crucially, the RG relations derived earlier provide explicit functional mappings

$$\tilde{h}^{(n)} = \mathcal{R}^{(n)}\left(h^{(1)}, h^{(2)}, \dots\right), \tag{4.28}$$

linking effective couplings to those of the complete space. If enough effective parameters can be inferred in the reduced space, these relations define a system of equations whose unknowns are the microscopic Hamiltonian parameters.

In principle, if this system is invertible, one can reconstruct the parameters of the complete-space model from reduced observations. The feasibility of this inversion depends on both the richness of the effective model and the degree of information loss induced by marginalization. This perspective turns RG flows into a bridge between incomplete observations and microscopic reconstruction, suggesting a new route to network inference under uncertainty, which we start to explore in Section 4.4.2.

§4.4 Towards applications to Network Inference

In this Chapter we are interpreting marginalization as a renormalization procedure acting on probabilistic network models. The resulting RG equations describe how effective Hamiltonians transform when parts of the network are hidden or unobserved. A natural question is whether these transformations can be exploited in the opposite direction: instead of propagating parameters forward under uncertainty, can we use the RG flow backward to infer properties of the unseen degrees of freedom?

This perspective suggests a new class of inference problems. Rather than reconstructing missing links directly from data, one attempts to infer microscopic parameters by comparing the observed (renormalized) model with the predictions of the marginalization equations. In other words, the RG map becomes a constraint linking observable effective models to families of compatible underlying ones.

In the following we explore two concrete scenarios in which this idea can be made explicit. We begin with a temporal setting, where marginalization corresponds to coarse temporal sampling. In this case the RG equations admit a closed form and can be locally inverted, providing a simple proof of concept for inference via RG inversion.

§4.4.1 Temporal Network reconstruction at higher resolution

Many complex systems of scientific and societal relevance, including economic, financial, ecological, and social systems, are inherently dynamic [104]. Their internal organization is networked, and both the state of the system and its interaction structure evolve over time. Understanding how network topology changes is essential for assessing resilience, systemic risk, and propagation of shocks.

A major obstacle in practice is the mismatch between the intrinsic time scales of the dynamics and the temporal resolution of available data. For example, international trade networks are often reported at yearly or quarterly frequencies, while shocks propagate over much shorter horizons. Similar issues arise in supply chains, interbank networks, and social interactions, where aggregation in time can mask rapid structural transitions [42, 105, 106]. As a result, one frequently observes a temporally coarse-grained version of an underlying network process evolving at a finer, unobserved resolution.

From a modelling perspective this creates a fundamental inference problem: one observes a renormalized trajectory in time, while the microscopic dynamics unfold at a higher sampling frequency. Standard approaches such as interpolation or state-

space models typically rely on ad hoc assumptions and do not explicitly account for how network structure transforms under temporal aggregation.

Here we show that the marginalization framework developed earlier naturally provides a principled description of temporal coarse-graining, and can therefore be used to formulate an RG-based inference strategy. We consider temporal networks in which each link switches between present and absent states according to a binary Markov process. That is, the probability that a link exists at time t depends only on its state at the previous time step. This class of models captures memory effects commonly observed in real data, while remaining analytically tractable. An explicit mapping between temporal networks and Markov link dynamics is discussed in [59, 60].

A key simplification arises from the fact that links evolve independently one from another. The probability of a full network trajectory therefore factorizes over links, allowing us to study single-link processes without loss of generality. From a statistical physics perspective, each link trajectory is equivalent to a one-dimensional spin chain with nearest-neighbour interactions in time, a system for which exact renormalization transformations are known.

The Hamiltonian of the complete temporal network over T time steps can be written as

$$H(\mathbf{A}) = -\sum_{t=1}^{T}\sum_{i<j}\epsilon_{ij}^{(t)}A_{ij}^{(t)} - \sum_{t=1}^{T}\sum_{i<j}g_{ij}^{(t,t-1)}A_{ij}^{(t)}A_{ij}^{(t-1)} = \sum_{i<j}H_{ij}(A_{ij}), \tag{4.29}$$

which factorizes into independent contributions from each link trajectory. Each term

$$H_{ij}(A_{ij}) = -\sum_{t=1}^{T}\epsilon_{ij}^{(t)}A_{ij}^{(t)} - \sum_{t=1}^{T}g_{ij}^{(t,t-1)}A_{ij}^{(t)}A_{ij}^{(t-1)} \tag{4.30}$$

describes a binary time series with local fields $\epsilon^{(t)}$ and nearest-neighbour temporal couplings $g^{(t,t-1)}$. This structure is formally identical to the Hamiltonians studied earlier in the context of link marginalization, with time playing the role of an additional degree of freedom. This observation allows us to directly apply the marginalization machinery in the temporal domain.

We can now reinterpret marginalization as temporal coarse-graining. Suppose that the link state at a given time t^* cannot be observed. Integrating over this hidden variable induces an effective interaction between the neighbouring time steps t^*-1 and t^*+1, exactly as topological marginalization induces effective higher-order couplings.

Carrying out the marginalization explicitly yields the renormalized parameters

$$\tilde{g}^{(t^*+1,t^*-1)} = \log\frac{\left(1+e^{\epsilon^{(t^*)}+g^{(t^*,t^*-1)}+g^{(t^*+1,t^*)}}\right)\left(1+e^{\epsilon^{(t^*)}}\right)}{\left(1+e^{\epsilon^{(t^*)}+g^{(t^*,t^*-1)}}\right)\left(1+e^{\epsilon^{(t^*)}+g^{(t^*+1,t^*)}}\right)} \tag{4.31}$$

and the local fields change as

$$\tilde{\epsilon}^{(t^*+1)} = \epsilon^{(t^*+1)} + \log \frac{\left(1 + e^{\epsilon^{(t^*)} + g^{(t^*+1,t^*)}}\right)}{\left(1 + e^{\epsilon^{(t^*)}}\right)} \tag{4.32}$$

and

$$\tilde{\epsilon}^{(t^*-1)} = \epsilon^{(t^*-1)} + \log \frac{\left(1 + e^{\epsilon^{(t^*)} + g^{(t^*,t^*-1)}}\right)}{\left(1 + e^{\epsilon^{(t^*)}}\right)} \tag{4.33}$$

Further insight can be gained by considering a time-homogeneous Markov chain, where parameters are stationary over time ($\epsilon^t = \epsilon$ and $g^{t,t-1} = g$ for all t). In this case the renormalization equations simplify considerably:

$$\tilde{g}^{(t^*+1,t^*-1)} = \log \frac{\left(1 + e^{\epsilon+2g}\right)\left(1 + e^{\epsilon}\right)}{\left(1 + e^{\epsilon+g}\right)^2} \tag{4.34}$$

and

$$\tilde{\epsilon}^{(t^*\pm1)} = \epsilon + \log \frac{\left(1 + e^{\epsilon+g}\right)}{\left(1 + e^{\epsilon}\right)}. \tag{4.35}$$

Even in this simple setting, an important effect emerges: marginalizing a single time step breaks temporal homogeneity. The renormalized parameters depend on the location of the missing observation, creating an inhomogeneous effective process. While this appears problematic, it also suggests a structured way to recover homogeneity.

If instead one marginalizes every other time step (for instance all even times) the resulting dynamics becomes homogeneous again. In this case the RG transformation closes within the space of stationary Markov chains and takes the simple form

$$\tilde{g} = \log \frac{\left(1 + e^{\epsilon+2g}\right)\left(1 + e^{\epsilon}\right)}{\left(1 + e^{\epsilon+g}\right)^2} \tag{4.36}$$

and

$$\tilde{\epsilon} = \epsilon + 2\log \frac{\left(1 + e^{\epsilon+g}\right)}{\left(1 + e^{\epsilon}\right)}, \tag{4.37}$$

These equations admit a natural interpretation: they describe how the effective parameters of a link process change when the sampling frequency is reduced by a factor of two. The original parameters (ϵ, g) correspond to the microscopic dynamics, while $(\tilde{\epsilon}, \tilde{g})$ describe the effective model observed under temporal coarse-graining.

Figure 4.3 illustrates this idea schematically. When all time steps are observed, the link dynamics is described by nearest-neighbour couplings. After frequency halving, the effective chain lives on a sparser temporal lattice, but retains the same structural form with renormalized parameters.

Inferring Higher Resolution Parameters

The renormalization map defined by Eqs. (4.36)–(4.37) induces a nonlinear flow in the parameter space. As shown in Fig. 4.4, the flow exhibits a characteristic shearing structure with a well-defined attractive line separating different dynamical regimes.

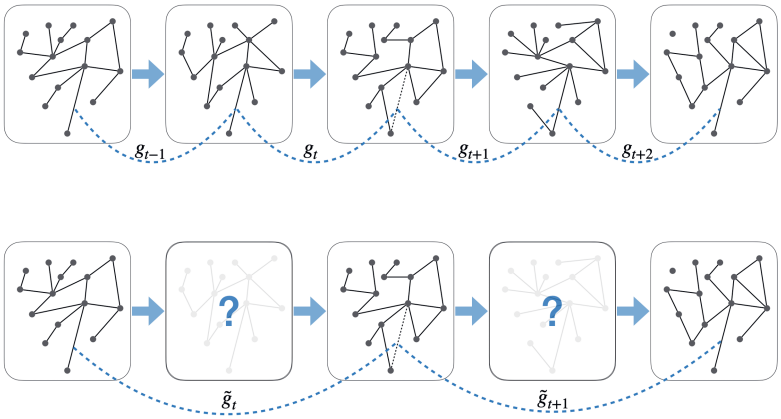


Figure 4.3: Linear coupling between the same link of a network over different time steps. While in the first row the sampling frequency is such that every time instance of the network can be seen, in the second row the sampling frequency is just half of the rate of change of the network, but the coupling structure between each time step and its precursor is preserved, and the value change according to eq. (4.36).

A key observation is that repeated coarse-graining drives correlations toward zero. This is expected: as temporal resolution decreases, short-time memory effects are progressively washed out, and the dynamics approaches an effectively memoryless process. In RG language, the line $g = 0$ plays the role of an attractive manifold corresponding to loss of temporal correlations. Another notable feature is the behaviour for negative couplings. If successive states are anticorrelated ($g < 0$), marginalization over alternating time steps induces an effective positive correlation at larger time scales. Intuitively, two consecutive anticorrelations combine into a positive correlation over longer temporal separations. This explains why the RG flow rapidly maps the $g < 0$ region into the $g > 0$ half-plane.

Most importantly, the structure of the flow in the region $g > 0$ opens the possibility of inverting the RG transformation and using it for inference. In this regime the renormalization map is locally one-to-one, meaning that the coarse-grained parameters $(\tilde{\epsilon}, \tilde{g})$ uniquely determine a compatible pair of microscopic parameters (ϵ, g) in a neighbourhood of parameter space.

From an inference perspective, this suggests the following strategy. Suppose that one observes a temporal network sampled at a coarse resolution and fits a stationary Markov model to the data, obtaining effective parameters $(\tilde{\epsilon}, \tilde{g})$. If the system is known (or assumed) to arise from a finer-scale Markov process undergoing frequency halving, then the microscopic parameters can be estimated by solving the coupled

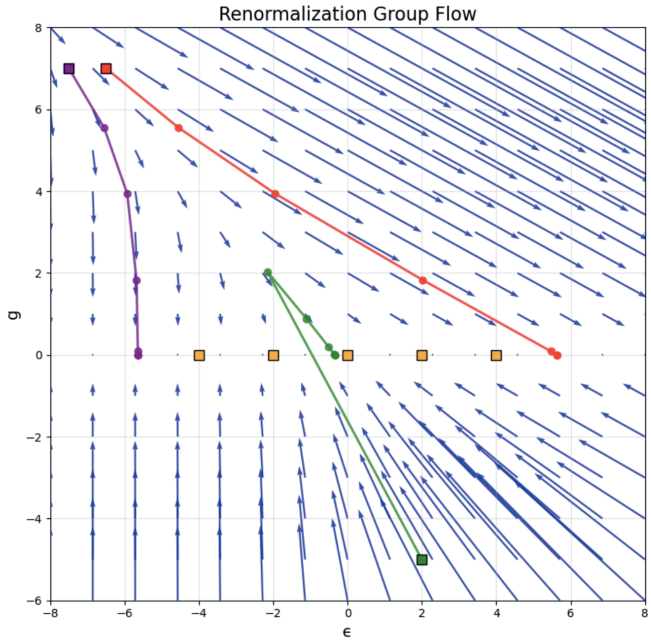


Figure 4.4: Renormalization flow of the Linear chain system when frequency halving is performed. There is a clear fixed, attractive line that separates the plane into two regions. Colored lines represent possible flows under successive transformation of halving frequencies, squares representing starting points.

nonlinear equations

$$\tilde{g} = \log \frac{(1 + e^{\epsilon+2g})(1 + e^{\epsilon})}{(1 + e^{\epsilon+g})^2}, \tag{4.38}$$

$$\tilde{\epsilon} = \epsilon + 2 \log \frac{1 + e^{\epsilon+g}}{1 + e^{\epsilon}}. \tag{4.39}$$

These equations define a two-dimensional nonlinear inverse problem: two equations for the two unknowns (ϵ, g). While no closed-form analytic inversion exists in general, the system can be solved numerically using standard root-finding methods. Importantly, the RG flow guarantees local invertibility in the upper-plane region of the flow, ensuring that the reconstruction is well-posed there.

This perspective reframes temporal super-resolution as an RG inversion problem. Instead of interpolating trajectories directly, one infers the microscopic dynamical parameters that are consistent with the observed coarse-grained statistics. Once the underlying parameters are reconstructed, they can be used to simulate higher-resolution trajectories or to assess how network dynamics would respond under finer temporal sampling.

§4.4.2 All-to-all and multiplex networks

Our marginalization-inversion framework finds a second natural domain of application in the context of international trade networks. Previous works [**?** 63] have shown that a sensible probabilistic representation of the International Trade Network is provided by a multiplex model. In this picture, each commodity defines a layer, while countries are replicated across layers.

The key modeling assumption is the following: trade between two countries is independent across different country pairs, but the probability of trading a given commodity between the same pair of countries depends on which other commodities they already exchange. This induces inter-layer correlations while preserving independence across country pairs.

For a fixed pair of countries (i, j), and the vector encoding for the different realizations of the link across layers $\vec{A}_{ij} = (A^1_{ij}, A^2_{ij}, \cdots)$ the Hamiltonian can therefore be written, assuming homogeneous inter-layer coupling, as

$$H(\vec{A}_{ij}) = -\sum_a \epsilon^a_{ij} A^{(a)}_{ij} - \beta \sum_{a \neq b} A^{(a)}_{ij} A^{(b)}_{ij}, \tag{4.40}$$

where the superscript a labels the commodity (layer). A schematic representation is shown in Fig. 4.5.

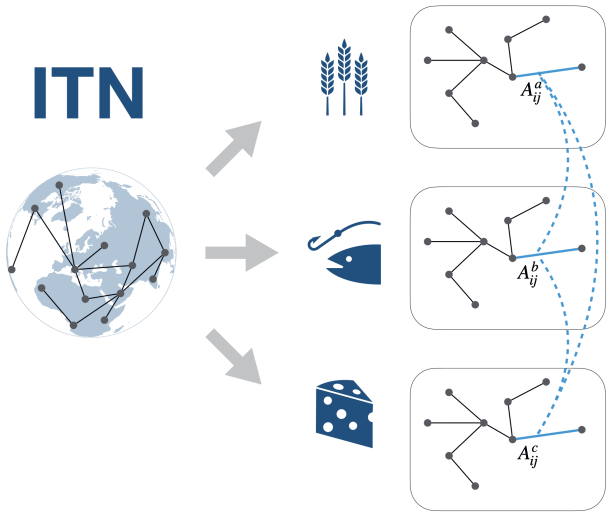


Figure 4.5: Multiplex representation of the International Trade Network. Each commodity defines a layer, and links are independent across country pairs but coupled across layers for the same pair.

Since marginalization acts locally on degrees of freedom, the relevant renormalization flow is not defined over countries but over the set of inter-layer replicas of a single country pair. We can therefore drop the (i, j) indices and focus on the inter-layer Hamiltonian.

We now ask: what happens if one layer is unobserved? For instance, suppose the trade of a specific commodity (e.g. fish) is missing, while all other commodities are observed. Using the general marginalization formulas (4.18) and (4.20), we obtain the renormalization flow for the observable parameters:

$$\tilde{\epsilon}^{(a)} = \epsilon^{(a)} + \log \frac{1 + e^{\epsilon^{(\bar{a})}+\beta}}{1 + e^{\epsilon^{(\bar{a})}}}, \tag{4.41}$$

$$\tilde{\beta} = \beta + \log \frac{\left(1 + e^{\epsilon^{(\bar{a})}+2\beta}\right)\left(1 + e^{\epsilon^{(\bar{a})}}\right)}{\left(1 + e^{\epsilon^{(\bar{a})}+\beta}\right)^2}. \tag{4.42}$$

An example of the resulting flow is shown in Fig. 4.6.

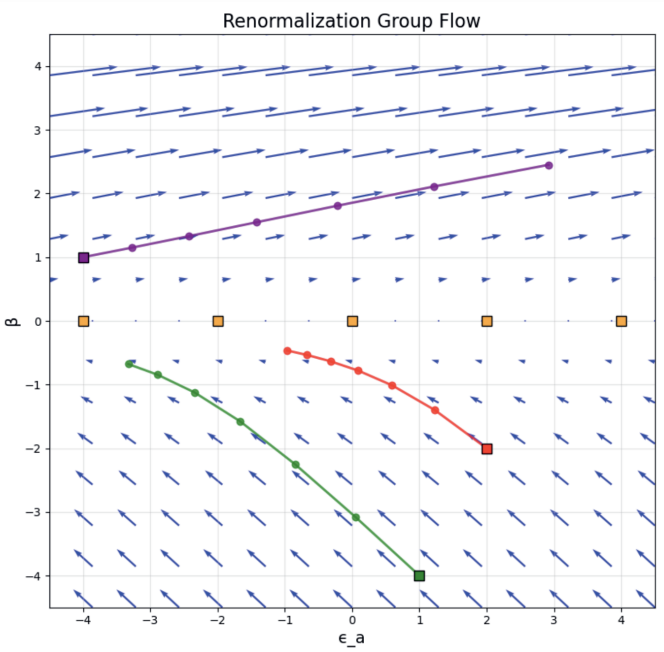


Figure 4.6: RG flow of the equations (4.41) *and* (4.42) *with a choice of* $\epsilon_{\bar{a}} = -5$.

Inference on the unobservable space

We now turn to the inverse problem. Suppose that the parameters of the observable multiplex (all layers except $\bar{a}$) can be estimated from data. The renormalization equations can then be interpreted as constraints relating observable effective parameters to the underlying microscopic ones:

$$E^{(a)} = \epsilon^{(a)} + \log \frac{1 + e^{\epsilon^{(\bar{a})}+\beta}}{1 + e^{\epsilon^{(\bar{a})}}}, \tag{4.43}$$

$$B = \beta + \log \frac{\left(1 + e^{\epsilon^{(\bar{a})}+2\beta}\right)\left(1 + e^{\epsilon^{(\bar{a})}}\right)}{\left(1 + e^{\epsilon^{(\bar{a})}+\beta}\right)^2}, \tag{4.44}$$

where $(E^{(a)}, B)$ are measured, while the unknowns are the true microscopic parameters $\epsilon^{(a)}$, β, and the hidden-layer field $\epsilon^{(\bar{a})}$.

These two equations for three unknowns define a one-dimensional manifold in parameter space, implying that the inversion problem is underdetermined. A naive closure would be to assume homogeneous local fields, $\epsilon^{(a)} = \epsilon$, reducing the number of unknowns. However, such an approximation is overly restrictive: inter-layer coupling already acts as a homogenizing mechanism, and forcing identical local fields risks erasing genuine heterogeneity.

A more principled resolution emerges from the discussion in Sec. 4.3.1. Marginalization does not only renormalize pairwise couplings: it systematically generates higher-order interactions. These additional terms encode information that is implicitly present in the reduced model and can therefore be used as extra constraints in the inverse problem. In the present case, the leading correction is the induced third-order interaction

$$\Delta_{III} \equiv \tilde{h}^{(3)} = \log \frac{\left(1 + e^{\epsilon^{(\bar{a})}+\beta}\right)^3 \left(1 + e^{\epsilon^{(\bar{a})}+3\beta}\right)}{\left(1 + e^{\epsilon^{(\bar{a})}}\right)\left(1 + e^{\epsilon^{(\bar{a})}+2\beta}\right)^3}. \tag{4.45}$$

The effective Hamiltonian in the observed space therefore reads

$$\begin{aligned} \tilde{H}(A_{ij}) = & -\sum_{a\neq\bar{a}} \tilde{\epsilon}^{(a)} A_{ij}^{(a)} - \tilde{\beta} \sum_{\substack{a\neq b \\ a,b\neq\bar{a}}} A_{ij}^{(a)} A_{ij}^{(b)} \\ & - \Delta_{III} \sum_{\substack{a,b,c\neq\bar{a} \\ \text{distinct}}} A_{ij}^{(a)} A_{ij}^{(b)} A_{ij}^{(c)} + \dots \end{aligned} \tag{4.46}$$

together with higher-order terms.

Crucially, this Hamiltonian describes the statistics of the observable system. Therefore, empirical evidence that the reduced multiplex is well captured by a pairwise model implies that higher-order effective couplings must be small. This provides a natural closure condition: instead of imposing homogeneity assumptions, we constrain the magnitude of the induced higher-order terms.

We thus introduce the inequality

$$\Delta_{III} \leq R, \tag{4.47}$$

where R quantifies the tolerance with which the observed system is described by a second-order model. Operationally, R can be interpreted as a model-selection threshold, e.g. derived from goodness-of-fit tests or information criteria comparing pairwise and higher-order null models.

Geometrically, Eqs. (4.43)–(4.44) define a curve of admissible microscopic parameters. The additional constraint (4.47) restricts this curve to a smaller subset, effectively regularizing the inversion. As illustrated in Fig. 4.7, decreasing R progressively shrinks the admissible region, collapsing the degeneracy and yielding a nearly unique reconstruction when the reduced system is strongly pairwise-dominated.

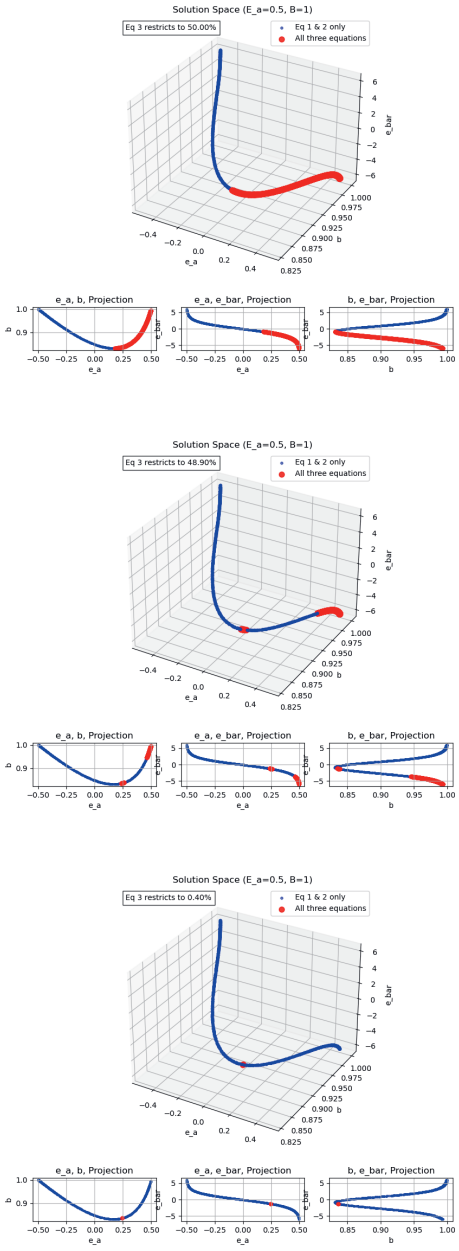


Figure 4.7: The reconstruction problem defined by (4.43) *and* (4.44)*. In blue the line of points in the parameter space* $\epsilon^{(a)}$*,* β*,* $\epsilon^{(\bar{a})}$ *that are compatible with the RG equations, with measured parameters* $E^{(a)} = 0.5$*,* $B = 1$*. In red the sub-manifold defined by adding the third equation* (4.47) *with* R *respectively* $1, 0.1, 0.01$ *from top to bottom.*

§4.5 Final remarks

In this chapter we tried to introduce a novel perspective on network reconstruction problems in which missing information is localized on individual links or layers. The framework was developed with information loss as primary direction, yet it naturally reveals a deeper connection with renormalization ideas familiar from theoretical physics. In particular, marginalization over unobserved variables plays the role of a decimation procedure, generating a flow in model space analogous to renormalization group transformations.

Starting from a general random network model, we derived explicit flow equations for the effective parameters under link marginalization. A central result is that renormalization generically induces higher-order effective interactions: even when the original model is pairwise, the reduced description typically involves many-body couplings.

Rather than viewing this complexity as an obstruction, we proposed to invert the perspective and use renormalization as an inference tool. In this "top-down" approach, one starts from an effective model fitted on the observable space and exploits the renormalization equations as constraints on the underlying microscopic parameters. We illustrated this strategy in two complementary scenarios. In the Markovian case (Sec. 4.4.1), the renormalization flow closes exactly, allowing for direct inversion. In more complex settings such as multiplex all-to-all models (Sec. 4.4.2), closure is lost but additional information can be extracted from the structure of induced higher-order terms. There, controlled heuristics based on the smallness of effective many-body interactions provide a principled way to regularize the inverse problem and recover otherwise hidden parameters. Beyond these specific examples, the broader message of this chapter is that marginalization-induced renormalization provides a unifying language to study reconstruction under partial observability. The framework suggests a new class of inverse problems in which missing degrees of freedom are not treated as noise, but as generators of structured flows in parameter space. This viewpoint opens several directions for future work. For instance, extending the formalism to temporal networks could help address the time-resolution problem that often arises in empirical data, where coarse temporal sampling effectively marginalizes over fast interactions. Understanding how uncertainty flows across temporal scales may therefore offer new tools for reconstructing dynamical networks.

Finally, this chapter is closely connected with the next Chapters in this thesis. Namely, in Chapter 5, we analyze in detail the class of models that admit a mapping to Markov chains and study the possibility of renormalization inversion in closed form, providing the conceptual foundation for the exact reconstruction results discussed here. In Chapter 6, we develop a general solution for second-order random network models by deriving computable expressions for their partition functions. This makes maximum-likelihood inference feasible for a wide class of pairwise models, thereby supplying the computational backbone required to implement the reconstruction schemes proposed in the all-to-all setting to virtually any other effectively binary Random Network Model.

CHAPTER 5

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 priors[1]. 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

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

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:

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

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:

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

[1] 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 [111, 112].

subject to maximization of the Gibbs-Shannon entropy

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

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 $h^{(1)}_{ij} = h^{(1)}_{ji} = \log[p/(1-p)]$ for $i \neq j$ (see e.g. [114]). A disordered variant of the Erdős-Renyi graph is obtained by allowing the $h^{(1)}_{ij}$ to vary link-wise, so that

$$h^{(1)}_{ij} = \log\left(\frac{p_{ij}}{1-p_{ij}}\right). \tag{5.5}$$

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

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

where each individual factor corresponds to the partition function for each possible individual link[2] — 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

[2]For 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 accuracy[3]. 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

[3]This 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 *ir*relevant. 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:

$$\frac{d\rho}{dt} = \sigma\rho + D_2\nabla^2\rho + D_4\ell^2\nabla^4\rho + D_6\ell^4\nabla^6\rho + ... \tag{5.7}$$

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

[4]The 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 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 theories[5], 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 categoric 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].

[5] 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 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,

$$Z \neq Z_1 Z_2 ... Z_{\mathcal{N}}, \tag{5.8}$$

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:

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

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 $h^{(2)}_{ijkl}$ 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 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:

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

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 graph[6]. 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:

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

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

so that

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

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 couplings[7]:

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

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:

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

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

[7] Up to an overall rescaling of the partition function that is equivalent to a localized additive free energy contribution given by $F_n = \log\left[1+\varepsilon_n\right]$, corresponding to the individual contribution of the marginalized over link.

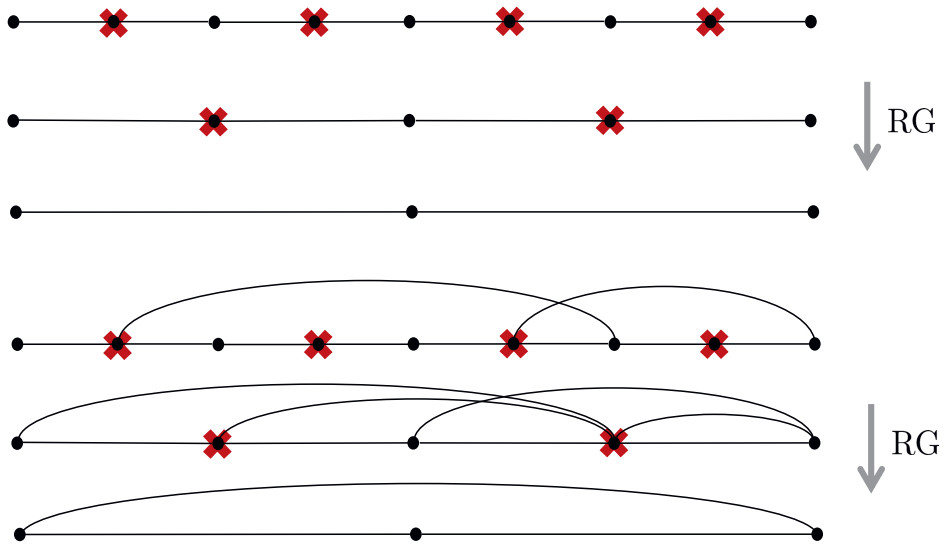


Figure 5.1: Schematic rendering of maximum coordination two couplings (top) and a single bare coordination number three coupling (bottom) under RG transformations. The red crosses indicate the sites that are decimated at each RG step. More generally, successive RG steps increase the maximum coordination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

$$H \supset \left[\varepsilon_m + g_m A_{m+1} + g'_m A_{m+k}\right] A_a \tag{5.15}$$

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 ≡ 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:

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

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

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

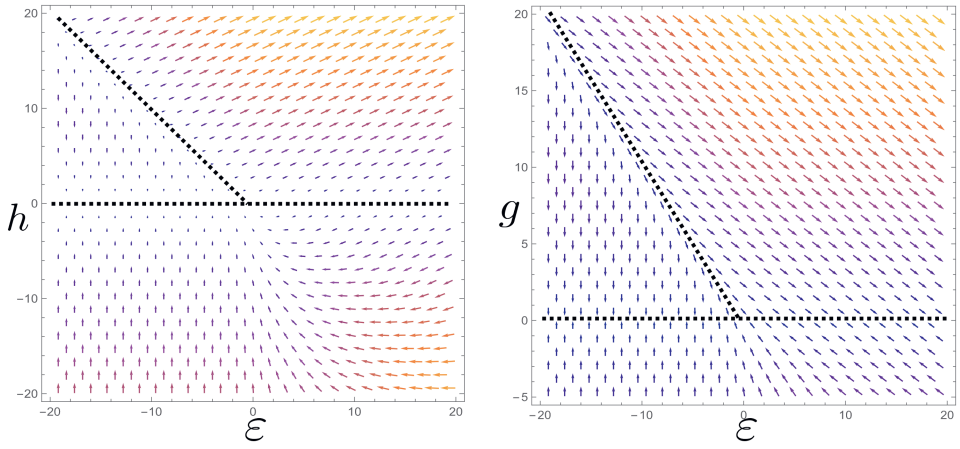


Figure 5.2: RG flow for homogeneous coordination number two interactions towards an Erdős-Renyi fixed line, where arrows indicate direction of flow and color determine the its intensity, lighter being stronger. Upper panel: A shearing behavior for positively correlated conditioning is apparent, along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

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

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-Renyi 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. $\widetilde{g} > 0$) 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

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

the result will be a Hamiltonian of the form:

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

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 plane[9]. 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.

[9]This is made explicit by considering the beta function for the linear combination $h = \varepsilon + g$, which is given by $\beta_h = \log\left\{\left(e^{g+\varepsilon} + e^{-g}\right)/\left(1 + e^{\varepsilon}\right)\right\}$, 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:

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

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 $\widetilde{y}_n = \widetilde{y}_n(y_n, y_{n+1}, x_{n+1})$, and $\widetilde{x}_n = \widetilde{x}_n(x_{n-1}, x_n, x_{n+1}, y_{n-1}, y_n)$, it must be the case that under decimation, successive RG transformations induce a flow on the probability assignments given by

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

and

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

where the dependent variables $\widetilde{y}_n$ and $\widetilde{x}_n$ 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 $\widetilde{y}_n$. To do this, we must first invert $\widetilde{y}_n$ 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

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

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

integral, one ends up with

$$\widetilde{\psi}_n(y) = \int_0^\infty\!\!\int_0^\infty dy_{n+1}dx_{n+1}\Big\{\Theta(y_n)\psi_n(y_n[x_{n+1},y_{n+1},y]) \\ \frac{\psi_{n+1}(y_{n+1})\chi_{n+1}(x_{n+1})(1+x_{n+1})(1+y_{n+1}x_{n+1})|y_{n+1}-1|}{x_{n+1}\left[(1+y_{n+1}x_{n+1})y-(1+x_{n+1})y_{n+1}\right]^2}\Big\}, \tag{5.24}$$

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

$$\widetilde{\chi}_n(\widetilde{x}_n\equiv x) = \int_0^\infty\!\!...\int_0^\infty dy_{n-1}dy_n dx_{n-1}dx_{n+1}\Big\{\chi_n(x_n[...]) \\ \psi_{n-1}(y_{n-1})\psi_n(y_n)\chi_{n-1}(x_{n-1})\chi_{n+1}(x_{n+1}) \\ \times\frac{(1+x_{n+1})(1+x_{n-1})}{(1+y_n x_{n+1})(1+y_{n-1}x_{n-1})}\Big\}, \tag{5.25}$$

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

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

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:

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

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

$$\widetilde{P}_n(\quad \widetilde{g}_n\equiv g) = \int_{-\infty}^\infty\!\!\int_{-\infty}^\infty dg_{n+1}d\varepsilon_{n+1}\Big\{\Theta(g_n)P_{n+1}(g_{n+1}) \\ \frac{e^g(1+e^{\varepsilon_{n+1}})(1+e^{g_{n+1}+\varepsilon_{n+1}})|e^{g_{n+1}}-1|}{e^{\varepsilon_{n+1}}\left[(1+e^{g_{n+1}+\varepsilon_{n+1}})e^g-(1+e^{\varepsilon_{n+1}})e^{g_{n+1}}\right]^2} \\ \times C_{n+1}(\varepsilon_{n+1})P_n(g_n[\varepsilon_{n+1},g_{n+1},\widetilde{g}_n\equiv g])\Big\}, \tag{5.28}$$

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:

$$\widetilde{C}_n(\widetilde{\varepsilon}_n\equiv\varepsilon) = \int_{-\infty}^\infty\!\!...\int_{-\infty}^\infty dg_{n-1}dg_n d\varepsilon_{n-1}d\varepsilon_{n+1}\Big\{P_n(g_n) \\ P_{n-1}(g_{n-1})C_{n-1}(\varepsilon_{n-1})\frac{(1+e^{\varepsilon_{n+1}})(1+e^{\varepsilon_{n-1}})}{(1+e^{g_n+\varepsilon_{n+1}})(1+e^{g_{n-1}+\varepsilon_{n-1}})} \\ \times C_{n+1}(\varepsilon_{n+1})C_n(\varepsilon_n[g_{n-1},g_n,\varepsilon_{n-1},\varepsilon_{n+1},\widetilde{\varepsilon}_n\equiv\varepsilon])\Big\}. \tag{5.29}$$

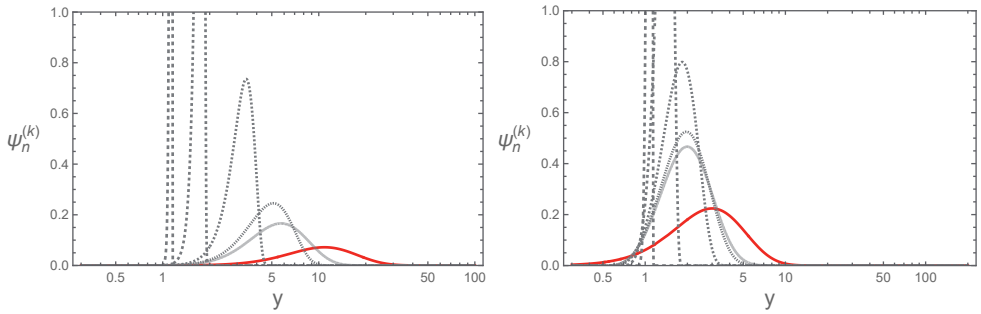


Figure 5.3: Top panel: Renormalization group induced deformation of the bare probability assignment $\Pi(y)$ *given by Eq. 5.35 for* $\alpha_\pi = 5$, $\lambda_\pi = e^{-1}$ *(red line) for* k *from 1 to 5 (gray dashed lines with successively coarser dashing). Bottom panel: The same, but for* $\Pi(y)$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

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

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:

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

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

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:

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

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 Erdös-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 sites[10]. 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]).

[10] In 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:

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

We now presume the disorder to be uniformly assigned as:

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

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 site[11]:

$$\begin{aligned} \chi^{(k+1)}(x) = \int_0^\infty \!\!\ldots \int_0^\infty du\, dv\, dy\, dw \Bigg\{ & \frac{(1+w)(1+v)}{(y+uw)} \\ & \psi^{(k)}\Big[\frac{x(1+w)(1+v)}{(y+wu)} - 1\Big]_+ \psi^{(k)}(u)\chi^{(k)}(v)\chi^{(k)}(y)\chi^{(k)}(w) \Bigg\}, \end{aligned} \tag{5.39}$$

and

$$\begin{aligned} \psi^{(k+1)}(z) = \int_0^\infty \!\!\ldots \int_0^\infty du\, dv\, dy\, dw \Bigg\{ & \psi^{(k)}(v)\psi^{(k)}(y)\psi^{(k)}(w) \\ & \frac{(u+yw)(1+v)}{z^2(1+w)} \chi^{(k)}\Big[\frac{(u+yw)(1+v)}{z(1+w)} - 1\Big]_+ \chi^{(k)}(u) \Bigg\}, \end{aligned} \tag{5.40}$$

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

[11] Here, 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.

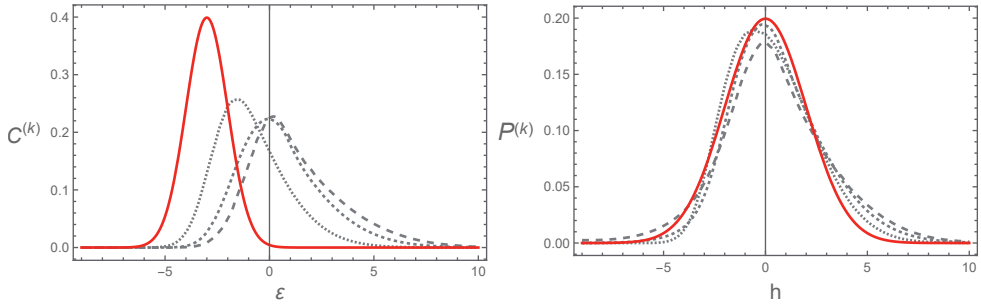


Figure 5.4: Renormalization group induced deformation of the uniform probability assignments $C^{(k)}(\varepsilon)$ (top) and $P^{(k)}(h)$ (bottom). Bold red lines denote the bare probability assignments, taken to be Gaussians with means and variances given by $\mu_\varepsilon = -3, \sigma_\varepsilon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

cases by noticing that both convolutions can be expressed as

$$\chi^{(k+1)}(x) = \frac{d}{dx} \int_0^\infty \!\!\ldots \int_0^\infty du\, dv\, dy\, dw \Big\{ \Psi^{(k)}\left[f(x|u,v,y,w)\right] \times \psi^{(k)}(u)\chi^{(k)}(v)\chi^{(k)}(y)\chi^{(k)}(w) \Big\}, \tag{5.41}$$

$$\psi^{(k+1)}(z) = -\frac{d}{dz} \int_0^\infty \!\!\ldots \int_0^\infty du\, dv\, dy\, dw \Big\{ \Xi^{(k)}\left[g(z|u,v,y,w)\right] \times \chi^{(k)}(u)\psi^{(k)}(v)\psi^{(k)}(y)\psi^{(k)}(w) \Big\}. \tag{5.42}$$

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:

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

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:

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

and

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

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:

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

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

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:

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

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:

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

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. 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

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

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 link[12]. 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

[12] With 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.

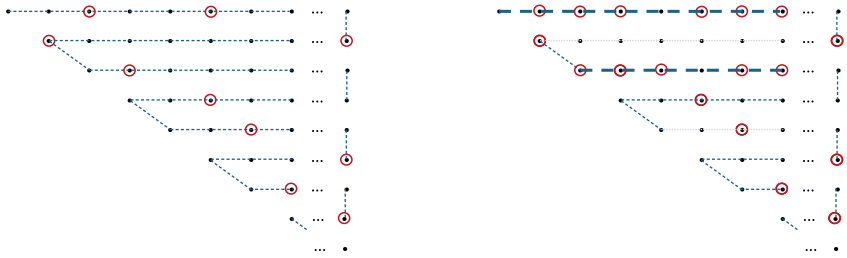


Figure 5.6: Representation of a random graph with maximum coordination two interactions (top panel). Enhanced or diminished attachment and preference effects can be implemented via the sign and magnitude of the relevant bilinear couplings $\{g_a\}$ (bottom panel: bold and faint lines, respectively).

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 $\widetilde{N} = N/2$ nodes[13]. 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 $\widetilde{N} = N/2^m$ 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.

[13]There 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-Renyi 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 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-gained' 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].

CHAPTER 6

The Harmonic Oscillator of Random Graphs

This chapter is based on:
Alessio Catanzaro, Subodh Patil, and Diego Garlaschelli. *"A General Solution for Network Models with Pairwise Edge Coupling"*. arXiv e-prints (2025): arXiv-2503.

Abstract

Network Models with couplings between link pairs are the simplest models for a class of networks with Higher Order interactions. In this paper we give an analytic, general solution to this family of Random Graph Models extending previous results obtained for specific interaction structures. We use the Hubbard-Stratonovich transform to show that such higher order models can be decoupled into families of effectively separable random graph models, which might be enhanced with auxiliary, possibly multi-valued, parameters identifying the thermodynamical phases of the system. We study the diagonalizability of the matrix of the interactions between links, that is itself an adjacency matrix on the dual network whose vertices are the nodes pairs, that can be dealt with using tools of algebraic graph theory. By orthogonalizing the interaction, we carry out the full computation of the partition functions and discuss why in some cases interactions effectively reduce to their Mean-Field versions.

§6.1 Introduction

Complex system science is heavily reliant on networks to model a multitude of systems. Random networks are among the most used tools to encode both the random nature of such objects as well as their evident heterogeneity. A statistical mechanics approach to random networks has been widely used by the network science community in the last 20 years [48, 47, 90, 110]. Yet a fundamental hurdle in the modelling of such systems with the aforementioned approaches is that few among the innumerable conceivable models of random networks can be solved, and if so it is often through approximate techniques that are model-dependent.

First order networks are easily solved, as the Hamiltonian describing them has only link specific parameters, which gives rise to a separable partition function and hence independent, link-specific probabilities. Regretfully when any form of coupling among links are introduced one has often to resort to either trivial toy-models of the interaction, or to be satisfied with approximated solutions only [48, 50].

The purpose of this paper is to partially cover this knowledge gap by showing how a general solution for any second order unweighted, undirected random network can be written in an exact way, and to solve for some cases of interest by showing some non-trivial behaviour. We also note that the analytical techniques we use could be extended to provide for a solution of weighted and directed random networks, but we do not believe the physical behaviour of such models to be fundamentally different than the one we present here in terms of phase transitions.

Our approach relies on the implementation of the multidimensional Hubbard-Stratonovich transform to account for any coupling structure between links. This also implies an effective description of the model through independent links, conditional on a given order parameter that can be treated as a random variable (and being a function of the model given parameters). Our solution pipeline also extends some results already found in the literature in order to put them under a single umbrella of explicative framework.

A further motivation for obtaining such a general solution comes from network inference. Ordinarily, exponential random graph models with pairwise dependence between links have an intractable normalizing constant, which forces practitioners to resort to approximate estimationschemes such as Markov Chain Monte Carlo maximum likelihood or related curved exponential-family techniques, both of which are known to be computationally expensive and prone to model degeneracy. Because the solution we derive gives explicit access to the partition function and hence to the full likelihood of any second order model in this class, it removes this bottleneck: the log-likelihood of an observed network can be evaluated, and in many cases maximized, directly and exactly, rather than estimated through simulation. This places second order, pairwise-coupled models on the same operationally tractable footing that first order maximum-entropy models already enjoy for network reconstruction and null-model based inference [90].

The Chapter is organized as follows: in Section 6.2 we introduce the techniques to solve the general problem, the pipelines to compute the partition function and write the single link probability function in an exact, although not explicit, manner.

In Section 6.3 we show how the machinery works for some cases of second order Random Network Models (RNM) to give meaning to our formulae, and we give explicit solutions for a few models of interest. Some models for which we solve were actually already found in the literature, but as stand-alone works and without the general, explicative framework we provide. Moreover, in the case of the model of 6.3.4, we highlight a connection between the model already analyzed in [49**?**] to the field of Coding Theory [155]. Finally in Section 6.4 we comment on how to give a rule of thumb to determine the possible phase structure of the models at stake and if the mean field approximation can give a satifying picture of the whole phase space.

§6.2 Solution to II order Random Network

An Random Network Model (RNM) is a probability law $P(\mathbf{A})$ for the realizations of the adjacency matrix $\mathbf{A}$ of a network. We choose to deal with unweighted, undirected networks so that $\mathbf{A}$ is symmetric and its entries can be either 0 or 1. This latter fact implies the fundamental relation

$$A_{ij} = A_{ij}^m \quad \forall m \tag{6.1}$$

that simplifies enormously the computations that follow along the entire chapter.

For instance, we can use relation (6.1) to show that we can always represent the probability function describing the model in the exponential form

$$P(\mathbf{A}) = \frac{e^{-H(\mathbf{A})}}{Z} \quad Z = \sum_{\{\mathbf{A}\}} e^{-H(\mathbf{A})} \tag{6.2}$$

where the Hamiltonian is in the form to eventually enforce constraints

$$H(A) = \sum_l \theta_l f_l(\mathbf{A}). \tag{6.3}$$

If one desires to enforce the full ERG formalism [48, 47, 90, 110], then has to add the constraint equations such that

$$\sum_{\{\mathbf{A}\}} f_l(\mathbf{A}) P(\mathbf{A}) = f_l^* \quad \forall l, \tag{6.4}$$

where f_l^* is the 'target' expected value of the property f_l, which can be interpreted in two ways. The first interpretation is to take f_l^* as the empirical value $f_l^* = f_l(\mathbf{A}^*)$ from a real observed matrix $\mathbf{A}^*$. A second interpretation is one where the parameter space of the model is explored exhaustively and f_l^* is the expected value of the property over the ensemble implied by the chosen parameters. We will follow this latter interpretation throughout this chapter.

Moreover, again making use of relation (6.1) the Hamiltonian function can always be written as a multilinear function of the adjacency matrix elements, in the form

$$H(\mathbf{A}) = h^0 + \theta_1 \sum_{i,j} h^1_{ij} A_{ij} + \theta_2 \sum_{i,j,k,l} h^2_{ijkl} A_{ij} A_{kl} + \\ + \theta_3 \sum_{i,j,k,l,m,n} h^3_{ijklmn} A_{ij} A_{kl} A_{mn} + \dots . \quad (6.5)$$

The above expansion is analogous to similar computation performed in the context of QFT for Grassmannian variable, where the relation (6.1) is replaced by the anticommuting relation.

We now proceed to show how a solution to the second order approximation of a model defined by the Hamiltonian (6.5) can be found in general.

§6.2.1 Evaluation of the partition function

Our task is to compute the partition function or equivalently the single link probability function of the model at stake. We have to maipulate the second order term of equation (6.5). The term h^0 can obviously be ignored as it is just a multiplicative constant in front of the partition function. We are left with

$$H(\mathbf{A}) = -\sum_{i<j} h^1_{ij} A_{ij} - \sum_{i<j,k<l} h^2_{ijkl} A_{ij} A_{kl} \quad (6.6)$$

First and foremost we change the labels of the indices to get a more manageable expression, since we see that the sums can be expressed in the couples of indices (i, j) which we order with index a going from 1 to $V = \frac{N(N-1)}{2}$ (without self-loops). We then express the interaction coefficient in order to scale with V as $h^2_{ijkl} = \frac{\beta_{ijkl}}{2V}$. The new Hamiltonian becomes

$$H(\mathbf{A}) = -\sum_{a=1}^{V} h^1_a A_a - \frac{1}{2V} \sum_{a \neq b}^{V} \beta_{ab} A_a A_b \quad (6.7)$$

we further manipulate this object to have a compact term to operate with. First we see that

$$\frac{1}{2} \sum_{a \neq b}^{V} \beta_{ab} A_a A_b = \frac{1}{2V} \sum_{a,b}^{V} \beta_{ab} A_a A_b - \frac{1}{2V} \sum_{a} \beta_{aa} A_a = \frac{1}{2V} \langle A|\beta|A\rangle - \frac{1}{2V} \langle A|\vec{\beta}\rangle \quad (6.8)$$

where $\vec{\beta}^T = (\beta_{11}, \beta_{22}, \cdots, \beta_{VV})$. We can readily see that this term can be reabsorbed in the local field term (and moreover that it is vanishing in the large V limit).

Then we write also the first term in the braket notation so that the whole Hamiltonian is in the form

$$H(\mathbf{A}) = -\sum_{a}^{V} A_a \left(h_a^1 + \frac{\beta_{aa}}{2V} \right) - \frac{1}{2V} \sum_{a,b}^{V} \beta_{ab} A_a A_b = -\langle A|\epsilon\rangle - \frac{1}{2V} \langle \mathbf{A}|\beta|\mathbf{A}\rangle \quad (6.9)$$

where β is a matrix that couple the links between them selves. This matrix is an adjacency matrix (either expected or realized) in disguise, over the network of links of the actual graph (dual or line graph in the literature), and we will make use of the properties of such a matrix to derive the behaviour of the models.

We assume the β matrix is symmetric (as the correlation between links is). Then it admits a spectral representation

$$\beta = \sum_{k=1}^{r} \lambda_k \left|\omega_k\right\rangle \left\langle\omega_k\right| \quad (6.10)$$

where r is the rank of the matrix, λ_k are the eigenvalues and $|\omega_k\rangle$ their respective orthonormal eigenvectors. We also split the positive spectrum part from the negative one. Let then r_+ be the number of positive eigenvalues, (including their multiplicities) and $r - r_+$ the number of the negative ones.

The Hamiltonian can be hence written as

$$H(\mathbf{A}) = -\langle A|\epsilon\rangle - \sum_{k=1}^{r_+} \frac{\lambda_k}{2V} \langle \mathbf{A}|\omega_k\rangle^2 - \sum_{k=r_++1}^{r} \frac{\lambda_k}{2V} \langle \mathbf{A}|\omega_k\rangle^2 \quad (6.11)$$

where β is the interaction matrix in which all the coupling between links is encoded, and V is the total number of possible links. Note also that the local parameter is $\epsilon_{ij} = h_{ij}^1 + \frac{\beta_{ijij}}{2V}$, clearly reducing to h_{ij}^1 in the large N limit.

All the ingredients are now in place to evaluate the partition function, as this latter form of equation (6.11) can be exponentiated as

$$e^{-H(\mathbf{A})} = e^{\langle \mathbf{A}|\epsilon\rangle + \sum_{k=1}^{r_+} \frac{\lambda_k}{2V} \langle \mathbf{A}|\omega_k\rangle^2 + \sum_{k=r_++1}^{r} \frac{\lambda_k}{2V} \langle \mathbf{A}|\omega_k\rangle^2} \quad (6.12)$$

$$= e^{\langle \mathbf{A}|\epsilon\rangle} \prod_{k=1}^{r_+} e^{\frac{\lambda_k}{2V} \langle \mathbf{A}|\omega_k\rangle^2} \prod_{k=r_++1}^{r} e^{\frac{\lambda_k}{2V} \langle \mathbf{A}|\omega_k\rangle^2} \quad (6.13)$$

and this form is very amenable to an Hubbard Stratonovich transformation. This integral representation allows us to disentangle the squares as

$$e^{a^2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx e^{-\frac{x^2}{2} + \sqrt{2}ax} \quad (6.14)$$

if $a > 0$ and

$$e^{-a^2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx e^{-\frac{x^2}{2} + i\sqrt{2}ax} \quad (6.15)$$

if $a < 0$. This allows to write

$$e^{-H(A)} = e^{\langle A|\epsilon\rangle} \prod_{k=1}^{r_+} \left[\frac{1}{\sqrt{2\pi}} \int dx e^{-\frac{x^2}{2} + x\sqrt{\frac{\lambda_k}{V}}\langle A|\omega_k\rangle} \right] \prod_{k=r_++1}^{r} \left[\frac{1}{\sqrt{2\pi}} \int dx e^{-\frac{x^2}{2} + ix\sqrt{\frac{\lambda_k}{V}}\langle A|\omega_k\rangle} \right] \tag{6.16}$$

which we want to compactly write as

$$\prod_{k=1}^{r} \left[\frac{1}{\sqrt{2\pi}} \int dx e^{-\frac{x^2}{2} + x\sqrt{\frac{\lambda_k}{V}}\langle A|\omega_k\rangle + \frac{\langle A|\epsilon\rangle}{r}} \right] \tag{6.17}$$

by making sure we take the proper square root, so that

$$\sqrt{\lambda_k} = \begin{cases} |\sqrt{\lambda_k}| \text{ if } \lambda_k > 0 \\ i|\sqrt{|\lambda_k|}| \text{ if } \lambda_k < 0 \end{cases} \tag{6.18}$$

The last manipulation of this term is to realize that this product of integrals can be written as a repeated integral as

$$\begin{aligned} e^{-H(\mathbf{A})} &= \left(\frac{1}{\sqrt{2\pi}}\right)^r \int \cdots \int dx_1, \cdots dx_r e^{-\sum_{k=1}^r \frac{x_k^2}{2}} \prod_{k=1}^r e^{x_k\sqrt{\frac{\lambda_k}{V}}\langle A|\omega_k\rangle + \frac{\langle A|\epsilon\rangle}{r}} = \\ &= \left(\frac{1}{\sqrt{2\pi}}\right)^r \int \cdots \int dx_1, \cdots dx_r e^{-\sum_{k=1}^r \frac{x_k^2}{2}} \prod_{k=1}^r e^{\sum_{a=1}^V x_k\sqrt{\frac{\lambda_k}{V}} A_a(\omega_k)_a + \frac{A_a\epsilon_a}{r}} = \\ &= \left(\frac{1}{\sqrt{2\pi}}\right)^r \int \cdots \int dx_1, \cdots dx_r e^{-\sum_{k=1}^r \frac{x_k^2}{2}} \prod_{k=1}^r \prod_{a=1}^V e^{x_k\sqrt{\frac{\lambda_k}{V}} A_a(\omega_k)_a + \frac{A_a\epsilon_a}{r}} \end{aligned} \tag{6.19}$$

Now it is the time to integrate over the configurations

$$\begin{aligned} Z &= \sum_{\{A_a=0,1\}} e^{-H(\mathbf{A})} = \\ &= \sum_{\{A_a=0,1\}} \left(\frac{1}{\sqrt{2\pi}}\right)^r \int \cdots \int dx_1, \cdots dx_r e^{-\sum_{k=1}^r \frac{x_k^2}{2}} \prod_{k=1}^r \prod_{a=1}^V e^{x_k\sqrt{\frac{\lambda_k}{V}} A_a(\omega_k)_a + \frac{A_a\epsilon_a}{r}} = \\ &= \left(\frac{1}{\sqrt{2\pi}}\right)^r \int \cdots \int dx_1, \cdots dx_r e^{-\sum_{k=1}^r \frac{x_k^2}{2}} \prod_{k=1}^r \prod_{a=1}^V \sum_{\{A_a=0,1\}} e^{x_k\sqrt{\frac{\lambda_k}{V}} A_a(\omega_k)_a + \frac{A_a\epsilon_a}{r}} = \\ &= \left(\frac{1}{\sqrt{2\pi}}\right)^r \int \cdots \int dx_1, \cdots dx_r e^{-\sum_{k=1}^r \frac{x_k^2}{2}} \prod_{k=1}^r \prod_{a=1}^V \left[1 + e^{x_k\sqrt{\frac{\lambda_k}{V}}(\omega_k)_a + \frac{\epsilon_a}{r}}\right] \end{aligned} \tag{6.20}$$

and finally we can regroup the integrals in the form

$$Z = \prod_{k=1}^r \left[\sqrt{\frac{1}{2\pi}} \int dx e^{-\frac{x^2}{2}} \prod_{a=1}^V \left[1 + e^{x\sqrt{\frac{\lambda_k}{V}}(\omega_k)_a + \frac{\epsilon_a}{r}}\right] \right] \tag{6.21}$$

which allows for some further computation when the structure of the matrix β is known.

§6.2.2 Effective Separability of the partition function

We know from general Statistical Mechanics theory that the model defined by (6.6) might undergo phase transitions, with a phase structure dictated by the structure of the interaction matrix β. In the Hubbard-Stratonovich approach, the multiplicity of the possible phases of the model can be encoded in the vector of values $(x_1^*, x_2^*, \ldots)$ where each x_k^* can be either single-valued or multi-valued depending on the other parameters of the model. Conditionally on the values of $(x_1^*, x_2^*, \ldots)$, the model can be rewritten in an effectively noninteracting form where the full partition function (6.21) can be factorized.

A separable partition function in the form

$$Z = \prod_{a=1}^{V} z_a \tag{6.22}$$

is indeed what one would like to work with, since it has a straightforward connection to the probability of link formation. As a matter of facts, when the link variables are not coupled there is an identification between the terms in the product of Z and the bare link probabilities. This can be shown from

$$Z = \sum_{\{A_a=0,1\}} e^{-H(\mathbf{A})} = \sum_{\{A_a=0,1\}} \prod_{a=1}^{V} e^{h_a^{(1)} A_a} = \prod_{a=1}^{V} \left(1 + e^{h_a^{(1)}}\right) = \prod_{a=1}^{V} \frac{1}{1-p_a} \tag{6.23}$$

The formula (6.21) is amenable to show that this is the case also for second order interacting networks. As a matter of fact say that you isolate x_1 as

$$Z = \left(\frac{1}{\sqrt{2\pi}}\right)^{r-1} \int \cdots \int dx_2, \cdots dx_r e^{-\sum_{k=2}^{r} \frac{x_k^2}{2}} \prod_{k=2}^{r} \prod_{a=1}^{V} \left[1 + e^{x_k \sqrt{\frac{\lambda_k}{V}} (\omega_k)_a + \frac{\epsilon_a}{r}}\right] \cdot$$
$$\cdot \underbrace{\frac{1}{\sqrt{2\pi}} \int dx_1 e^{-\frac{x_1^2}{2}} \prod_{a=1}^{V} \left[1 + e^{x_1 \sqrt{\frac{\lambda_1}{V}} (\omega_1)_a + \frac{\epsilon_a}{r}}\right]}_{J_1}. \tag{6.24}$$

Now J_1 can be seen as a Gaussian expected value of a certain function, and it is clearly convergent, indeed

$$J_1 = \int dx_1 \mathcal{N}(0,1) \prod_{a=1}^{V} \left[1 + e^{x_1 \sqrt{\frac{\lambda_1}{V}} (\omega_1)_a + \frac{\epsilon_a}{r}}\right] = E_{\mathcal{N}(0,1)}\left[f(x_1)\right] = \mu \tag{6.25}$$

where we defined

$$f(x_1) \equiv \prod_{a=1}^{V} \left[1 + e^{x_1 \sqrt{\frac{\lambda_1}{V}} (\omega_1)_a + \frac{\epsilon_a}{r}}\right]. \tag{6.26}$$

Now we can subtract the expected value in the integral so that

$$0 = \int dx_1 \mathcal{N}(0,1) \left[\prod_{a=1}^{V} \left(1 + e^{x_1 \sqrt{\frac{\lambda_1}{V}} (\omega_1)_a + \frac{\epsilon_a}{r}} \right) - \mu \right]. \tag{6.27}$$

This last integral evaluates to 0, hence there must be at least one change of sign of the function

$$g(x_1) = \prod_{a=1}^{V} \left(1 + e^{x_1 \sqrt{\frac{\lambda_1}{V}} (\omega_1)_a + \frac{\epsilon_a}{r}} \right) - \mu \tag{6.28}$$

therefore, since the function is continuous, there must be at least a x_1^* in which $g(x_1^*) = 0$, so that there is at least one x_1^* in which

$$J_1 = f(x_1^*) = \prod_{a=1}^{V} \left[1 + e^{x_1^* \sqrt{\frac{\lambda_1}{V}} (\omega_1)_a + \frac{\epsilon_a}{r}} \right] \tag{6.29}$$

By repeating this argument for all the k, we can write the partition function in the form

$$Z = \prod_{k=1}^{r} \prod_{a=1}^{V} \left[1 + e^{x_k^* \sqrt{\frac{\lambda_k}{V}} (\omega_k)_a + \frac{\epsilon_a}{r}} \right] \tag{6.30}$$

then by switching the products, we can identify the single link partition function as

$$z_a^{\text{eff}} = \prod_{k=1}^{r} \left[1 + e^{x_k^* \sqrt{\frac{\lambda_k}{V}} (\omega_k)_a + \frac{\epsilon_a}{r}} \right]. \tag{6.31}$$

Finally, we can further manipulate this object to have a nicer representation

$$\begin{aligned} z_a^{\text{eff}} &= \sum_{A_a=0,1} \prod_{k=1}^{r} e^{A_a \left(x_k^* \sqrt{\frac{\lambda_k}{V}} (\omega_k)_a + \frac{\epsilon_a}{r} \right)} = \sum_{A_a=0,1} e^{A_a \sum_{k=1}^{r} \left(x_k^* \sqrt{\frac{\lambda_k}{V}} (\omega_k)_a + \frac{\epsilon_a}{r} \right)} \\ &= \sum_{A_a=0,1} e^{A_a \left(\epsilon_a + \sum_{k=1}^{r} x_k^* \sqrt{\frac{\lambda_k}{V}} (\omega_k)_a \right)} = 1 + e^{\epsilon_a} \prod_{k=1}^{r} e^{x_k^* \sqrt{\frac{\lambda_k}{V}} (\omega_k)_a} \equiv \frac{1}{1 - p_a^{\text{eff}}} \end{aligned} \tag{6.32}$$

or for the link probability itself

$$p_a^{\text{eff}} = \frac{e^{\epsilon_a + \sum_k x_k^* \sqrt{\frac{\lambda_k}{V}} (\omega_k)_a}}{1 + e^{\epsilon_a + \sum_k x_k^* \sqrt{\frac{\lambda_k}{V}} (\omega_k)_a}} \tag{6.33}$$

as

$$Z(x_1^*, x_2^*, \ldots) = \prod_{i<j} z_{ij}^{\text{eff}}(x_1^*, , x_2^*, \ldots) \tag{6.34}$$

where

$$z_{ij}^{\text{eff}}(x_1^*, x_2^*, \ldots) = 1 + e^{\epsilon_{ij}} \prod_{k=1}^{r} e^{x_k^* \sqrt{\frac{\lambda_k}{V}} (\omega_k)_{ij}} \tag{6.35}$$

$$= 1 + e^{\epsilon_{ij} + \sum_{k=1}^{r} x_k^* \sqrt{\frac{\lambda_k}{V}} (\omega_k)_{ij}} \tag{6.36}$$

in which the contribution of each link has been isolated. The last expression is equivalent to that corresponding to a modified, non-interacting ($\beta = 0$) version of the Hamiltonian in eq.(6.6), i.e.

$$H(\mathbf{A}|x_1^*, x_2^*, \ldots) = -\langle \mathbf{A}|\epsilon'(x_1^*, x_2^*, \ldots)\rangle \tag{6.37}$$

where we have introduced the modified parameters

$$|\epsilon'(x_1^*, x_2^*, \ldots)\rangle = |\epsilon\rangle + \sum_{k=1}^{r} x_k^* \sqrt{\frac{\lambda_k}{V}} \, |\omega_k\rangle \, . \tag{6.38}$$

for the partition function and

$$p_{ij}^{\text{eff}} = \frac{e^{\epsilon_{ij} + \sum_{k=1}^{r} x_k^* \sqrt{\frac{\lambda_k}{V}} (\omega_k)_{ij}}}{1 + e^{\epsilon_{ij} + \sum_{k=1}^{r} x_k^* \sqrt{\frac{\lambda_k}{V}} (\omega_k)_{ij}}} \tag{6.39}$$

for the link probability expression. Formulae (6.36) and (6.39), together with equation (6.21) are the core results of this paper. These two latter formulae however need some explanation as a new key parameter x_k^* is introduced.

§6.2.3 Phase structure

The price to pay to write Equations (6.36) (6.39) with such a product over node couples and to have an effectively separable partition function over links is that a phase structure emerges. In our case we encode this behaviour in the parameter x_k^*. In the derivation of Equation (**??**) (see SI), we learn that x_k^* is the value(s) of the auxiliary field x_k for which $f(x_k^*)$ assumes the value of the Gaussian average of $f(x_k)$ over the real axis. This parameter is global and depends only on the specific structure of the interaction, combined with the local fields ϵ_{ij}, and most importantly its unicity is not guaranteed. Actually, when the parameters combination allows, x_k^* splits into multiple values (found numerically) that, when insterted in Equation (6.39), give rise to different values of the single link probability. A scheme of what happens is shown in Figure 6.1.

In the cases in which the parameter splits, the actual value of the link probability changes, hence identifying different phases (possibly denser or sparser). When drawing a realization of the ensemble, one has to keep in mind that also the various values of x_k^* come each with its own probability and can be computed as we show in Section 6.3.1.

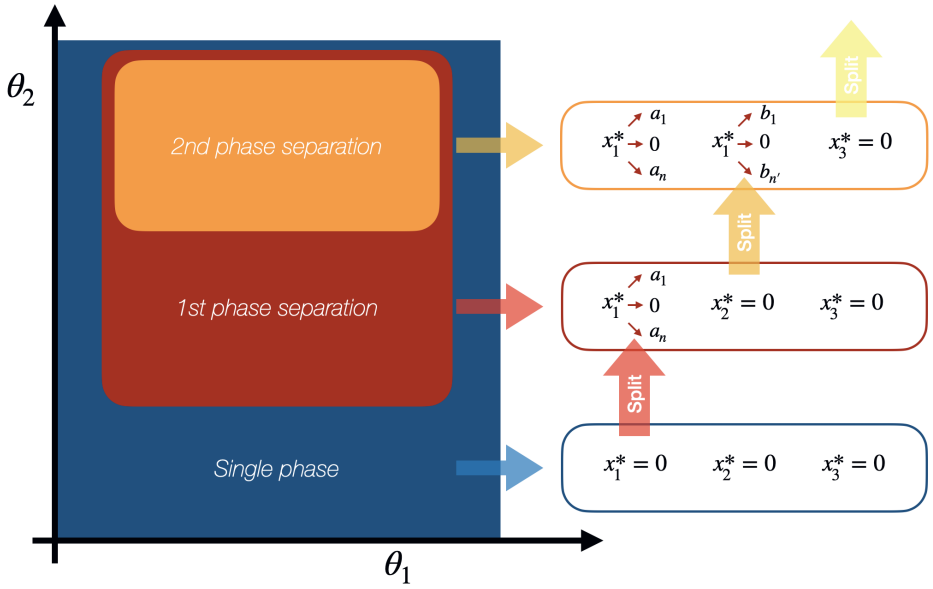


Figure 6.1: A schematic representation of the phase diagram of a II order RNM and the parameter x_k^* in the sparse case (large, negative ϵ_{ij} implying $x_k^* \simeq 0$). Firstly one has to look if the region of the parameter space whom is looking at may be of single phase or with phase separation. This can be done by actually computing the partition function but in general is done by looking at the multiplicity of solutions of the equation defined by putting Formula (6.28) *equal to zero.*

§6.3 II order interacting models

We will now show how equations (6.21), (6.36) and (6.39) can be used to actually solve some II order RNMs. We will start with a very simple, homogeneous one to use as a lamppost for the general behaviour and explore more complex models going on.

§6.3.1 RNM with Homogeneous interactions

Say that the Hamiltonian is in the form

$$H(\mathbf{A}) = -\theta_1 \langle \mathbf{A}|\epsilon\rangle - \frac{\theta_2}{2V} \langle \mathbf{A}|\beta|\mathbf{A}\rangle \text{ with } \beta = |1\rangle\langle 1| \tag{6.40}$$

so that every link interacts with any other link in the same way. The the eigenvalues of this matrix are all 0 except from 1 whose value is $\lambda = V$. The relative eigenvector is also easy to find as it is just $\frac{|1\rangle}{\sqrt{V}}$. Moreover we set ourselves in the case $\epsilon_a = \epsilon$, which can be further reabsorbed in the θ_1 parameter. The partition function then becomes easy to compute explicitly in the large N limit. By formula (6.21) we have

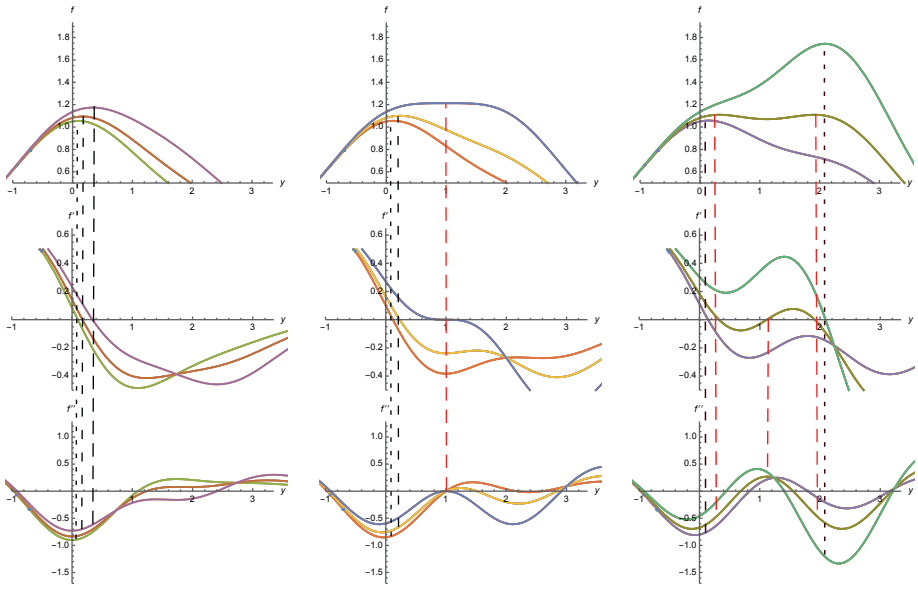


Figure 6.2: Plots of the function z^{eff} for different values of θ_1,θ_2. From top to bottom we plot z^{eff}, its first and second derivative to locate the maxima and show that there is a regime in which 2 maxima are present. The parameter θ_1 always takes the values $-3, -2.5, -2$, while θ_2 is, from left to right $3, 4, 5$ respectively.

$$z_a^{\text{eff}} = \lim_{V\to\infty}\left[\sqrt{\frac{1}{2\pi}}\int dx e^{-\frac{x^2}{2}}\left[1+e^{x\sqrt{\frac{\theta_2}{V}}+\theta_1}\right]^V\right]^{\frac{1}{V}} =$$

$$= \lim_{V\to\infty}\left[\sqrt{\frac{V}{2\pi}}\int dy\left[e^{-\frac{y^2}{2}}\left(1+e^{y\sqrt{\theta_2}+\theta_1}\right)\right]^V\right]^{\frac{1}{V}} =$$

$$= \max_{-\infty\le y\le\infty}\left\{e^{-\frac{y^2}{2}}(1+e^{y\sqrt{\theta_2}+\theta_1})\right\} \quad (6.41)$$

where the substitution $x = y\sqrt{V}$ has been made in the second row. The last step comes from a Laplace saddle point evaluation.

To evaluate the z^{eff} in the large limit we need to see where its derivative vanishes and its second derivative is negative, which can be done graphically and is shown in figure 6.2. We can see that for $\theta_2 \geq 4$ there is a range of θ_1 for which the partition function has more than one maximum. Indeed, in the parameter space we can locate a whole region in which the maxima are two and hence there are two possible values of the link probability. This is shown in figure 6.3.

Phase coexistence

In figure 6.4 we can see the effective link probability for two "slices" of the phase diagram, one across the single phase region (red, $\theta_1 = -4$) and another one across the single phase, then a bifurcation point, then again single phase (blue, $\theta_1 = -1$).

The key ingredients for the analysis of this simple, interacting model are now in place. The single link probability $p_{\text{eff}} = 1 - \frac{1}{z}$ is the probability by which links are assigned, yet we know that z can have two different values that will lead to two different phases. From figure 6.4 it is clear that when $\theta_2 >> 4$ and $\theta_2 < 4$ only one phase is present, the sparse one and the dense one respectively, while in the regime in which $\theta_2 \gtrsim 4$ the two phases coexist.

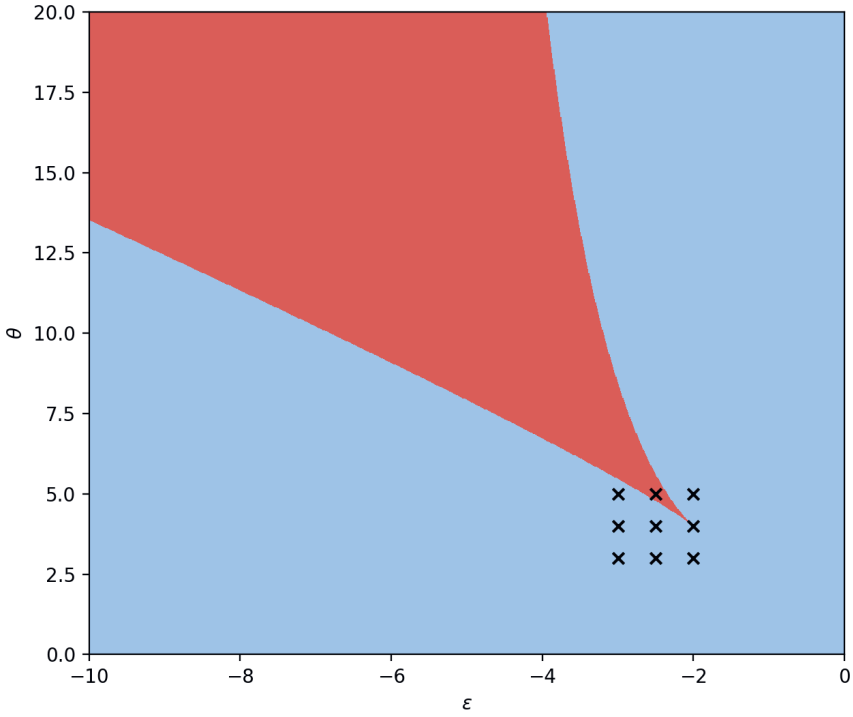


Figure 6.3: Phase diagram for the system. The blue region corresponds to the place where there is only one maximum, hence only one link probability. In the red region the maxima of equation (6.41) *are two and the link probabilities as well. The crosses indicate the coordinates of Fig. 6.2*

This behavior can be effectively described by the scheme in which, every time a realization of the networks has to be generated, a coin is tossed. This coin determines if the realization will be a dense or a sparse one, *i.e.* this first coin toss chooses the coin that will be tossed for each link in the realization. Then, when another realization has to be generated, this "phase coin" is tossed again and so on. For each instance of the network there will be one previous coin toss that will determine its phase.

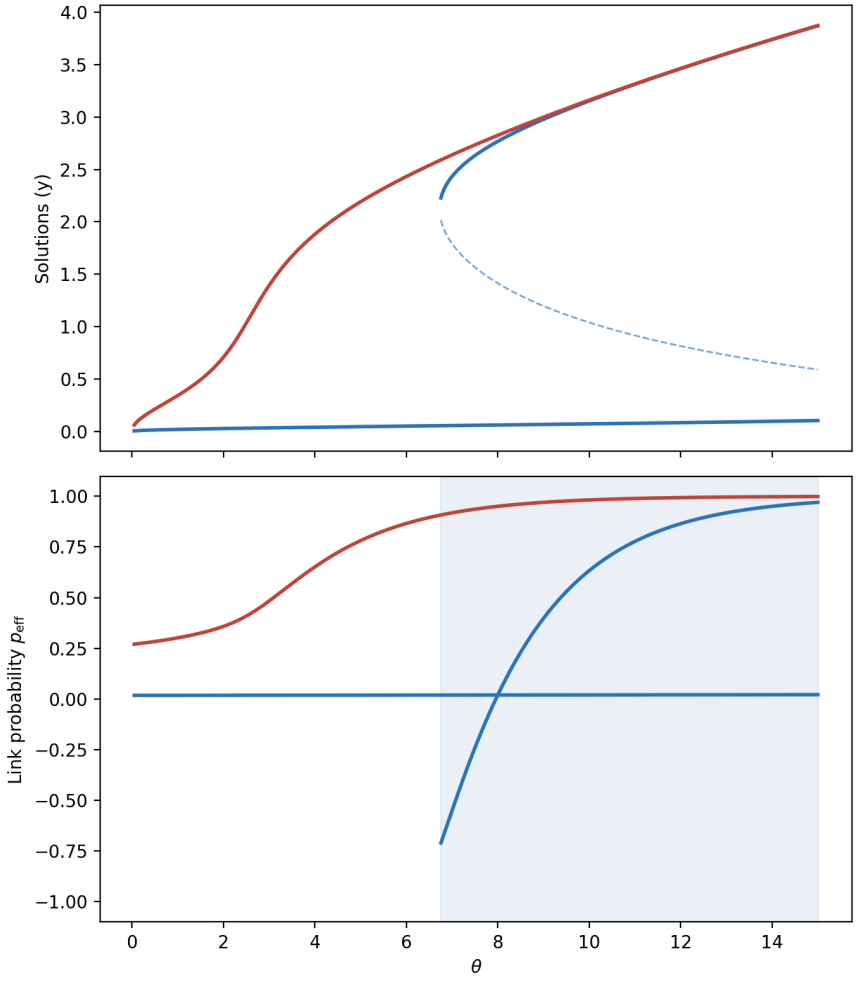


Figure 6.4: Possible solutions of equating to zero the derivative (top panel, dashed line indicate minima) of 6.41 and effective link probability (bottom panel, shaded region is where two solutions are possible) vs θ from 0 to 15 with $\epsilon = -4$ (blue) and $\epsilon = -1$ (red).

The relative abundance of the two phases (and hence the weights of the "phase coin") can be computed in the large V limit by plugging the respective values of p in the formula for the overall $P(\mathbf{A})$. For instance, for the dense phase in which p^{eff}

$$P \quad \left\{ \sum_{i<j} A_{ij} = V p_d \right\} \propto \binom{V}{V\frac{1+p_d}{2}} e^{\theta_1 V p_d + \frac{\theta_2}{2} V p_d^2} \tag{6.42}$$

$$\simeq \quad \exp\left\{ V \left[\theta_1 p_d + \frac{\theta_2}{2} p_d^2 + + \frac{1-p_d}{2} \ln \frac{2}{1-p_d} + \frac{1+p_d}{2} \ln \frac{2}{1+p_d} \right] \right\} \tag{6.43}$$

and for the sparse phase the same with p_d replaced with $p_s \equiv p_{\text{eff}}|_{x_1^*=0}$, and the exact proportionality constant is immediately computed by imposing

$$P \left\{ \sum_{i<j} A_{ij} = V p_d \right\} + P \left\{ \sum_{i \leq j} A_{ij} = V p_s \right\} = 1. \tag{6.44}$$

§6.3.2 A square-like interaction

After the previous simple model, that is given as a lamppost to highlight the main features of the systems at stake, we examine a slightly more complicated 1-rank model. We will show that this interaction structure give rise to an interesting phenomena that captures the power of single link probability representation found in equation (6.39). Let us consider the Hamiltonian with the general rank-one interaction matrix $\beta = |v\rangle \langle v|$

$$H(\mathbf{A}) = -\sum_{i<j} \epsilon_{ij} A_{ij} - \frac{C}{2V} \sum_{\substack{i<j \\ k<l}} v_{ij} v_{kl} A_{ij} A_{kl} \tag{6.45}$$

In this case the link probability would simply read

$$p_{ij} = \frac{e^{\epsilon_{ij} + x^* \sqrt{C} v_{ij}}}{1 + e^{\epsilon_{ij} + x^* \sqrt{C} v_{ij}}} \tag{6.46}$$

Heterogeneity as a low temperature phenomenon

Let us assume then that the model is even simpler and the local variable $\epsilon_{ij} = \epsilon$ is constant. Then the possible values of p_{ij} are of two kinds: a constant/homogeneous one in the high temperature/non effectively interacting phase, and an heterogeneous value for the phases in which $x^* \neq 0$.

In such a simple case already we see that the two phases describe two very different models: one is a standard homogeneous Erdos Renyi with probability given by $p = \ln \frac{e^\epsilon}{1+e^\epsilon}$, and the other in which heterogeneity comes into place. We can also construct the interaction in such a way that $v_{ij} = v_i + v_j$. Then this reads as an effective configuration model in the low temperature phase as

$$p_{ij} = \frac{K x_i x_j}{1 + K x_i x_j} \tag{6.47}$$

where $K = e^{\epsilon + x^* \sqrt{C}}$ and $x_i = e^{v_i}$, $x_j = e^{v_j}$. If we bring back the heterogeneity also in the ϵ_{ij} parameter as well, assuming the same separation $\epsilon_{ij} = \epsilon_i + \epsilon_j$ we will have a family of configuration models, described by the probability

$$p_{ij} = \frac{x_i x_j}{1 + x_i x_j} \tag{6.48}$$

where $x_i = e^{\epsilon_i + \sqrt{C} x^* v_i}$.

§6.3.3 Community Induced Interactions

We now extend the complexity of the simple models we have solved to include interacting link communities. Such communities, involving link instead of nodes, can arise for different reasons and link community detection has been already a quite active field [152, 153].

We will solve for a model in which link communities are induced by the nodes' ones, but the computations can be readily generalized to the desired case where communites are defined by some other property.

Among the most used (simple) models in network theory there are the Stochastic Block Model (SBM) and Core-Periphery one (CP). They both posit the partition of nodes into communities, 2 in the CP case or more for the SBM. We want to show here that such partitions can induce quite naturally a correlation structure between links so that already a quite complex behaviour arises. The process to construct the correlation structure is the following:

- for simplicity, let us assume there are two communities of nodes, with N_1 and N_2 nodes in each;
- links between nodes of community 1 has a certain probability $\frac{e^{\epsilon_1}}{1+e^{\epsilon_1}}$, links between nodes of community 2 has a probability $\frac{e^{\epsilon_2}}{1+e^{\epsilon_2}}$ and links between the two communities has probability $\frac{e^{\epsilon_3}}{1+e^{\epsilon_3}}$; creating a block matrix of the form

$$\epsilon = \begin{pmatrix} \epsilon_1 \mathbf{1}_{N_1 \times N_1} & \epsilon_3 \mathbf{1}_{N_1 \times N_2} \\ \epsilon_3 \mathbf{1}_{N_2 \times N_1} & \epsilon_2 \mathbf{1}_{N_2 \times N_2} \end{pmatrix} \tag{6.49}$$

  where $1_{N \times M}$ is the all-ones rectuangular $N \times M$ matrix.
- this community structure for the nodes induces a community structure for the links: the links between nodes of community 1, that are $V_1 = N_1(N_1 - 1)/2$, the links between nodes of community 2 that are $V_2 = N_2(N_2 - 1)/2$, and the links between nodes belonging to different communities that are $V_3 = N_1 N_2$;
- this will result in a 3×3 block structure of the interaction matrix, as if there was a SBM for three communities of links in the form

$$\beta = \begin{pmatrix} \beta_1 \mathbf{1}_{V_1 \times V_1} & \beta_{12} \mathbf{1}_{V_1 \times V_2} & \beta_{13} \mathbf{1}_{V_1 \times V_3} \\ \beta_{12} \mathbf{1}_{V_2 \times V_1} & \beta_2 \mathbf{1}_{V_2 \times V_2} & \beta_{23} \mathbf{1}_{V_2 \times V_3} \\ \beta_{13} \mathbf{1}_{V_3 \times V_1} & \beta_{23} \mathbf{1}_{V_3 \times V_2} & \beta_3 \mathbf{1}_{V_3 \times V_3} \end{pmatrix} \tag{6.50}$$

A scheme of how this community induced interaction is created is shown in figure 6.5.

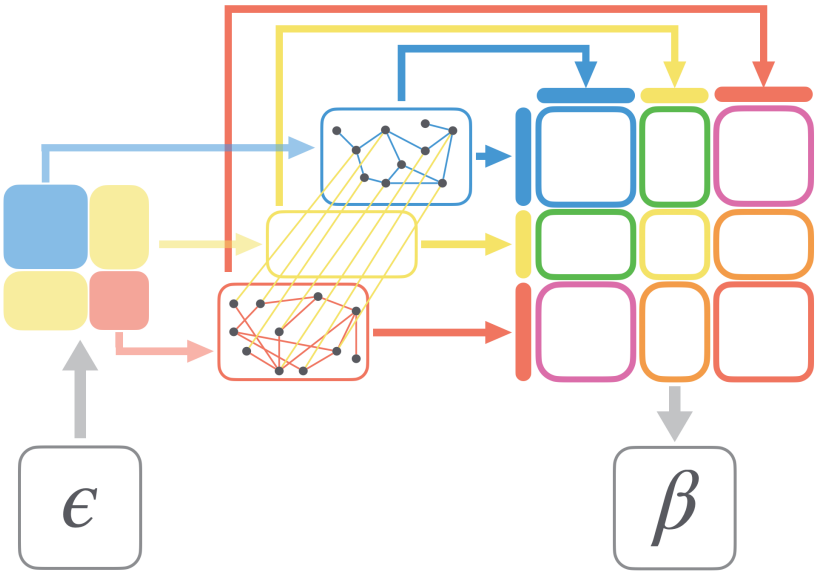


Figure 6.5: A pictorial representation of how the 3×3 correlation matrix is induced. The two nodes community structure allows for three values of link probability: in the first community (blue), in the second community (red), across the two communities (yellow). The possible interactions between these links are then blue/blue, blue/yellow (green), blue/red (purple), and so on.

This three by three matrix is the first example we can devise of phase space that has more than one region of phase coexistence, as indeed we will show that up to 27 phases can coexist. The rank three interaction defined by matrix β can be again decomposed with a basis of orthonormal vectors.

We now proceed to compute also the single link partition function for each one of the three kind of links that the community structure induces. The first big difference from the homogeneous interacting model is that there is a mixture of the eigenmodes and the equations that define the maxima are not the very same partition functions. This means that there are three distinct maximization procedure, each one returning a parameter that can have multiple values, and then the link-specific partition function is a function of all three parameters, hence possibly multifurcating in many different values.

Laplace large size approximation

The theorem we will make use of is the following [154]:

Let $g(x)$ and $f(x)$ be continuous and positive in the range $]\infty, \infty[$. Then

$$\lim_{N\to\infty}\left(\int dx g(x) f(x)^N\right)^{\frac{1}{N}} = \max_{-\infty\leq x\leq\infty}\{f(x)\} \tag{6.51}$$

The integrals we are interested in this work are of the kind

$$\frac{1}{\sqrt{2\pi}}\int dx\left[e^{-\frac{x^2}{2}}\prod_{a=1}^{V}\left(1+e^{x\sqrt{\frac{\lambda_k}{V}}(\omega_k)_a+\frac{\epsilon_a}{r}}\right)\right] = \sqrt{\frac{V}{2\pi}}\int dy\left[e^{-V\frac{y^2}{2}}\prod_{a=1}^{V}\left(1+e^{y\sqrt{\lambda_k}(\omega_k)_a+\frac{\epsilon_a}{r}}\right)\right] \tag{6.52}$$

where the substitution $x = y\sqrt{V}$ was made. We will assume for simplicity that all ϵ_a are the same[1]. Now we represent the eigenvector entries in the following way. Say that we are interested in the k^{th} eigenmode, we say that the corresponding eigenvector will be made of the following entries

$$(\omega_k)_1 \text{ with multiplicity } m_1 \tag{6.53}$$

$$(\omega_k)_2 \text{ with multiplicity } m_2 \tag{6.54}$$

$$(\omega_k)_3 \text{ with multiplicity } m_3 \tag{6.55}$$

and so on. Then we introduce the fraction of equal entries as $c_i = \frac{m_i}{V}$. Of course $\sum_{i=1}^{T} c_i = 1$.

Now we can write the product in the integral as

$$\prod_{a=1}^{V}\left(1+e^{y\sqrt{\lambda_k}(\omega_k)_a+\frac{\epsilon_a}{r}}\right) = \prod_{i=1}^{T}\left(1+e^{y\sqrt{\lambda_k}(\omega_k)_i+\frac{\epsilon}{r}}\right)^{c_i V} \tag{6.56}$$

We are now ready to compute the limit

[1] Or at least are coupled to the respective entry of the eigenvector as in the model described in 6.3.3. However a generalization to the case in which they also have their respective multiplicity and heterogeneity can be done.

$$\lim_{V\to\infty}\left[\sqrt{\frac{V}{2\pi}}\int dy\left[e^{-\frac{y^2}{2}}\prod_{i=1}^{T}\left(1+e^{y\sqrt{\lambda_k}(\omega_k)_i+\frac{\epsilon}{r}}\right)^{c_i}\right]^V\right]^{\frac{1}{V}}$$
$$=\underbrace{\lim_{V\to\infty}\left[\sqrt{\frac{V}{2\pi}}\right]^{\frac{1}{V}}}_{=1}\lim_{V\to\infty}\left[\int dy\left[e^{-\frac{y^2}{2}}\prod_{i=1}^{T}\left(1+e^{y\sqrt{\lambda_k}(\omega_k)_i+\frac{\epsilon}{r}}\right)^{c_i}\right]^V\right]^{\frac{1}{V}}$$
$$=\max_{-\infty\leq x\leq\infty}\left\{e^{-\frac{y^2}{2}}\prod_{i=1}^{T}\left(1+e^{y\sqrt{\lambda_k}(\omega_k)_i+\frac{\epsilon}{r}}\right)^{c_i}\right\} \quad (6.57)$$

from which the exact computations performed for the various models follow. In particular we want to highlight the fact that if, there are entries whose multiplicities are extensive, *i.e.* grow as a fraction of V, and other sets with subextensive multiplicities, then the only surviving part is that of the extensive entries, as they clearly dominate the product. In a more formal way, let us have M different values of the entries, and let the large V structure be $c_1(V)\to_{V\to\infty} C_1>0, c_2(V)\to_{V\to\infty} C_2>0,\ldots,c_M(V)\to_{V\to\infty} C_M>0, c_{M+1}(V)\to_{V\to\infty} C_{M+1}=0, c_{M+2}(V)\to_{V\to\infty} C_{M+2}=0,\ldots c_T(V)\to_{V\to\infty} C_T=0$. Then

$$\lim_{V\to\infty}\left[\int dy\left[e^{-\frac{y^2}{2}}\prod_{i=1}^{T}\left(1+e^{y\sqrt{\lambda_k}(\omega_k)_i+\frac{\epsilon}{r}}\right)^{c_i}\right]^V\right]^{\frac{1}{V}}$$
$$=\max_{-\infty\leq x\leq\infty}\left\{e^{-\frac{y^2}{2}}\prod_{i=1}^{M}\left(1+e^{y\sqrt{\lambda_k}(\omega_k)_i+\frac{\epsilon}{r}}\right)^{c_i}\right\} \quad (6.58)$$

A basis for the community interaction matrix

We can write a rank three matrix in the following form

$$\beta=\lambda_1|\omega_1\rangle\langle\omega_1|+\lambda_2|\omega_2\rangle\langle\omega_2|+\lambda_3|\omega_3\rangle\langle\omega_3| \quad (6.59)$$

where the eigenvectors are

$$\langle\omega_1|=\frac{1}{\sqrt{n_1}}(\underbrace{(\omega_1)_1,\ldots,(\omega_1)_1}_{V_1},\underbrace{(\omega_1)_2,\ldots,(\omega_1)_2}_{V_2},\underbrace{(\omega_1)_3,\ldots,(\omega_1)_3}_{V_3}) \quad (6.60)$$

$$\langle\omega_2|=\frac{1}{\sqrt{n_2}}(\underbrace{(\omega_2)_1,\ldots,(\omega_2)_1}_{V_1},\underbrace{(\omega_2)_2,\ldots,(\omega_2)_2}_{V_2},\underbrace{(\omega_2)_3,\ldots,(\omega_2)_3}_{V_3}) \quad (6.61)$$

$$\langle\omega_3| = \frac{1}{\sqrt{n_3}}(\underbrace{(\omega_3)_1,\ldots,(\omega_3)_1}_{V_1},\underbrace{(\omega_3)_2,\ldots,(\omega_3)_2}_{V_2},\underbrace{(\omega_3)_3,\ldots,(\omega_3)_3}_{V_3}). \tag{6.62}$$

where the norms are

$$n_1 = V_1\,(\omega_1)_1^2 + V_2\,(\omega_1)_2^2 + V_3\,(\omega_1)_3^2 \tag{6.63}$$

$$n_2 = V_1\,(\omega_2)_1^2 + V_2\,(\omega_2)_2^2 + V_3\,(\omega_2)_3^2 \tag{6.64}$$

$$n_3 = V_1\,(\omega_3)_1^2 + V_2\,(\omega_3)_2^2 + V_3\,(\omega_3)_3^2 \tag{6.65}$$

The eigenvalues can be computed from the secular equation

$$\beta\,|\omega_i\rangle = \lambda_i\,|\omega_i\rangle \tag{6.66}$$

so that

$$\begin{cases} \lambda_1 = V_1\beta_1 + V_2\beta_{12}\frac{(\omega_1)_2}{(\omega_1)_1} + V_3\beta_{13}\frac{(\omega_1)_3}{(\omega_1)_3} \\ \lambda_2 = V_1\beta_1 + V_2\beta_{12}\frac{(\omega_2)_2}{(\omega_2)_1} + V_3\beta_{13}\frac{(\omega_2)_3}{(\omega_2)_1} \\ \lambda_3 = V_1\beta_1 + V_2\beta_{12}\frac{(\omega_3)_2}{(\omega_3)_1} + V_3\beta_{13}\frac{(\omega_3)_3}{(\omega_3)_1} \end{cases} \tag{6.67}$$

The equation for the spectral decomposition defines the relations between the eigenvector components and the β parameters

$$\begin{cases} \beta_1 = \lambda_1\,(\omega_1)_1^2 + \lambda_2\,(\omega_2)_1^2 + \lambda_3\,(\omega_3)_1^2 \\ \beta_2 = \lambda_1\,(\omega_1)_2^2 + \lambda_2\,(\omega_2)_2^2 + \lambda_3\,(\omega_3)_2^2 \\ \beta_3 = \lambda_1\,(\omega_1)_3^2 + \lambda_2\,(\omega_2)_3^2 + \lambda_3\,(\omega_3)_3^2 \\ \beta_{12} = \lambda_1\,(\omega_1)_1\,(\omega_1)_2 + \lambda_2\,(\omega_2)_1\,(\omega_2)_2 + \lambda_3\,(\omega_3)_1\,(\omega_3)_2 \\ \beta_{13} = \lambda_1\,(\omega_1)_1\,(\omega_1)_3 + \lambda_2\,(\omega_2)_1\,(\omega_2)_3 + \lambda_3\,(\omega_3)_1\,(\omega_3)_3 \\ \beta_{23} = \lambda_1\,(\omega_1)_2\,(\omega_1)_3 + \lambda_2\,(\omega_2)_2\,(\omega_2)_3 + \lambda_3\,(\omega_3)_2\,(\omega_3)_3 \end{cases} \tag{6.68}$$

and these are supplemented by the orthogonality equations so that

$$\begin{cases} V_1\,(\omega_1)_1\,(\omega_2)_1 + V_2\,(\omega_1)_2\,(\omega_2)_2 + V_3\,(\omega_1)_3\,(\omega_2)_3 = 0 \\ V_1\,(\omega_1)_1\,(\omega_3)_1 + V_2\,(\omega_1)_2\,(\omega_3)_2 + V_3\,(\omega_1)_3\,(\omega_3)_3 = 0 \\ V_1\,(\omega_2)_1\,(\omega_3)_1 + V_2\,(\omega_2)_2\,(\omega_3)_2 + V_3\,(\omega_2)_3\,(\omega_3)_3 = 0 \end{cases} \tag{6.69}$$

which completely defines the ω and the eigenvalues. To compute the partition function we also define the quantities $\alpha = \frac{V_1}{V}$, $\beta = \frac{V_2}{V}$ and $\gamma = \frac{V_3}{V}$. The partition function to be computed will be

$$Z = \tag{6.70}$$

$$\underbrace{\frac{1}{\sqrt{2\pi}}\int dx e^{-\frac{x^2}{2}}\left[1+e^{x\sqrt{\frac{\lambda_1}{Vn_1}}(\omega_1)_1+\frac{\epsilon_1}{3}}\right]^{\alpha V}\left[1+e^{x\sqrt{\frac{\lambda_1}{Vn_1}}(\omega_1)_2+\frac{\epsilon_2}{3}}\right]^{\beta V}\left[1+e^{x\sqrt{\frac{\lambda_1}{Vn_1}}(\omega_1)_3+\frac{\epsilon_3}{3}}\right]^{\gamma V}}_{I_1}. \tag{6.71}$$

$$\underbrace{\frac{1}{\sqrt{2\pi}}\int dx e^{-\frac{x^2}{2}}\left[1+e^{x\sqrt{\frac{\lambda_2}{Vn_2}}(\omega_2)_1+\frac{\epsilon_1}{3}}\right]^{\alpha V}\left[1+e^{x\sqrt{\frac{\lambda_2}{Vn_2}}(\omega_2)_2+\frac{\epsilon_2}{3}}\right]^{\beta V}\left[1+e^{x\sqrt{\frac{\lambda_2}{Vn_2}}(\omega_2)_3+\frac{\epsilon_3}{3}}\right]^{\gamma V}}_{I_2}. \tag{6.72}$$

$$\underbrace{\frac{1}{\sqrt{2\pi}}\int dx e^{-\frac{x^2}{2}}\left[1+e^{x\sqrt{\frac{\lambda_3}{Vn_3}}(\omega_3)_1+\frac{\epsilon_1}{3}}\right]^{\alpha V}\left[1+e^{x\sqrt{\frac{\lambda_3}{Vn_3}}(\omega_3)_2+\frac{\epsilon_2}{3}}\right]^{\beta V}\left[1+e^{x\sqrt{\frac{\lambda_3}{Vn_3}}(\omega_3)_3+\frac{\epsilon_3}{3}}\right]^{\gamma V}}_{I_3} \tag{6.73}$$

$$= (z_1)^{\alpha V}(z_2)^{\beta V}(z_3)^{\gamma V} \tag{6.74}$$

In the large V we have to compute

$$\lim_{V\to\infty}(Z)^{\frac{1}{V}} = (z_1)^{\alpha}(z_2)^{\beta}(z_3)^{\gamma} = \lim_{V\to\infty}(I_1)^{\frac{1}{V}}\lim_{V\to\infty}(I_2)^{\frac{1}{V}}\lim_{V\to\infty}(I_3)^{\frac{1}{V}} \tag{6.75}$$

that gives

$$\lim_{V\to\infty}(I_1)^{\frac{1}{V}}$$
$$= \max_{y_1}\left\{e^{-\frac{y_1^2}{2}}\left(1+e^{y_1\sqrt{\frac{\lambda_1}{n_1}}(\omega_1)_1+\frac{\epsilon_1}{3}}\right)^{\alpha}\left(1+e^{y_1\sqrt{\frac{\lambda_1}{n_1}}(\omega_1)_2+\frac{\epsilon_2}{3}}\right)^{\beta}\left(1+e^{y_1\sqrt{\frac{\lambda_1}{n_1}}(\omega_1)_3+\frac{\epsilon_3}{3}}\right)^{\gamma}\right\}$$
$$\equiv \max f(y_1) \tag{6.76}$$

$$\lim_{V\to\infty}(I_2)^{\frac{1}{V}}$$
$$= \max_{y_2}\left\{e^{-\frac{y_2^2}{2}}\left(1+e^{y_2\sqrt{\frac{\lambda_2}{n_2}}(\omega_2)_1+\frac{\epsilon_1}{3}}\right)^{\alpha}\left(1+e^{y_2\sqrt{\frac{\lambda_2}{n_2}}(\omega_2)_2+\frac{\epsilon_2}{3}}\right)^{\beta}\left(1+e^{y_2\sqrt{\frac{\lambda_2}{n_2}}(\omega_2)_3+\frac{\epsilon_3}{3}}\right)^{\gamma}\right\}$$
$$\equiv \max g(y_2) \tag{6.77}$$

$$\lim_{V\to\infty}(I_1)^{\frac{1}{V}}$$
$$= \max_{y_3}\left\{e^{-\frac{y_3^2}{2}}\left(1+e^{y_3\sqrt{\frac{\lambda_3}{n_3}}(\omega_3)_1+\frac{\epsilon_1}{3}}\right)^{\alpha}\left(1+e^{y_3\sqrt{\frac{\lambda_3}{n_3}}(\omega_3)_2+\frac{\epsilon_2}{3}}\right)^{\beta}\left(1+e^{y_3\sqrt{\frac{\lambda_3}{n_3}}(\omega_3)_3+\frac{\epsilon_3}{3}}\right)^{\gamma}\right\}$$
$$\equiv \max h(y_3) \tag{6.78}$$

we now introduce the notation

$$y_1^* = \arg\max f(y_1) \tag{6.79}$$

$$y_2^* = \arg\max g(y_2) \tag{6.80}$$

$$y_3^* = \arg\max h(y_3) \tag{6.81}$$

$$\begin{aligned}
&(z_1)^\alpha (z_2)^\beta (z_3)^\gamma \\
&= \left[e^{-\frac{(y_1^*)^2}{2}} \left(1 + e^{y_1^* \sqrt{\frac{\lambda_1}{n_1}} (\omega_1)_1 + \frac{\epsilon_1}{3}}\right)^\alpha \left(1 + e^{y_1^* \sqrt{\frac{\lambda_1}{n_1}} (\omega_1)_2 + \frac{\epsilon_2}{3}}\right)^\beta \left(1 + e^{y_1^* \sqrt{\frac{\lambda_1}{n_1}} (\omega_1)_3 + \frac{\epsilon_3}{3}}\right)^\gamma\right] \\
&\left[e^{-\frac{(y_2^*)^2}{2}} \left(1 + e^{y_2^* \sqrt{\frac{\lambda_2}{n_2}} (\omega_2)_1 + \frac{\epsilon_1}{3}}\right)^\alpha \left(1 + e^{y_2^* \sqrt{\frac{\lambda_2}{n_2}} (\omega_2)_2 + \frac{\epsilon_2}{3}}\right)^\beta \left(1 + e^{y_2^* \sqrt{\frac{\lambda_2}{n_2}} (\omega_2)_3 + \frac{\epsilon_3}{3}}\right)^\gamma\right] \\
&\left[e^{-\frac{(y_3^*)^2}{2}} \left(1 + e^{y_3^* \sqrt{\frac{\lambda_3}{n_3}} (\omega_3)_1 + \frac{\epsilon_1}{3}}\right)^\alpha \left(1 + e^{y_3^* \sqrt{\frac{\lambda_3}{n_3}} (\omega_3)_2 + \frac{\epsilon_2}{3}}\right)^\beta \left(1 + e^{y_3^* \sqrt{\frac{\lambda_3}{n_3}} (\omega_3)_3 + \frac{\epsilon_3}{3}}\right)^\gamma\right]
\end{aligned} \tag{6.82}$$

by which we can recognize the terms

$$z_1 = \left[e^{-\frac{(y_1^*)^2}{2\alpha}} \left(1 + e^{y_1^* \sqrt{\frac{\lambda_1}{n_1}} (\omega_1)_1 + \frac{\epsilon_1}{3}}\right) \left(1 + e^{y_2^* \sqrt{\frac{\lambda_2}{n_2}} (\omega_2)_1 + \frac{\epsilon_1}{3}}\right) \left(1 + e^{y_3^* \sqrt{\frac{\lambda_3}{n_3}} (\omega_3)_1 + \frac{\epsilon_1}{3}}\right)\right] \tag{6.83}$$

$$z_2 = \left[e^{-\frac{(y_2^*)^2}{2\beta}} \left(1 + e^{y_1^* \sqrt{\frac{\lambda_1}{n_1}} (\omega_1)_2 + \frac{\epsilon_2}{3}}\right) \left(1 + e^{y_2^* \sqrt{\frac{\lambda_2}{n_2}} (\omega_2)_2 + \frac{\epsilon_2}{3}}\right) \left(1 + e^{y_3^* \sqrt{\frac{\lambda_3}{n_3}} (\omega_3)_2 + \frac{\epsilon_2}{3}}\right)\right] \tag{6.84}$$

$$z_3 = \left[e^{-\frac{(y_3^*)^2}{2\gamma}} \left(1 + e^{y_1^* \sqrt{\frac{\lambda_1}{n_1}} (\omega_1)_3 + \frac{\epsilon_3}{3}}\right) \left(1 + e^{y_2^* \sqrt{\frac{\lambda_2}{n_2}} (\omega_2)_3 + \frac{\epsilon_3}{3}}\right) \left(1 + e^{y_3^* \sqrt{\frac{\lambda_3}{n_3}} (\omega_3)_3 + \frac{\epsilon_3}{3}}\right)\right]. \tag{6.85}$$

In figure 6.6 we show an instance of interaction matrix value for which there is a complete separation of phases. Each one of the modes of the interaction give rise to a three-partite split in the corresponding part of the partition function.

§6.3.4 Node mediated interaction

Up until now we considered models in which some kind of community structure was induced between links, and this fact was the responsible for the phase transition. Now we consider a new kind of interaction, in which links are correlated via nodes' property, namely links are correlated if they share one node in common. This structure is highly

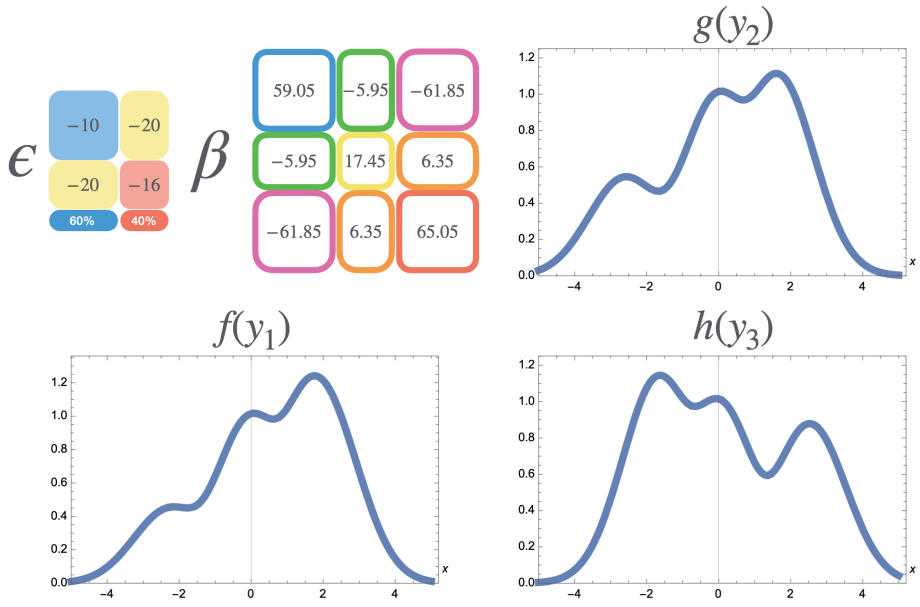


Figure 6.6: Instance of a crowded phase space: each one of the functions defined in (6.76), (6.77), (6.78) *has multiple maxima, hence the each link probability has 27 possible values.*

non-trivial and has been considered in the literature of ERGs as the "Two star" model [49]. In the following we will solve it in a different fashion, making us of some results coming from coding theory. A scheme of how this interaction works is depicted in figure 6.7.

The Johnson graph

Again we assume that all ϵ are the same for simplicity, but note that once one has diagonalized the matrix of the correlations, formula (6.39) allows to write the solution of a heterogeneous degree two-star model. Let us concentrate on the matrix then. The interaction we want to implement is such that links are correlated if they share one node. This kind of interaction can mimic for instance traffic jam propagation. We can define β as follows

$$\beta_{ijkl} = \begin{cases} 1 \text{ if } |(i,j) \cap (k,l)| = 1 \\ 0 \text{ otherwise.} \end{cases} \tag{6.86}$$

and this matrix is actually a known object in Algebraic Coding Theory known as the Johnson Graph $J(N,2)$ adjacency matrix. The Johnson Graph $J(N,k)$ is a graph in which the nodes are the subsets of $\{1,2,\ldots,N\}$ elements of dimension k. Edges between these nodes exist if the intersection of the two nodes has dimension $k-1$. If $k=2$ the adjacency matrix of $J(N,2)$ reduces exactly to the β we are looking for.

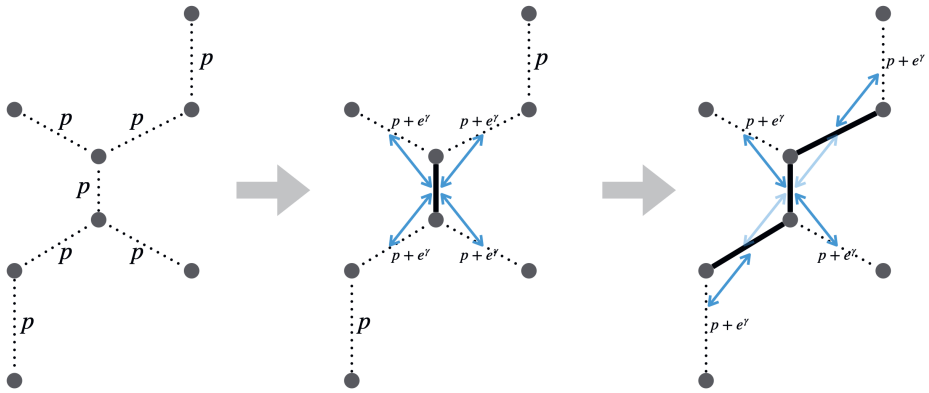


Figure 6.7: A scheme of how the interaction is mediated by nodes. At first, any link has the same (low) probability p to exist. If one flips up and connects two nodes, then it start interacting with the sites of the links that share one node with that link. Then more links might be promoted and propagate the interaction with their node-neighbours.

Moreover if one wants to treat this model as the Two-Star ERG, can see that the Hamiltonian can be readily parametrized and put in the more familiar form

$$\begin{aligned} H(\mathbf{A}) = -\frac{\theta_1 - \theta_2}{2} \sum_{i<j} A_{ij} - \frac{\theta_2}{2} \sum_{i<j} \sum_{k<l} \beta_{ijkl} A_{ij} A_{kl} = \\ = -\frac{\theta_1 - \theta_2}{2} \sum_{i<j} A_{ij} - \frac{\theta_2}{2} \sum_{i<j<k} A_{ij} A_{jk} = \\ = -\theta_1 m(\mathbf{A}) - \theta_2 s(\mathbf{A}) \end{aligned} \tag{6.87}$$

From literature it is known that the Johnson Graph Adjacency matrix is full rank, and has known eigenspaces [155, 156, 157]. We choose to reabsorb the parameters so that $\frac{\theta_1 - \theta_2}{2} = \epsilon$ and $\theta_2 = \gamma$. For our case of interest $k = 2$, the eigenvalues of β are

$$\begin{cases} \lambda = 2N - 4 \text{ with multiplicity } \mu_1 = 1; \\ \mu = N - 4 \text{ with multiplicity } \mu_2 = N - 1; \\ \nu = -2 \text{ with multiplicity } \mu_3 = \frac{(N-2)(N-3)}{2}; \end{cases} \tag{6.88}$$

Before showing how the Partition function decouples in the eigenmodes, we must pay attention to the fact that the scaling of the largest eigenvalue is of the order $N \simeq \sqrt{V}$. This means that while in equation (5.2) the term to keep confined the interaction term was generically put to V, in this case that would kill the interaction

entirely. As a matter of fact, we have to rescale by a quantity of the same order of the largest eigenvalue of the interaction, *i.e.* in this case $\frac{1}{N}$.

Then partition function decouples as

$$Z = \underbrace{\left[\sqrt{\frac{1}{2\pi}}\int dx e^{-\frac{x^2}{2}} \prod_{a=1}^{V}\left[1+e^{x\sqrt{\frac{\gamma\lambda}{N}}(\omega_k)_a+\frac{\epsilon}{r}}\right]\right]}_{I_1}\cdot$$
$$\cdot \underbrace{\prod_{k=2}^{N}\left[\sqrt{\frac{1}{2\pi}}\int dx e^{-\frac{x^2}{2}} \prod_{a=1}^{V}\left[1+e^{x\sqrt{\frac{\gamma\mu}{N}}(\omega_k)_a+\frac{\epsilon}{r}}\right]\right]}_{I_2}\cdot$$
$$\cdot \underbrace{\prod_{k=N+1}^{V}\left[\sqrt{\frac{1}{2\pi}}\int dx e^{-\frac{x^2}{2}} \prod_{a=1}^{V}\left[1+e^{x\sqrt{\frac{\gamma\nu}{N}}(\omega_k)_a+\frac{\epsilon}{r}}\right]\right]}_{I_3}. \quad (6.89)$$

The large N limit

We now want compute the terms of the partition function in the large N, V limit. The single-link partition function will be

$$z = \lim_{V\to\infty}(Z)^{\frac{1}{V}} = \lim_{V\to\infty}(I_1)^{\frac{1}{V}}\lim_{V\to\infty}(I_2)^{\frac{1}{V}}\lim_{V\to\infty}(I_3)^{\frac{1}{V}} \quad (6.90)$$

We have already seen the first term in previous cases, that is

$$\lim_{V\to\infty}(I_1)^{\frac{1}{V}} = \max_y\left\{e^{-\frac{y^2}{2}}\left(1+e^{y\sqrt{2\gamma}+\frac{\epsilon}{r}}\right)\right\} = \sum_{A_{ij}=0,1} e^{-\frac{y_*^2}{2}}\left(e^{\left[y_*\sqrt{2\gamma}+\frac{\epsilon}{r}\right]A_{ij}}\right) \quad (6.91)$$

where $y_* = \arg\max\left\{e^{-\frac{y^2}{2}}\left(1+e^{y\sqrt{2\gamma}+\frac{\epsilon}{r}}\right)\right\}$.

We see there is a phase transition if $\gamma > 2$ (and the proper ϵ is there).

For the third term we see that the scaling term N kills any effect coming from this eigenspace. As a matter of facts we can see that, when N is large, we can approximate

$$\left[\sqrt{\frac{1}{2\pi}}\int dx e^{-\frac{x^2}{2}} \prod_{a=1}^{V}\left[1+e^{x\sqrt{-\frac{2\gamma}{N}}(\omega_k)_a+\frac{\epsilon}{r}}\right]\right] \simeq$$
$$\simeq \max_{y_3}\left\{e^{-\frac{y^2}{2}}\left(1+e^{\frac{\epsilon}{r}}\right)\right\} = \left[1+e^{\frac{\epsilon}{r}}\right] \quad (6.92)$$

This term is a constant and does not depend on any parameter of the model, hence it does not give rise to any separation of phases. If one is interested in the correction that arise in the finite size case however, a orthogonal basis of the third eigenspace is needed, and we will briefly provide a method to do so.

Last we need to compute the second term I_2. For this term we need to be careful. As a matter of facts, the representation for which we are able to decompose the interaction matrix requires an orthonormal basis of eigenvectors to be performed. While in the all-to-all and link community we did not need to pay too much attention to this, since eigenvectors of different eigenvalues are orthogonal between them, in this case we need again to ensure that we find an orthonormal basis for the eigenspace of eigenvector μ.

The eigenvalue equation in this case becomes

$$\sum_{(k,l)\neq(i,j)} \beta_{(ij),(kl)} v^a_{kl} = \sum_{l\neq i,j} \left(v^a_{il} + v^a_{lj}\right) = \mu v^a_{ij} \tag{6.93}$$

We now introduce the following vector for $a = 2, 3, \cdots, N$

$$v^a_{ij} = \begin{cases} 0 \text{ if } (i,j) = (a, a+1) \\ 1 \text{ if } i = a \wedge j = a \\ -1 \text{ if } i = a+1 \wedge j = a+1 \\ 0 \text{ otherwise} \end{cases} \tag{6.94}$$

and can be checked that they are indeed eigenvectors and a basis. However they are not orthogonal, which can be seen by computing the scalar product between two of them

$$\langle v^a | v^b \rangle = 2(N-2)\delta_{ab} - (N-3)\delta_{ab\pm 1} \tag{6.95}$$

Orthogonalization of the basis

From a set of nonorthogonal basis vectors we can find a set of orthogonal ones by simultaneously orthogonalizing them according to the Lowdin transformations [158]. If we define the overlap matrix $\mathbf{S}$

$$S_{ab} = \langle v^a | v^b \rangle \tag{6.96}$$

we can get a set of orthogonal vectors as

$$|\omega^a\rangle = \sum_b S^{-\frac{1}{2}}_{ab} |v^b\rangle \tag{6.97}$$

Let us focus on the overlap matrix. Since it is symmetric, we can decompose it as

$$\mathbf{S} = \mathbf{U}\Lambda\mathbf{U}^T \tag{6.98}$$

and

$$\mathbf{S}^{-\frac{1}{2}} = \mathbf{U}\Lambda^{-\frac{1}{2}}\mathbf{U}^T. \tag{6.99}$$

This is a banded tridiagonal matrix. It is known that it can be diagonalized as

$$\Lambda_{cd} = \frac{\delta_{cd}}{\sqrt{2(N-2) - 2(N-3)\cos\left(\frac{c\pi}{N}\right)}} \tag{6.100}$$

$$U_{ac} = \sin\left(\frac{ac\pi}{N}\right) \tag{6.101}$$

so that we can compute

$$S_{ab}^{-\frac{1}{2}} = \sum_{cd} \sin\left(\frac{ac\pi}{N}\right)\sin\left(\frac{db\pi}{N}\right)\frac{\delta_{cd}}{\sqrt{2(N-2) - 2(N-3)\cos\left(\frac{c\pi}{N}\right)}} = \sum_{c} \frac{\sin\left(\frac{ac\pi}{N}\right)\sin\left(\frac{cb\pi}{N}\right)}{\sqrt{2(N-2) - 2(N-3)\cos\left(\frac{c\pi}{N}\right)}} \tag{6.102}$$

So that the orthogonalized vectors can be written as

$$|\omega^a\rangle = \sum_{a,c}^{N} \frac{\sin\left(\frac{ac\pi}{N}\right)\sin\left(\frac{cb\pi}{N}\right)}{\sqrt{2(N-2) - 2(N-3)\cos\left(\frac{c\pi}{N}\right)}} |v^b\rangle \tag{6.103}$$

One can realize that the structure of this vector is such that only $2N-2$ entries are nonzero. Hence we can apply the consideration at the end of section 6.3.3 to claim that we do not expect a phase separation coming from this term. If one is interested in the finite size system effects, an orthogonal basis for the third space is needed as well. In this case we refer to the work [156] that has a complete description of all the eigenvectors. Nonetheless they are not orthogonal but can be made so by following the same procedure of Lowdin transformations we just performed. This means that the second term can be evaluated as

$$\left[\sqrt{\frac{1}{2\pi}}\int dx e^{-\frac{x^2}{2}} \prod_{a=1}^{V}\left[1 + e^{x\sqrt{\frac{\gamma\mu}{N}}(\omega_k)_a + \frac{\epsilon}{r}}\right]\right] \simeq \simeq \max_y \left\{e^{-\frac{y^2}{2}}\left(1 + e^{\frac{\epsilon}{r}}\right)\right\} = \left[1 + e^{\frac{\epsilon}{r}}\right] \tag{6.104}$$

Again, this does not produce any phase separation.

Then the term resulting from the second and third term can then be evaluated as

$$\lim_{V\to\infty}(I_2)^{\frac{1}{V}} \lim_{V\to\infty}(I_3)^{\frac{1}{V}} \simeq \left[1 + e^{\frac{\epsilon}{r}}\right]^{\mu_2+\mu_3} = \sum_{A_{ij}=0,1} e^{\left[\frac{\mu_2+\mu_3}{r}\epsilon\right]A_{ij}}. \tag{6.105}$$

Bringing back all terms into a single one, and making p the subject of the formula, we get

$$p = 1 - \frac{1}{e^{-\frac{(y^*)^2}{2}}\left(1 + e^{y^*\sqrt{2\gamma}+\epsilon}\right)} \tag{6.106}$$

In figures 6.8 and 6.9 we show the Phase Diagram and the link probability of the model for a certain range of parameters, which both resemble the simple, homogeneous case.

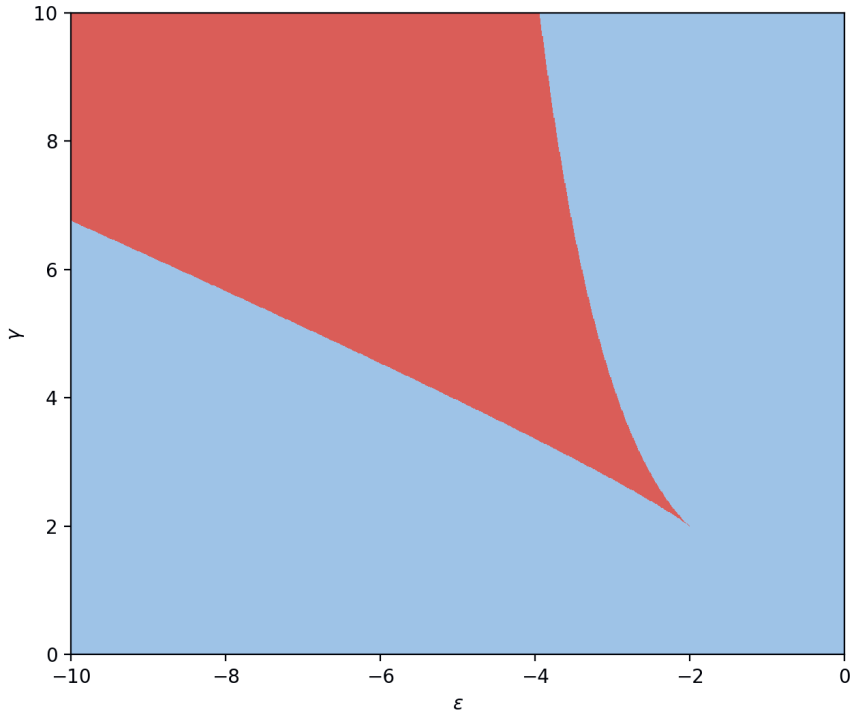


Figure 6.8: Phase Diagram of the Node-mediated Interaction model. In red the region where multiple phases are present, in blue the region where only one phase is possible.

Lastly, we want to comment on the unexpected similarity between this model and the simple, homogeneous one. This is due to the fact that while the mean field approximation is far from exact, since a large number of eigenvalues of the same size as the largest one exist, nonetheless they localize in a few nodes, and the respective eigenvectors are basically everywhere zero. This is not true if we want to consider the finite size effects on the system. In that case one would need to resort to the general single-link probability formula (6.39) with the eigenvector entries that are found using the method of appendix 6.3.4. The phase structure would have to be found with the equation described in appendix 6.2.2, by putting equation (6.25) equal to zero with the found entries. After that, it could also be possible to correct the model for an heterogeneous ϵ term.

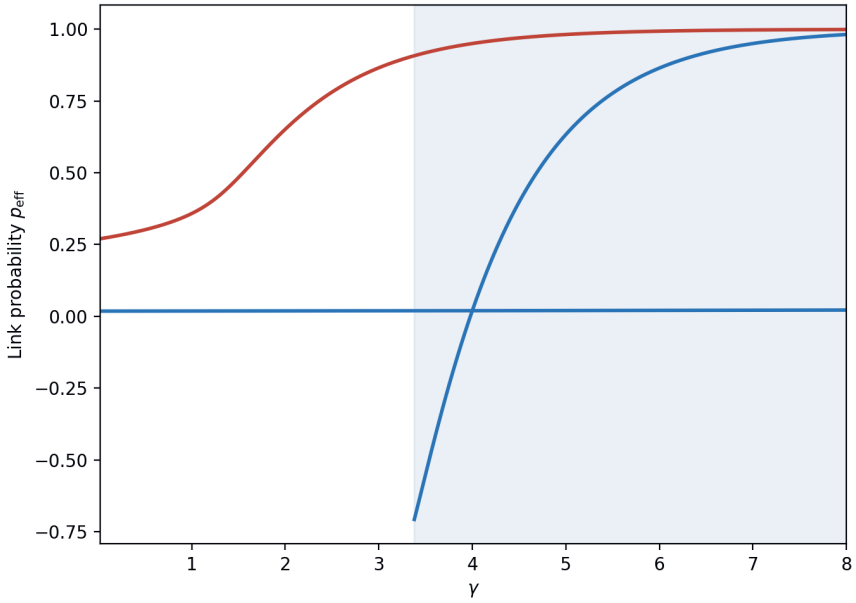


Figure 6.9: Single-link probability function, in the range $\gamma = [0, 15]$, with a large negative ϵ in blue, shaded region indicate where two solutions are possible.

§6.4 On the Mean Field Approximation and Phase Structure

Before the final remarks, we want to comment on the unexpected similarity between the node-mediated interaction model 6.3.4 and the simple, homogeneous one 6.3.1. One would expect that with such a non-trivial corrrelation structure, the mean field computation, which is essentialy the model of Section 6.3.1 with the right scalings for the parameters, would be a wrong approximation of the system. However that is not the case, and they basically coincide in the large N limit. Explicitly, the correspondence follows from Equation (6.91): the node-mediated model's mean-field mode carries effective coupling $\theta_{\text{eff}} = 2\gamma$, so that the homogeneous model of Section 6.3.1 with $\theta = 2\gamma$ reproduces the leading behavior of the node-mediated model of Sec. 6.3.4 exactly in this limit. This is due to the fact that while the mean field approximation is far from exact, since a large number of eigenvalues of the same size as the largest one exist, nonetheless they localize in a subextensive set of links, and the respective eigenvectors are zero everywhere else.

We can state this more precisely as a two-part criterion, already implicit in the derivation of Equation (6.91). An eigenmode k of β, with eigenvalue λ_k and eigenvector $|\omega_k\rangle$, drives genuine non-mean-field phase structure only if both of the following hold in the $V \to \infty$ limit:

(i) *Extensive eigenvalue:* λ_k/V tends to a finite, nonzero constant;

(ii) *Delocalized eigenvector:* the number of entries a for which $(\omega_k)_a$ is bounded away from zero is a macroscopic (non-vanishing) fraction of all V links participates in the mode.

If either condition fails, the corresponding auxiliary field is pinned to its mean-field/trivial value in the thermodynamic limit, and the mode contributes at most a finite-size correction rather than an independent order parameter. This is exactly what happens for the non-principal eigenspaces of the Johnson graph coupling in Sec.6.3.4: the ν-eigenspace fails condition (i) outright ($\nu/V \to 0$), while the μ-eigenspace satisfies (i) ($\mu/V \to \gamma$, finite) but fails (ii), since its eigenvector has only $2N-2$ nonzero entries out of $V = N(N-1)/2$, a fraction that vanishes as $N \to \infty$. Both mechanisms independently force the corresponding auxiliary field to its trivial value, which is why only the mean-field mode survives to shape the phase diagram of Fig. 6.8.

This is not true if we want to consider the finite size effects on the system, or the sparse limit in which the scaling of the parameters is different from the one we used. In both such cases one would need to resort to the general single-link probability Formula (6.39) with the eigenvector entries that are found using the method of SI. The phase structure would have to be found with the equation described in SI, by putting Equation (6.25) equal to zero with the found entries. After that, it could also be possible to correct the model for an heterogeneous ϵ term.

We also want to highlight the connection between our approach and Algebraic Graph Theory. When we recast the computation of the Partition functions in terms of the spectral decomposition of the coupling structure we are actually looking for the diagonalization of link coupling matrix. Given that this is a matrix describing binary relations in a set of V elements, it is itself an adjacency matrix and all theorems of Algebraic Graph Theory apply.

In this respect, a rough procedure to determine the possible phase separation structure of a model can be outlined by generalizing the argument done for the goodness of the mean field approximation. In the thermodynamic limit, only eigenvalues with the right scaling, *i.e.* outliers of the link-link adjacency matrix, will be relevant to the the computation of the partition function. From literature it is known that in most cases outliers other than the mean field one are symptoms of either a community structure or a hub presence. One would indeed be interested in those generated by the community structure, as they are the one that does not localize and create the long-range correlations needed for the phase transition to appear.

This is consistent with results on the detectability of community structure via spectral methods, where a community-induced eigenvalue must separate from the bulk of the spectrum by an amount related to the community strength before the corresponding (delocalized) eigenvector becomes informative, exactly mirroring the criterion above.

This heuristics for the outliers to be responsible for the phase transitions can be further argued for in the case of the regular random graph, as we hinted at in Section 6.3.1. Indeed looking at the spectra of regular random graphs, the first one in which a separation of the spectrum appears is when the degree is 3, as can be seen in Figure 6.10

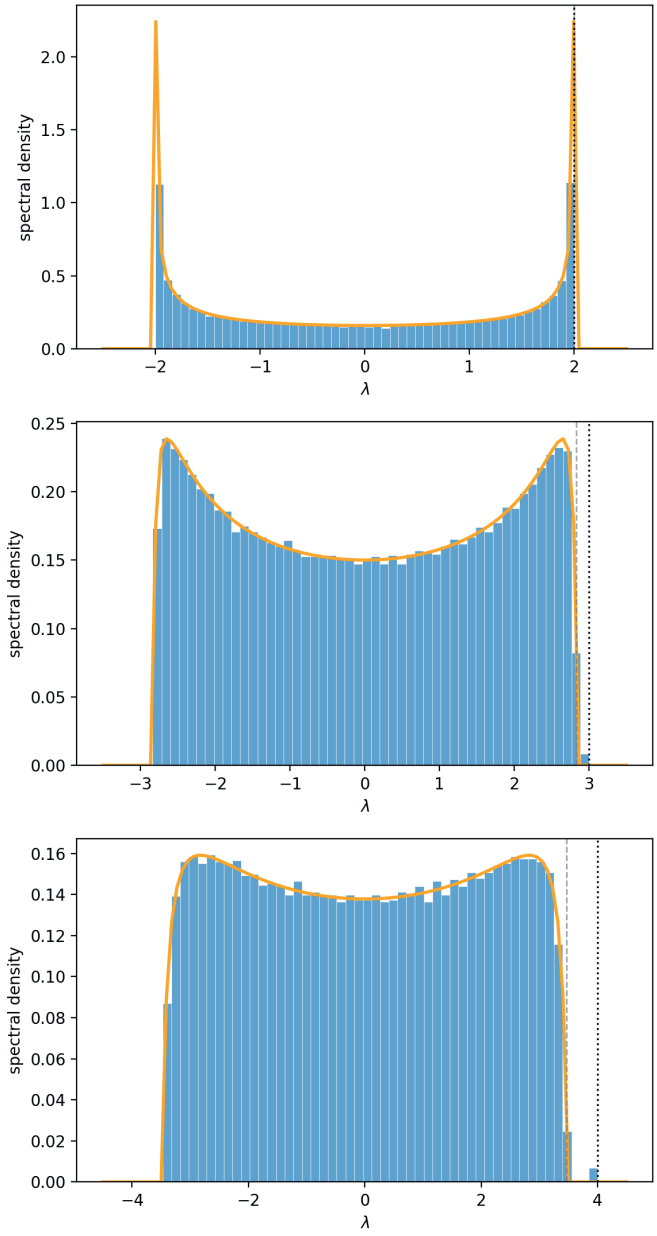


Figure 6.10: Spectrum of the adjacency matrix of a random regular graph with 1000 *nodes. In orange the Kensten-McKay law describing the distribution of the bulk of the spectrum. From top to bottom the degree k is* 2, 3, 4, *and we can see that the outlier starts to appear from the second figure only.*

§6.5 Final Remarks

The solution of Random Network Models is a fundamental, analytical tool to tackle many among the problems in the science of Complex Systems. In this paper we have shown how to solve the models that mimic a link-link coupling of real world systems. The power of our approach lies in the fact that is quite universal. Using the general Formula (6.39), heterogeneity in local features can be added at will while still retaining the same coupling structure. On the condition of having a viable way to decompose the interaction structure, we have also shown that Equation (6.21) can provide an exact, explicit solution of the partition function and the phase structure as well. An immediate practical consequence of this explicit solution is that it makes maximum likelihood estimation directly accessible for this class of second order models. Since the partition function of Equation (6.21) and the single-link probability of Equation (6.39) are available in closed, or at least saddle-point-exact, form, the likelihood of an observed network under any of the models we solve can be written down and maximized without resorting to the Markov Chain Monte Carlo or pseudo-likelihood approximations ordinarily required for pairwise-dependent exponential random graph models, and without incurring the model-degeneracy pathologies such approximate schemes are known to suffer from. We regard this as one of the more promising practical directions opened up by the present work: using the class of second order models solved here as tractable, exactly-normalizable null models for network reconstruction and hypothesis testing, in the same spirit already established for first order maximum-entropy models [90].

Moreover, in the various cases that we explored, we have seen that the phase structure can become easily crowded as the phase landscape can be quite hilly. Higher order features of networks are increasingly thought to be key features of complex systems, and the present work constitutes a little brick in the wall of understanding and solving coupled, complex structures. But we have also shown in Section 6.3.2 how seemingly higher order structures can be effectively reduced to lower order network models and be, in principle, indistinguishable from them. Lastly, the study of even higher order structure has still many more steps to undertake, and we believe some of the methods developed here, and properly generalized, could be of help in that direction of research, for example as null models to distinguish between true higher order interaction and the one just described by link-link coupling that is extensively dealt with in the present work.

Bibliography

Bibliography

[1] E. Garuccio, M. Lalli, and D. Garlaschelli, Multiscale network renormalization: Scale-invariance without geometry, *Physical Review Research* **5**, 043101 (2023).

[2] M. Lalli and D. Garlaschelli, Geometry-free renormalization of directed networks: scale-invariance and reciprocity, *arXiv preprint* arXiv:2403.00235 (2024).

[3] L. Avena, D. Garlaschelli, R. S. Hazra, and M. Lalli, Inhomogeneous random graphs with infinite-mean fitness variables, *arXiv preprint* arXiv:2212.08462 (2022).

[4] M. Newman, *Networks* (Oxford University Press, 2018).

[5] R. Albert and A.-L. Barabási, Statistical mechanics of complex networks, *Reviews of Modern Physics* **74**, 47–97 (2002).

[6] G. Caldarelli, *Scale-Free Networks: Complex Webs in Nature and Technology* (Oxford University Press, 2007).

[7] A.-L. Barabási and R. Albert, Emergence of scaling in random networks, *Science* **286**, 509–512 (1999).

[8] D. J. Watts and S. H. Strogatz, Collective dynamics of 'small-world' networks, *Nature* **393**, 440-442 (1998).

[9] M. E. Newman, Properties of highly clustered networks, *Physical Review E* **68**, 026121 (2003).

[10] P. Erdős and A. Rényi, On the evolution of random graphs, *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* (1960).

[11] G. Caldarelli, A. Capocci, P. De Los Rios, and M. A. Munoz, Scale-free networks from varying vertex intrinsic fitness, *Physical Review Letters* **89**, 258702 (2002).

[12] D. Garlaschelli and M. I. Loffredo, Fitness-dependent topological properties of the world trade web, *Physical Review Letters* **93**, 188701 (2004).

[13] M. Boguná and R. Pastor-Satorras, Class of correlated random networks with hidden variables, *Physical Review E* **68**, 036112 (2003).

[14] M. Á. Serrano and M. Boguna, Clustering in complex networks. I. General formalism, *Physical Review E* **74**, 056114 (2006).

[15] M. Á. Serrano and M. Boguná, Clustering in complex networks. II. Percolation properties, *Physical Review E* **74**, 056115 (2006).

[16] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguná, Hyperbolic geometry of complex networks, *Physical Review E* **82**, 036106 (2010).

[17] R. Michielan, N. Litvak, and C. Stegehuis, Detecting hyperbolic geometry in networks: Why triangles are not enough, *Physical Review E* **106**, 054303 (2022).

[18] D. Krioukov, Clustering implies geometry in networks, *Physical Review Letters* **116**, 208302 (2016).

[19] R. Aliakbarisani, M. Boguñá, and M. Á. Serrano, Clustering does not always imply latent geometry, *Physical Review Letters* **135**, 197402 (2025).

[20] G. Garcia-Perez, M. Boguna, and M. A. Serrano, Multiscale unfolding of real networks by geometric renormalization, *Nature Physics* **14**, 583 (2018).

[21] M. Boguna, I. Bonamassa, M. De Domenico, S. Havlin, D. Krioukov, and M. Á. Serrano, Network geometry, *Nature Reviews Physics* **3**, 114 (2021).

[22] J. Nolan, *Univariate Stable Distributions: Models for Heavy Tailed Data*, Springer Series in Operations Research and Financial Engineering (Springer, 2020).

[23] A. Baptista, R. J. Sánchez-García, A. Baudot, and G. Bianconi, Zoo guide to network embedding, *Journal of Physics: Complexity* **4**, 042001 (2023).

[24] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, *et al.*, Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study, *Proceedings of the National Academy of Sciences* **105**, 1232-1237 (2008).

[25] P. Van Mieghem, *Graph Spectra for Complex Networks* (Cambridge University Press, 2023).

[26] X. Zhang, R. R. Nadakuditi, and M. E. Newman, Spectra of random graphs with community structure and arbitrary degrees, *Physical Review E* **89**, 042816 (2014).

[27] F. Krzakala, C. Moore, E. Mossel, J. Neeman, A. Sly, L. Zdeborová, and P. Zhang, Spectral redemption in clustering sparse networks, *Proceedings of the National Academy of Sciences* **110**, 20935 (2013).

[28] C. Bordenave, M. Lelarge, and L. Massoulié, Non-backtracking spectrum of random graphs: Community detection and non-regular Ramanujan graphs, in *2015 IEEE 56th Annual Symposium on Foundations of Computer Science* (IEEE, 2015), pp. 1347–1357.

[29] F. Benaych-Georges and A. Knowles, Lectures on the local semicircle law for Wigner matrices, *arXiv preprint* arXiv:1601.04055 (2016).

[30] Z. Bai and J. W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, Vol. 20 (Springer, 2010).

[31] F. Chung, L. Lu, and V. Vu, Spectra of random graphs with given expected degrees, *Proceedings of the National Academy of Sciences* **100**, 6313 (2003).

[32] J. Baik, G. Ben Arous, and S. Péché, Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices, *Annals of Probability* (2005).

[33] R. R. Nadakuditi and M. E. Newman, Spectra of random graphs with arbitrary expected degrees, *Physical Review E* **87**, 012803 (2013).

[34] F. Benaych-Georges and R. R. Nadakuditi, The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices, *Advances in Mathematics* **227**, 494 (2011).

[35] D. Sornette, Discrete-scale invariance and complex dimensions, *Physics Reports* **297**, 239 (1998).

[36] D. Sornette, A. Johansen, and J.-P. Bouchaud, Stock market crashes, precursors and replicas, *Journal de Physique I* **6**, 167–175 (1996).

[37] A. Johansen, D. Sornette, and O. Ledoit, Predicting financial crashes using discrete scale invariance, *Journal of Risk* **1**, 5–32 (1999).

[38] D. Sornette, *Why Stock Markets Crash: Critical Events in Complex Financial Systems* (Princeton University Press, 2003).

[39] M. E. J. Newman, The structure and function of complex networks, *SIAM Review* **45**(2), 167–256 (2003).

[40] M. E. Newman and J. Park, Why social networks are different from other types of networks, *Physical Review E* **68**, 036122 (2003).

[41] T. Squartini, G. Caldarelli, G. Cimini, A. Gabrielli, and D. Garlaschelli, Reconstruction methods for networks: The case of economic and financial systems, *Physics Reports* **757**, 1-47 (2018).

[42] J.-P. Fouque and J. A. Langsam, *Handbook on Systemic Risk* (Cambridge University Press, 2013).

[43] S. Battiston, M. Puliga, R. Kaushik, P. Tasca, and G. Caldarelli, DebtRank: Too central to fail? Financial networks, the FED and systemic risk, *Scientific Reports* **2**, 541 (2012).

[44] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, Network motifs: Simple building blocks of complex networks, *Science* **298**(5594), 824–827 (2002).

[45] U. Alon, Network motifs: Theory and experimental approaches, *Nature Reviews Genetics* **8**(6), 450–461 (2007).

[46] Z. N. Oltvai and A.-L. Barabási, Life's complexity pyramid, *Science* **298**(5594), 763–764 (2002).

[47] G. Cimini, T. Squartini, F. Saracco, D. Garlaschelli, A. Gabrielli, and G. Caldarelli, The statistical physics of real-world networks, *Nature Reviews Physics* **1**, 58-71 (2019).

[48] J. Park and M. E. J. Newman, Statistical mechanics of networks, *Physical Review E* **70**(6), 066117 (2004).

[49] A. Annibale and O. T. Courtney, The two-star model: Exact solution in the sparse regime and condensation transition, *Journal of Physics A: Mathematical and Theoretical* **48**, 365001 (2015).

[50] J. Park and M. E. J. Newman, Solution of the two-star model of a network, *Physical Review E* **70**, 066146 (2004).

[51] C. P. Burgess, *Introduction to Effective Field Theory: Thinking Effectively about Hierarchies of Scale* (Cambridge University Press, 2020).

[52] A. V. Manohar, Introduction to effective field theories, *arXiv preprint* arXiv:1804.05863 (2018).

[53] R. Shankar, Effective field theory in condensed matter physics, *arXiv preprint* arXiv:cond-mat/9703210 (1997).

[54] A. Gabrielli, D. Garlaschelli, S. P. Patil, and M. Á. Serrano, Network renormalization, *Nature Reviews Physics*, **7**, 203-219 (2025).

[55] H. Stanley, L. Amarala, P. Gopikrishnana, P. Ivanova, T. Keittb, V. Pleroua, Scale invariance and universality: organizing principles in complex systems, *Physica A: Statistical Mechanics and its Applications*, **281** 60-68 (2000).

[56] F. Radicchi, A. Barrat, S. Fortunato1, and J. Ramasco, Renormalization flows in complex networks, *Physical Review E* **79**, 026104 (2009).

[57] K. G. Wilson and J. Kogut, The renormalization group and the ϵ expansion, *Physics Reports* **12**, 75 (1974).

[58] A. M. Schakel, *Boulevard of Broken Symmetries: Effective Field Theories of Condensed Matter* (World Scientific, 2008).

[59] G. Clemente, C. Tessone, D. Garlaschelli, Temporal networks with node-specific memory: Unbiased inference of transition probabilities, relaxation times, and structural breaks, *Physical Research Review* **6**, 043257 (2024).

[60] G. Clemente, D. Garlaschelli, Linking through time: Memory-enhanced community discovery in temporal networks, *Physical Research Review* **6**, 043204 (2024).

[61] T. Squartini, G. Fagiolo, and D. Garlaschelli (2013), Null Models of Economic Networks: The Case of the World Trade Web, *Journal of Economic Interaction and Coordination* **8**, 75-107 (2013).

[62] V. Gemmetto and D. Garlaschelli Multiplexity versus correlation: the role of local constraints in real multiplexes, *Scientific Reports* **5**, 9120 (2015).

[63] V. Gemmetto and D. Garlaschelli Multiplexity versus correlation: the role of local constraints in real multiplexes, *Scientific Reports* **5**, 9120 (2015).

[64] R. Van Der Hofstad, *Random Graphs and Complex Networks*, Vol. 1 (Cambridge University Press, 2017).

[65] R. Van Der Hofstad, *Random Graphs and Complex Networks*, Vol. 2 (Cambridge University Press, 2024).

[66] R. Milocco, F. Jansen, and D. Garlaschelli, Multi-scale node embeddings for graph modeling and generation, *arXiv preprint* arXiv:2412.04354 (2024).

[67] I. J. Farkas, I. Derényi, A.-L. Barabási, and T. Vicsek, Spectra of "real-world" graphs: Beyond the semicircle law, *Physical Review E* **64**, 026704 (2001).

[68] L. Poley, T. Galla, and J. W. Baron, Eigenvalue spectra of finely structured random matrices, *Physical Review E* **109**, 064301 (2024).

[69] J. W. Baron, T. J. Jewell, C. Ryder, and T. Galla, Eigenvalues of random matrices with generalized correlations: A path integral approach, *Physical Review Letters* **128**, 120601 (2022).

[70] P. Cizeau and J.-P. Bouchaud, Theory of Lévy matrices, *Physical Review E* **50**, 1810 (1994).

[71] F. Benaych-Georges, A. Guionnet, and C. Male, Central limit theorems for linear statistics of heavy-tailed random matrices, *Communications in Mathematical Physics* **329**, 641 (2014).

[72] G. B. Arous and A. Guionnet, The spectrum of heavy-tailed random matrices, *Communications in Mathematical Physics* **278**, 715 (2008).

[73] I. Afanasiev, L. Berlyand, and M. Kiyashko, Asymptotic behavior of eigenvalues of large rank perturbations of large random matrices, *arXiv preprint* arXiv:2507.12182 (2025).

[74] I. Alexeev, Stable random variables with complex stability index, I, *Theory of Probability & Its Applications* **67**, 335 (2022).

[75] I. Alexeev, Stable random variables with complex stability index, II, *Theory of Probability & Its Applications* **67**, 499 (2023).

[76] J. Huang, Mesoscopic perturbations of large random matrices, *Random Matrices: Theory and Applications* **7**, 1850004 (2018).

[77] R. Van Handel, On the spectral norm of Gaussian random matrices, *Transactions of the American Mathematical Society* **369**, 8161 (2017).

[78] R. Latała, Some estimates of norms of random matrices, *Proceedings of the American Mathematical Society* **133**, 1273 (2005).

[79] S. Riemer and C. Schütt, On the expectation of the norm of random matrices with non-identically distributed entries, *Electronic Journal of Probability* **18**, 1 (2013).

[80] D. Cheliotis and M. Louvaris, The limit of the operator norm for random matrices with a variance profile, *Electronic Journal of Probability* **30**, 1 (2025).

[81] F. Benaych-Georges, C. Bordenave, and A. Knowles, Spectral radii of sparse random matrices, in *Annales de l'Institut Henri Poincaré—Probabilités et Statistiques*, Vol. 56 (2020), pp. 2141–2161.

[82] F. Benaych-Georges, C. Bordenave, and A. Knowles, Largest eigenvalues of sparse inhomogeneous Erdős–Rényi graphs, *The Annals of Probability* **47**, 1653 (2019).

[83] A. Bandeira and R. van Handel, Sharp nonasymptotic bounds on the norm of random matrices with independent entries, *The Annals of Probability* **44**, 2479 (2016).

[84] R. Van Handel, Structured random matrices, *Convexity and Concentration*, **161**, pp. 107-165 (2017).

[85] J. McWhirter and E. R. Pike, On the numerical inversion of the Laplace transform and similar Fredholm integral equations of the first kind, *Journal of Physics A: Mathematical and General* **11**, 1729 (1978).

[86] N. Ostrowsky, D. Sornette, P. Parker, and E. Pike, Exponential sampling method for light scattering polydispersity analysis, *Optica Acta: International Journal of Optics* **28**, 1059 (1981).

[87] M. E. Newman, Random graphs with clustering, *Physical Review Letters* **103**, 058701 (2009).

[88] F. Battiston and G. Petri, *Higher-Order Systems* (Springer, 2022).

[89] S. Boccaletti, P. De Lellis, C. Del Genio, K. Alfaro-Bittner, R. Criado, S. Jalan, and M. Romance, The structure and dynamics of networks with higher order interactions, *Physics Reports* **1018**, 1-64 (2023).

[90] T. Squartini and D. Garlaschelli, *Maximum-Entropy Networks: Pattern Detection, Network Reconstruction and Graph Combinatorics* (Springer, 2017).

[91] A. Allard, M. Á. Serrano, and M. Boguñá, Geometric description of clustering in directed networks, *Nature Physics* **20**, 150-156 (2024).

[92] E. Candellero and N. Fountoulakis, Clustering and the hyperbolic geometry of complex networks, in *Algorithms and Models for the Web Graph: 11th International Workshop, WAW 2014* (Springer, 2014), 1–12.

[93] M. Boguna, D. Krioukov, and K. C. Claffy, Navigability of complex networks, *Nature Physics* **5**, 74-80 (2009).

[94] M. Boguná, F. Papadopoulos, and D. Krioukov, Sustaining the internet with hyperbolic mapping, *Nature Communications* **1**, 62 (2010).

[95] A. Catanzaro, R. S. Hazra, and D. Garlaschelli, Spectra of random graphs with discrete scale invariance, *arXiv preprint* arXiv:2509.12407 (2025).

[96] I. Norros and H. Reittu, On a conditionally Poissonian graph process, *Advances in Applied Probability* **38**, 59–75 (2006).

[97] G. J. Rodgers, K. Austin, B. Kahng, and D. Kim, Eigenvalue spectra of complex networks, *Journal of Physics A: Mathematical and General* **38**, 9431 (2005).

[98] S. Bhamidi, R. van der Hofstad, and J. van Leeuwaarden, Scaling limits for critical inhomogeneous random graphs with finite third moments, *Electronic Journal of Probability* **15**, 1682-1702, (2010).

[99] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, Vol. 1 (CRC Press, 1994).

[100] B. Bollobás, W. Fulton, A. Katok, *et al.*, *Random Graphs* (Cambridge University Press, 2001).

[101] S. Janson, T. Łuczak, and A. Ruciński, *Random Graphs* (Wiley, 2000).

[102] L. P. Kadanoff, Scaling laws for Ising models near T_c, *Physics Physique Fizika* **2**(6), 263–272 (1966).

[103] K. G. Wilson, The renormalization group: Critical phenomena and the Kondo problem, *Reviews of Modern Physics* **47**(4), 773–840 (1975).

[104] P. Holme and J. Saramäki, Temporal networks, *Physics Reports* **519**(3), 97–125 (2012).

[105] B. Ribeiro, N. Perra, and A. Baronchelli, Quantifying the effect of temporal resolution on time-varying networks, *Scientific Reports* **3**, 3006 (2013).

[106] G. Krings, M. Karsai, S. Bernhardsson, V. D. Blondel, and J. Saramäki, Effects of time window size and placement on the structure of an aggregated communication network, *EPJ Data Science* **1**, 4 (2012).

[107] D. Strauss, On a general class of models for interaction, *SIAM Review* **28**, 513–527 (1986).

[108] O. Häggström and J. Jonasson, Phase transition in the random triangle model, *Journal of Applied Probability* **36**, 1101–1115 (1999).

[109] J. Park and M. E. J. Newman, Solution for the properties of a clustered network, *Physical Review E* **72**, 026136 (2005).

[110] A. Coolen, A. Annibale, and E. Roberts, *Generating Random Networks and Graphs* (Oxford University Press, 2017).

[111] E. T. Jaynes, Information theory and statistical mechanics, *Physical Review* **106**, 620 (1957).

[112] E. T. Jaynes, Information theory and statistical mechanics. II, *Physical Review* **108**, 171 (1957).

[113] C. E. Shannon, A mathematical theory of communication, *The Bell System Technical Journal* **27**, 379 (1948).

[114] A. Catanzaro, S. Patil, and D. Garlaschelli, A general solution for network models with pairwise edge coupling, *arXiv preprint* arXiv:2503.20805 (2025).

[115] B. B. Machta, R. Chachra, M. K. Transtrum, and J. P. Sethna, Parameter space compression underlies emergent theories and predictive models, *Science* **342**, 604–607 (2013).

[116] 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 preprint* arXiv:2111.07176 (2022).

[117] C. R. Rao, Information and accuracy attainable in the estimation of statistical parameters, *Bulletin of the Calcutta Mathematical Society* **37**, 81 (1945).

[118] F. Nielsen, An elementary introduction to information geometry, *Entropy* **22**, 1100 (2020).

[119] S. Amari and H. Nagaoka, *Methods of Information Geometry*, Translations of Mathematical Monographs (American Mathematical Society, 2000).

[120] 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, *The Journal of Chemical Physics* **143**, 010901 (2015).

[121] L. P. Kadanoff, Scaling, universality and operator algebras, *Phase Transitions and Critical Phenomena* (1976).

[122] 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 Press, 1976).

[123] L. D. Landau and E. M. Lifshitz, *Physical Kinetics*, Course of Theoretical Physics, Vol. 10 (Pergamon Press, Oxford, 1981).

[124] S. Succi, *The Lattice Boltzmann Equation: For Fluid Dynamics and Beyond*, Numerical Mathematics and Scientific Computation (Oxford University Press, 2013).

[125] K. T. Grosvenor, N. A. Obers, and S. P. Patil, Hydrodynamics without boost-invariance from kinetic theory: From perfect fluids to active flocks, *SciPost Physics* **19**, 071 (2025).

[126] M. Helias and D. Dahmen, *Statistical Field Theory for Neural Networks* (Springer International Publishing, 2020).

[127] A. Catanzaro, S. P. Patil, and D. Garlaschelli, *To appear.*

[128] Y. Bar-Yam and S. P. Patil, Renormalization of sparse disorder in the Ising model, *arXiv preprint* arXiv:1805.12556 (2018).

[129] A. Migdal, Phase transitions in gauge and spin-lattice systems, *Zh. Eksp. Teor. Fiz.* **69**, 1457 (1975).

[130] L. P. Kadanoff, Notes on Migdal's recursion formulas, *Annals of Physics* **100**, 359 (1976).

[131] L. P. Kadanoff, *Statistical Physics: Statics, Dynamics and Renormalization* (World Scientific, 2000).

[132] L. Onsager, Crystal statistics. I. A two-dimensional model with an order-disorder transition, *Physical Review* **65**, 117 (1944).

[133] H. J. Hilhorst, M. Schick, and J. M. J. van Leeuwen, Exact renormalization group equations for the two-dimensional Ising model, *Physical Review B* **19**, 2749 (1979).

[134] G. G. Piva, F. L. Ribeiro, and A. S. Mata, Networks with growth and preferential attachment: Modeling and applications, *arXiv preprint* arXiv:2007.01349 (2020).

[135] V. D. P. Servedio, G. Caldarelli, and P. Buttà, Vertex intrinsic fitness: How to produce arbitrary scale-free networks, *Physical Review E* **70**, 056126 (2004).

[136] C. Schmeltzer, J. Soriano, I. M. Sokolov, and S. Rüdiger, Percolation of spatially constrained Erdős-Rényi networks with degree correlations, *Physical Review E* **89**, 012116 (2014).

[137] M. Alizadeh, C. Cioffi-Revilla, and A. Crooks, Generating and analyzing spatial social networks, *Computational and Mathematical Organization Theory* **23**, 362 (2017).

[138] X. Shi, L. A. Adamic, and M. J. Strauss, Networks of strong ties, *Physica A: Statistical Mechanics and its Applications* **378**, 33 (2007).

[139] C. Seshadhri, T. G. Kolda, and A. Pinar, Community structure and scale-free collections of Erdős-Rényi graphs, *Physical Review E* **85**, 056109 (2012).

[140] M. Molloy and B. Reed, A critical point for random graphs with a given degree sequence, *Random Structures & Algorithms* **6**, 161-180 (1995).

[141] G. Schoenebeck and F.-Y. Yu, Consensus of interacting particle systems on Erdös-Rényi graphs, in *Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2018)*, pp. 1945–1964 (2018).

[142] 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).

[143] F. D. Domenico, F. Caccioli, and G. Livan, Peer influence breaks ergodicity in an opinion dynamics model with external information, *arXiv preprint* arXiv:2507.06661 (2025).

[144] C. Fan, L. Yao, J. Zhang, Z. Zhen, and X. Wu, Advanced reinforcement learning and its connections with brain neuroscience, *Research* **6**, 0064 (2023).

[145] K. Doya, Reinforcement learning: Computational theory and biological mechanisms, *HFSP Journal* **1**, 30–40 (2007).

[146] T. Alkam, A. H. Van Benschoten, and E. Tarshizi, Reinforcement learning in artificial intelligence and neurobiology, *Neuroscience Informatics* **5**, 100220 (2025).

[147] M. Khoshkhou and A. Montakhab, Optimal reinforcement learning near the edge of a synchronization transition, *Physical Review E* **105**, 044312 (2022).

[148] N. Rungratsameetaweemana, R. Kim, T. Chotibut, and T. J. Sejnowski, Random noise promotes slow heterogeneous synaptic dynamics important for robust working memory computation, *Proceedings of the National Academy of Sciences* **122**, e2316745122 (2025).

[149] 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, *Nature Communications* **9**, 2383 (2018).

[150] Y. Arya and S. H. Lee, Effects of structural properties of neural networks on machine learning performance, *arXiv preprint* arXiv:2507.10005 (2025).

[151] G. Cimini, R. Mastrandrea, and T. Squartini, *Reconstructing Networks* (Cambridge University Press, 2021).

[152] Y.-Y. Ahn, J. P. Bagrow, and S. Lehmann, Link communities reveal multiscale complexity in networks, *Nature* **466**, 761-764 (2010).

[153] P. Villegas, A. Gabrielli, F. Santucci, G. Caldarelli, and T. Gili, Laplacian paths in complex networks: Information core emerges from entropic transitions, *Physical Review Research* **4**, 033196 (2022).

[154] G. Pólya, G. Szegö, *et al.*, *Problems and Theorems in Analysis: Series, Integral Calculus, Theory of Functions*, Vol. 1 (Springer, 1972).

[155] P. Delsarte, An algebraic approach to the association schemes of coding theory, *IEEE Transactions on Information Theory*, **44**(6), 2477-2504 (1998).

[156] K. Vorob'ev, On reconstruction of eigenfunctions of Johnson graphs, *Discrete Applied Mathematics* **276**, 166 (2020).

[157] Y. Filmus, An orthogonal basis for functions over a slice of the Boolean hypercube, *Electronic Journal of Combinatorics*, **23**(1), P1.23 (2016).

[158] P.-O. Löwdin, On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals, *The Journal of Chemical Physics* **18**, 365 (1950).

[159] F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain (eds.), *NIST Digital Library of Mathematical Functions*, Release 1.2.4 of 2025-03-15, `https://dlmf.nist.gov/`.

[160] R. LePage, M. Woodroofe, and J. Zinn, Convergence to a stable distribution via order statistics, *The Annals of Probability*, **9**, pp. 624-632, (1981).

[161] R. Leipus, J. Šiaulys, and D. Konstantinides, *Closure Properties for Heavy-Tailed and Related Distributions—An Overview*, SpringerBriefs in Statistics (Springer, 2023).

[162] E. N. Gilbert, Random graphs, *The Annals of Mathematical Statistics* **30**, 1141 (1959).

[163] H. Van Den Esker, R. Van Der Hofstad, G. Hooghiemstra, and D. Znamenski, Distances in random graphs with infinite mean degrees, *Extremes* **8**, 111-141 (2005).

[164] R. Van Der Hofstad, G. Hooghiemstra, and D. Znamenski, Random graphs with arbitrary i.i.d. degrees, *arXiv preprint* arXiv:math/0502580 (2005).

[165] A. Rogers, *Supermanifolds: Theory and Applications* (World Scientific, 2007).

[166] V. H. Can, Critical behavior of the annealed Ising model on random regular graphs, *Journal of Statistical Physics* **169**, 480 (2017).

[167] V. H. Can, Annealed limit theorems for the Ising model on random regular graphs, *The Annals of Applied Probability* **29**, 1398-1445 (2019).

[168] V. H. Can, C. Giardinà, C. Giberti, and R. van der Hofstad, Annealed Ising model on configuration models, in *Annales de l'Institut Henri Poincaré – Probabilités et Statistiques*, Vol. 58 (Institut Henri Poincaré, 2022), 134–163.

[169] F. Wegner, Some invariance properties of the renormalization group, *Journal of Physics C: Solid State Physics* **7**, 2098 (1974).

[170] J. Polchinski, Renormalization and effective Lagrangians, *Nuclear Physics B* **231**, 269 (1984).

[171] C. Bagnuls and C. Bervillier, Exact renormalization group equations: An introductory review, *Physics Reports* **348**, 91–157 (2001).

[172] J. Berges, N. Tetradis, and C. Wetterich, Non-perturbative renormalization flow in quantum field theory and statistical physics, *Physics Reports* **363**, 223–386 (2002).

[173] T. R. Morris, A gauge invariant exact renormalisation group. (I), *Nuclear Physics B* **573**, 97–126 (2000).

[174] A. Raju, B. B. Machta, and J. P. Sethna, Information loss under coarse graining: A geometric approach, *Physical Review E* **98**, 052112 (2018).

[175] W. Janke, D. Johnston, and R. Kenna, Information geometry and phase transitions, *Physica A: Statistical Mechanics and its Applications* **336**, 181–186 (2004).

[176] B. P. Dolan, D. A. Johnston, and R. Kenna, The information geometry of the one-dimensional Potts model, *Journal of Physics A: Mathematical and General* **35**, 9025–9035 (2002).

[177] J. Erdmenger, K. Grosvenor, and R. Jefferson, Information geometry in quantum field theory: Lessons from simple examples, *SciPost Physics* **8**, 073 (2020).

[178] V. Balasubramanian, J. J. Heckman, and A. Maloney, Relative entropy and proximity of quantum field theories, *arXiv preprint* arXiv:1410.6809 (2015).

[179] F. Frenet, Sur les courbes à double courbure, *Journal de Mathématiques Pures et Appliquées* **17**, 890 (1852).

[180] J. A. Serret, Sur quelques formules relatives à la théorie des courbes à double courbure, *Journal de Mathématiques Pures et Appliquées* **16**, 153 (1851).

[181] A. Negro and S. P. Patil, An Étude on the regularization and renormalization of divergences in primordial observables, *Rivista del Nuovo Cimento* **47**, 179 (2024).

[182] A. Andreassen, W. Frost, and M. D. Schwartz, Consistent use of the standard model effective potential, *Physical Review Letters* **113**, 241801 (2014).

[183] A. D. Jackson and S. P. Patil, Phases of small worlds: A mean field formulation, *Journal of Statistical Physics* **189**, 40 (2022).

[184] J. Cotler and S. Rezchikov, Renormalization group flow as optimal transport, *Physical Review D* **108**, 025003 (2023).

[185] D. S. Berman, M. S. Klinger, and A. G. Stapleton, Bayesian renormalization, *Machine Learning: Science and Technology* **4**, 045011 (2023).

[186] D. S. Berman and M. S. Klinger, The inverse of exact renormalization group flows as statistical inference, *Entropy* **26**, 389 (2024).

[187] T. Squartini, R. Mastrandrea, and D. Garlaschelli, Unbiased sampling of network ensembles, *New Journal of Physics* **17**, 023052 (2015).

[188] M. Z. Rácz and D. E. Rigobon, Towards consensus: Reducing polarization by perturbing social networks, *IEEE Transactions on Network Science and Engineering* **10**, 3450 (2023).

[189] K. Kandhway and J. Kuri, Campaigning in heterogeneous social networks: Optimal control of SI information epidemics, *IEEE/ACM Transactions on Networking* **24**, 383 (2016).

[190] C. Jiang, Y. Chen, and K. J. R. Liu, Evolutionary dynamics of information diffusion over social networks, *IEEE Transactions on Signal Processing* **62**, 4573 (2014).

Summary

This thesis studies the renormalization of random network models, with a focus on the topological and algebraic structures that emerge when networks are coarse-grained or partially observed. The work is divided into two parts.

The first part analyses the Multi-Scale Model (MSM), a random graph model that arises as the fixed point of a renormalization procedure based on aggregating nodes into supernodes while preserving the functional form of the connection probability. The MSM requires that node fitnesses follow a heavy-tailed stable distribution with infinite mean, and this unusual distributional property endows the model with a number of remarkable features.

Chapter 2 characterises the spectrum of the MSM adjacency matrix. The leading eigenvalues are shown to grow as the square root of the network size, to alternate in sign, and to lie at the intersections of the real axis with a logarithmic spiral in the complex plane, characterised analytically in terms of the Gamma function. The corresponding eigenvectors display log-periodic oscillations in the node fitness, a hallmark of discrete scale invariance. The bulk of the spectrum extends to the same scale as the outlier eigenvalues, in stark contrast with the behaviour of finite-mean models.

Chapter 3 proves that the MSM naturally generates a non-trivial clustering function without any appeal to an underlying geometric space. In the infinite-mean regime, all nodes except the highest-degree hubs have local clustering coefficients close to one, and the network becomes asymptotically fully clustered in a precise sense. The analysis also reveals a lack of self-averaging in the average clustering coefficient, a consequence of the heavy-tailed fitness distribution.

The second part develops a renormalization theory for random network models in which edges are correlated. The central observation of Chapter 4 is that marginalising a correlated network model over unobserved variables is equivalent to performing a decimation step in the renormalization group sense. This equivalence is formalised within the framework of exponential random graph models, and is used to derive exact RG flow equations, to identify universality classes, and to develop network inference methods that incorporate uncertainty about missing data in a principled way.

Chapter 5 solves the RG equations exactly for one-dimensional chain topologies, and extends the analysis to models with sparse and uniform disorder. The exact solutions are interpreted in terms of several network-relevant phenomena, including peer effects, preferential attachment, susceptibility, and reinforcement dynamics.

Chapter 6 analyses the class of second-order random network models, to which most correlated models flow under renormalization when higher-order interactions are irrelevant. It is shown that second-order models can be solved in full generality

by reducing the computation of the partition function to the diagonalization of a link-adjacency matrix. This approach reveals a rich phase structure, establishes a connection with algebraic graph theory, and provides a unified framework for solving a wide family of edge-correlated network models.

Samenvatting

Dit proefschrift bestudeert de renormalisatie van willekeurige netwerkmodellen, met bijzondere aandacht voor de topologische en algebraïsche structuren die ontstaan wanneer netwerken worden coarse-grained of gedeeltelijk worden waargenomen. Het werk is verdeeld in twee delen.

Het eerste deel analyseert het Multi-Scale Model (MSM), een willekeurig netwerkmodel dat voortkomt als het vaste punt van een renormalisatieprocedure waarbij knooppunten worden samengevoegd tot superknooppunten, terwijl de functionele vorm van de connection probability behouden blijft. Het MSM vereist dat de fitnesses van de nodes een heavy-tailed verdeling met oneindige verwachtingswaarde volgen, en deze eigenschap geeft het model een aantal opmerkelijke kenmerken.

Hoofdstuk 2 karakteriseert het spectrum van de bogenmatrix van het MSM. De dominante eigenwaarden groeien als de vierkantswortel van de netwerkgrootte, wisselen van teken, en liggen op de snijpunten van de reële as met een logaritmische spiraal in het complexe vlak, analytisch gekarakteriseerd via de Gammafunctie. De bijbehorende eigenvectoren vertonen log-periodieke oscillaties als functie van de node fitness, een kenmerk van discrete zelfgelijkvormigheid. De bulk van het spectrum strekt zich uit tot dezelfde schaal als de outlier eigenwaarden, in schril contrast met het gedrag van finite-mean modellen.

Hoofdstuk 3 bewijst dat het MSM van nature een niet-triviale clustering-functie genereert zonder gebruik te maken van een onderliggende geometrische ruimte. In het infinite-mean regime heeft nagenoeg iedere node, op de hoogst verbonden hubs na, een local clustering coefficient dicht bij één, en wordt het netwerk in precieze zin asymptotisch volledig geclusterd. De analyse onthult ook een gebrek aan self-averaging in de gemiddelde clustering coefficient, een gevolg van de heavy-tailed fitness-verdeling.

Het tweede deel ontwikkelt een renormalisatietheorie voor willekeurige netwerkmodellen met gecorreleerde links. De centrale observatie van Hoofdstuk 4 is dat het marginaliseren van een gecorreleerd netwerkmodel over ongeobserveerde variabelen equivalent is aan het uitvoeren van een decimatiestap in de renormalization group. Deze equivalentie wordt geformaliseerd binnen het kader van exponential random graph models, en wordt gebruikt om exacte RG-flow-vergelijkingen af te leiden, universaliteitsklassen te identificeren, en methoden te ontwikkelen voor netwerkinferentie met ontbrekende gegevens.

Hoofdstuk 5 lost de RG-vergelijkingen exact op voor eendimensionale keten topologieën, en breidt de analyse uit naar modellen met schaarse en uniforme wanorde. De exacte oplossingen worden geïnterpreteerd in termen van peer effects, preferential attachment, susceptibility en reinforcement-dynamica.

Hoofdstuk 6 analyseert de klasse van tweede-orde willekeurige netwerkmodellen, waarnaar de meeste gecorreleerde modellen stromen onder renormalisatie wanneer hogere-orde interacties irrelevant zijn. Er wordt aangetoond dat second-order modellen in volledige algemeenheid kunnen worden opgelost door de berekening van de partitiefunctie te reduceren tot de diagonalisatie van een link-adjacency matrix. Deze aanpak onthult een rijke fasestructuur, legt een verband met algebraïsche graaftheorie, en biedt een verenigend kader voor het oplossen van een brede familie van link-gecorreleerde netwerkmodellen.

Acknowledgements

I am bad enough at writing these acknowledgements, and the people that were close to me during my PhD trajectory are simply too many to thank them all as they should. Hopefully each one will find a trace of themselves in what follows.

This thesis wouldn't simply be here without my supervisors Diego Garlaschelli and Subodh Patil. Diego welcomed (and selected) me at IMT Lucca, nudged me to Leiden and finally sent me around the world to show my research and prove my worth. I am very thankful to him for all the discussions we had, both professional and personal ones and will always remember fondly those with a chalk at the blackboard. I would like to thank Subodh as well for hosting me in Leiden (more than once), teaching me to question always assumptions and frameworks, and to remember that fundamental questions are often more informative than partial answers. I have always appreciated the fact that they both treated me as a colleague more than a employee, respected my time, and that both intended my time as a PhD student not just as a moment to produce a research output but for me to grow, personally and as a scholar.

Growing as a scientist requires a community, and I am very thankful of having found great ones in these years. I would like to thank all members, past, present and future of the Networks Research Unit in Lucca, especially those who I got the privilege to meet and to learn from; all my colleagues in the ENBA track whom I shared the crazy economics exams with; all the people belonging to the wider Fostering Research Synergies community - I have always liked how we managed to tackle working together in a relaxed and easy-going way; and my co-authors Rajat Hazra and Remco van der Hofstad, who were so kind to suffer my lack of rigor in the mathematics we developed together.

I would not come back to Leiden this often, and with such a pleasure, if it weren't for the people that I met during my visiting period, that still offer me a bed from time to time and which made the time I spent here so enjoyable: the people from the Boerhavelaan student house; the Italian people in Leiden that are now scattered around the world; past and present colleagues from the Lorentz Institute; being a foreigner was a new and scary experience to me but it was great to be welcomed so wholeheartedly and to still be whenever I am back.

A special mention to my friends from uni time that managed to come to the thesis discussion: we had planned a big celebration for our common graduation back in the days, but Covid did not allow for it; hopefully we can make amends and compensate with this ceremony (as I am the last to get the PhD, I know).

A special thanks to my paranynphs. To Francesco: I will enjoy every minute in which I made you look like a penguin by wearing a tailcoat. To Francesca: thanks for

sharing basically the whole PhD with me and for bearing my apparently 'outlandish ideas'.

To my family, all the relatives that came to Leiden to cheer with me and those who couldn't, thanks for always supporting me, no matter what.

Finally, recent events made me aware of how many people can actually worry if something happens to me. It was truly heartwarming to read from everyone concerned about my situation and to feel the solidarity of a greater community of people around me. I thank you all again for your support and sympathy.

Curriculum Vitae

Alessio Catanzaro was born on June 19th 1995 in Bologna, Italy. He grew up there, where he attended the High School Luigi Galvani and obtained his International Scientific diploma in 2014.

He chose to study Physics at the Alma Mater Studiorum - Università of Bologna, but he was also lucky enough to be selected in the merits program of the University, the Collegio Superiore, where he first heard about network science. He obtained a Bachelor of Science in Physics in 2017, with a thesis on spectra of Random Matrices and their application to random walks, under the supervision of prof. A. Bazzani. Being fascinated by theoretical aspects of the field of physics, he obtained his Master of Sciences in Theoretical Physics with a thesis on the self-similarities properties of random networks, again under the supervision of prof. A. Bazzani, awarded by Bologna University in 2020, a few weeks into the pandemics lockdown.

In September 2020 he was admitted to the PhD program in System Science, track in Economics, Networks and Business Analytics at the IMT School for Advanced Studies in Lucca. There he joined the Networks research unit, under the supervision of prof. D. Garlaschelli, and started investigating scale invariant aspects of random network models. In 2022 he moved to the Lorentz Institute at Leiden University, where he started a double PhD program in theoretical physics about renormalization on random networks under the co-supervision of prof. S. P. Patil. He also worked closely with the Mathematical Institute, mainly with prof. R. S. Hazra, on the spectral properties of certain scale-invariant network models.

In 2024 he also started a junior research contract at the University of Palermo, under prof. R. Mantegna and prof. V. Latora, on the neuro-scientific aspects of higher-order interactions in coupled dynamical systems, under the framework of the PNRR-funded NeuNet project.

In 2026 he has been a visiting researcher at the Universidade Federal do Ceará, in Fortaleza, Brasil.

He attended several conferences, workshops and schools during his PhD trajectory. Among these, he got the chance to present his work in Italy, France, the Netherlands, India and Brazil.