Московский физико-технический институт (государственный университет)
Сколковский институт науки и технологий

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
«Низкоэнергетические возбуждения в сверхпроводниках вблизи перехода „сверхпроводник-изолятор“»

(магистерская диссертация)

Студент:
Антон Викторович Хвалюк
Научный руководитель:
Михаил Викторович Фейгельман, д. ф.-м. н., проф.

Москва 2020

Moscow Institute of Physics and Technology (State University)
Skolkovo Institute of Science and Technology

MASTER'S THESIS
"Low energy excitations in superconductors close to Superconductor-Insulator Transition"

Student:
Anton Khvalyuk
Scientific advisor:
Mikhail Feigel'man, D. Sc., Prof.

Moscow 2020

Abstract

The present theoretical research addresses the problem of the subgap excitations detected in the electromagnetic response of superconductors with strong microscopical disorder. Such materials have recently attracted much attention as promising materials with a wide variety of practical applications [1], ranging from single-photon detectors to qubits. The conditions for absence of the aforementioned low-energy modes is a particularly important qualitative question for all such applications and has not yet received a proper theoretical description. In the present work we use the Anderson pseudospin model on a random regular graph proposed in [1] to suggest a candidate for a physical mechanism behind the low-energy excitations. The developed theoretical approach provides full access to statistics and spatial structure of the order parameter in a strongly disordered superconductor away from the superconducting transition as well as a comprehensive theoretical model for the structure of phase fluctuations of the order parameter on top of the saddle-point configuration. The results of our analysis are backed by numerical experiments and suggest within a certain range of parameters that the phase fluctuations of the order parameter can indeed be found at arbitrarily low frequencies with the global superconducting order still present in the system.

Contents

Part I

Introduction

1 Motivation: experimental data on low energy excitations in dirty superconductors

1.1 From conventional superconductors...

Conventional superconductors with small concentration of non-magnetic impurities represent a well-studied object [2], both theoretically and experimentally. Among a wide set of known properties of such materials, there is a particularly interesting qualitative set that determines the great importance of conventional superconductors:

- The presence of a transition temperature T_c, characterizing the disappearance of static electrical resistance at $T < T_c$. The corresponding state of the material is called superconducting based on this feature.
- At temperatures well below the transition temperature $T < T_c$ no absorption of electromagnetic radiation is observed at frequencies below a certain threshold frequency

$$\omega_{\text{thr}} = \frac{2\Delta}{\hbar} \sim \frac{4T_c}{\hbar}, \tag{1}$$

  where Δ is a characteristic superconducting energy scale. This observation is an experimental manifestation of a hard gap in the excitation spectrum of a superconductor.
- The energy scale Δ also determines the gap in the local density of states and thus can also be measured as a gap Δ_{run} in a tunneling spectroscopy experiment, which is in perfect agreement with what can be inferred from the electrodynamics response measurements.

Both theoretically [3] and experimentally [4] it was established that small concentrations of non-magnetic impurities do not influence the above properties. While the parameter $k_F l$ remains large—with k_F being the Fermi momentum of a given material and l being the electron mean free path for a given impurity concentration—the presence of impurities merely renormalizes kinetic quantities and some response functions, preserving all thermodynamical properties, such as the structure of the excitations spectrum described above. The previous statement is usually referred to as the Anderson's theorem.

1.2 ...to superconductors close to SIT

However, as evidenced by numerous experimental data [4, 5, 6], upon further increase of level of disorder in the system, the qualitative picture mentioned above suffers noticeable changes:

- Despite being an explicitly thermodynamical quantity, the temperature of superconducting transition acquires an excessively pronounced dependence on the concentration of impurities. 1 presents a snippet of experimental data demonstrating such a behavior.

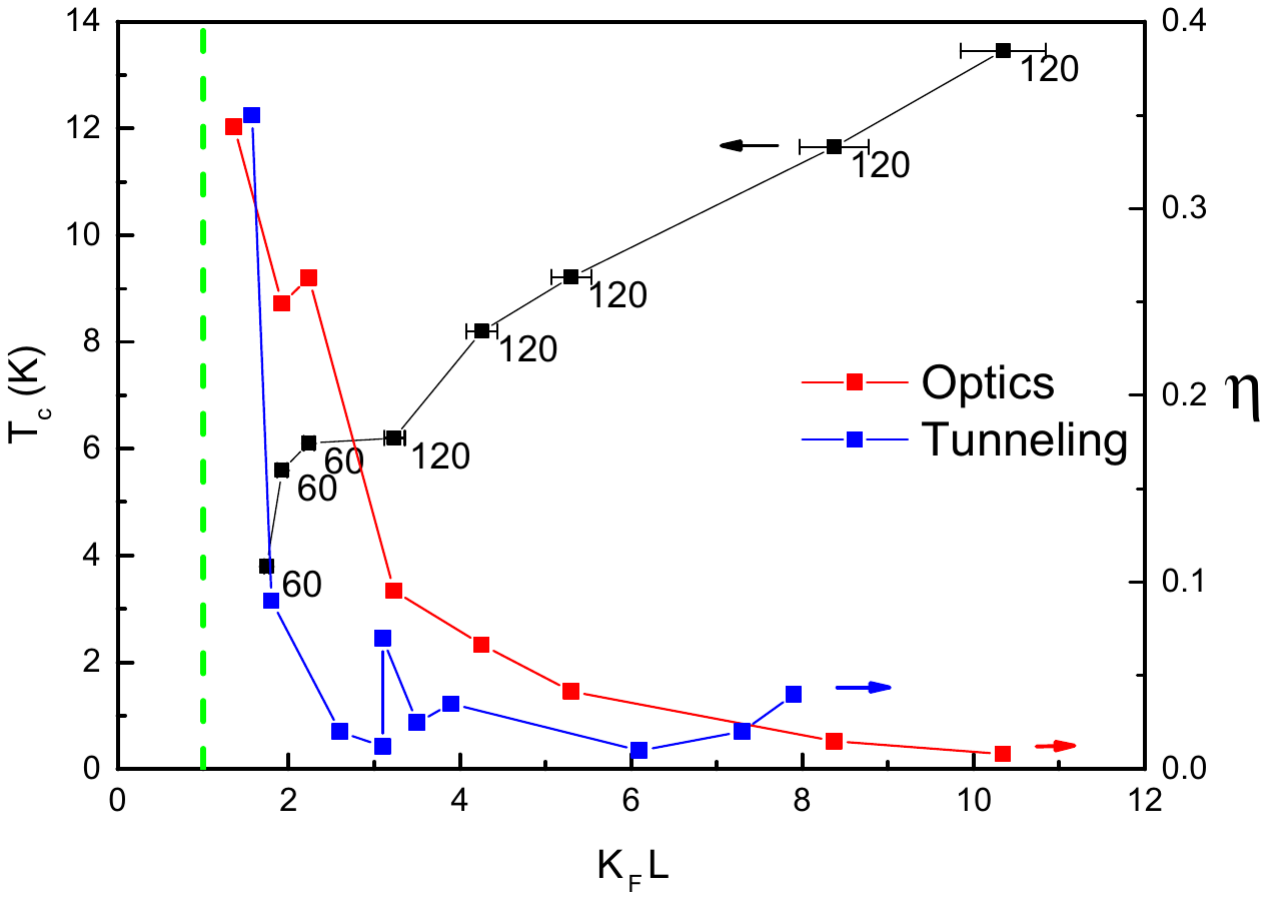


Figure 1: Experimental dependence of the temperature of the superconducting transition in NbN films on non-magnetic disorder, characterized by dimensionless parameter $k_F l$. The value of $k_F l$ was inferred from the room temperate resistance. The data is quoted from [4].

- The materials demonstrate considerable absorption of electromagnetic radiation at frequencies well below the superconducting scale of $2\Delta/\hbar$ as determined by tunneling spectroscopy, i.e, $\hbar\omega_{\rm thr} \ll 2\Delta$. Figure 2 showcases this phenomenon.

It is significant that the effects in question occur in macroscopically homogeneous samples, with the characteristic scales of inhomogeneities being essentially microscopic:

- The work [4] describes disordered NbN films, where inhomogeneity arises from to Nb vacancies in the crystal lattice.

- [6] presents observation of similar phenomena in films made of Aluminum nanogranules. The microscopic inhomogeneities in such a material are explicitly determined by the structure of aluminum granules with dimensions with typical sizes of order 4 nm separated by an oxide film. In this case, the microscopic nature of the inhomogeneities is manifested in the fact that a single granule, due to its small size, has too large level spacing to arrange superconducting order.

The effect of changing the transition temperature T_c in macroscopically homogeneous contaminated samples has already received its theoretical description [7], and is qualitatively explained by

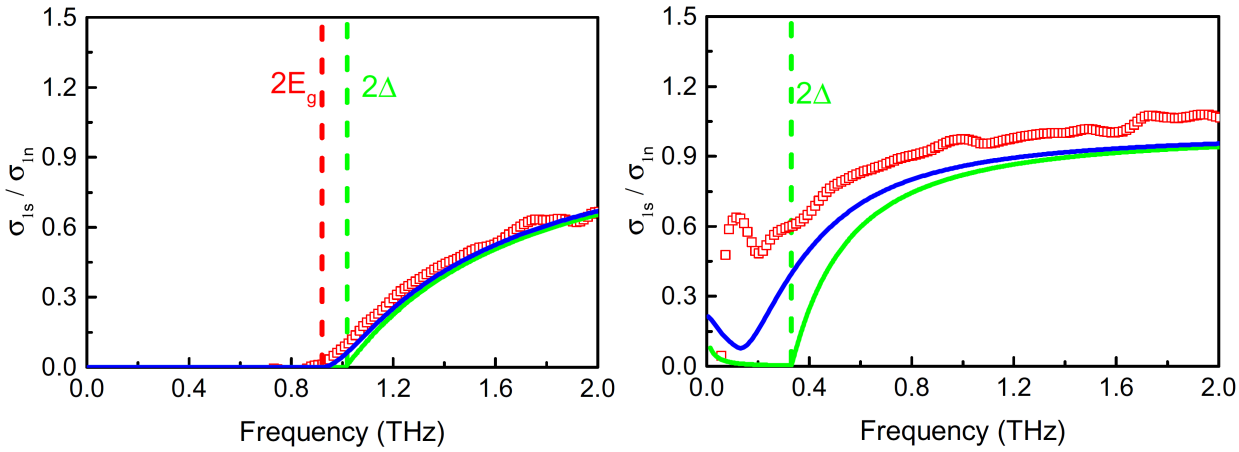


Figure 2: Dependence of the real part of the conductivity on the frequency of external electromagnetic radiation in NbN films. The values of the conductivity are normalized to their values in normal state. Relevant experimental data are presented by red dots, and the vertical green line indicates the position of the superconducting gap Δ_{tun} found by tunneling spectroscopy. Left plot: relatively clean sample with $k_F l \approx 8.37$; right plot: dirty sample with $k_F l \approx 1.74$. In both cases, the value of $k_F l$ was inferred from the room temperature resistance. Data quoted from [4].

the influence of Anderson localization effects and phenomenon of Cooper pair preformation—we will discuss these effects in more detail below. However, the same cannot be said about the presence of low-energy excitations: there only exist some empirical theories that fail to explain the microscopical nature of the observed phenomenon [4, 5].

1.3 Our aims

The ultimate goal of this research is to develop a quantitative theory of superconducting state in strongly disordered superconductors. This comprises 2 main objectives:

- Describe the nontrivial spatial structure of the superconducting order parameter that results from strong microscopical disorder.
- Suggest and demonstrate theoretically a physics mechanism capable of producing low energy excitations in superconducting state.

The theoretical description will be based on the ideas developed in [7], and will also rely on the ideas presented in [1] on the potential origin of the low energy excitations.

2 Structure of the Thesis

The paper is organized as follows:

- In Part II we discuss the general theoretical and empirical grounding of the present research. Section 3 reviews important aspects of phenomenology available from experimental data and previous theoretical findings on the topic and presents the main theoretical model of Anderson pseudospins on a random regular graph to be used throughout the paper.

Section 4 presents a concise review of main equations containing the relevant physics. Finally, Section 5 describes some important mathematical properties of random regular graphs that are going to be extensively used later.

- Part III contains an extensive study of the saddle-point equation, that describes the superconducting order parameter. Section 6 introduces the saddle-point equation and formulates the associated theoretical problem to be solved, discusses some classical limiting cases and previous studies on the topic, and also it describes the numerical approach to the problem of the saddle-point equation. Section 7 reviews some important approximate solutions to the target problem, thus providing the reader with some valuable insights on the structure of the solution. Finally, Section 8 introduces a framework for complete description of the saddle-point problem and provides comparison with numerical study, with excellent agreement in place.

- Part IV is devoted to the immediate topic of this research, as it studies the low-energy fluctuations of the order parameter on top of the saddle-point solution studied in Part III. Within Sections 9-10 the reader can find the formulation of the associated problem and the description of the numerical approach to the problem. Section 11 conducts a study of the simplified version of the problem that was previously investigated in most literature on the topic. In Section 12 the approximate approach of Section 11 is developed into a full theory, that is then compared with numerical experiments.

- Finally, Part V briefly concludes the main achievements and physical insights of the presented research and suggests some key investigative directions for the developed theoretical approach.

Part II
Theoretical model: Anderson pseudospins on a random regular graph

Our main aim in the current part will be to clarify a great deal of simplification separating a truly microscopic model of a superconductor close to superconductor-insulator transition (SIT) and a relatively naive model eventually used in the present study. At the end of the day, there exists only a fairly qualitative set of arguments leading to an analytically tractable instance of the theory. Therefore, below we focus on interpreting the available phenomenological data (both experimental and theoretical) in favor of the proposed simplifications. Nonetheless, to a great extent, the resulting model can only claim qualitative agreement with any experimental verification, and a first-principles model is yet to be developed based on our findings.

In Section 3 the reader can find relevant empirical considerations leading to a simple model of Anderson pseudospins on a random regular graph. The model is, in turn, formulated in Section 4. Within the same section presented there are expressions for standard functional techniques applied to our model. Finally, Section 5 reviews some specific properties of random regular graphs that are of high importance for further analytical consideration.

3 Relevant phenomenology of superconductors close to SIT

The microscopic description of an s-type superconducting state typically involves BCS-like Hamiltonian of the form:

$$H_{\mathrm{BCS}} = \sum_{j\sigma} \xi_j a^\dagger_{j\sigma} a_{j\sigma} - \frac{\lambda}{\nu_0} \sum_{ijkl} V_{ijkl} a^\dagger_{i\uparrow} a^\dagger_{j\downarrow} a_{k\downarrow} a_{l\uparrow}, \tag{2}$$

where indices enumerate eigenstates of a single-particle electron Hamiltonian with energy ξ_j (a generalization of momentum operator eigenstates with energy $p^2/2m - \mu$ in the traditional theory of clean superconductors); $a^\dagger_{j\sigma}$ and $a_{j\sigma}$ are creation and annihilation operators for a fermion in state j and spin σ; ν_0 is the single-particle density of states at the Fermi level; λ is a dimensionless Cooper constant that characterizes the attraction in the Cooper channel; and V_{ijkl} are the matrix elements of the attraction in the eigenbasis of the single-particle Hamiltonian, also typically characterized by a certain high energy cutoff ε_0 modeled as:

$$V_{ijkl} \sim \theta\left(\varepsilon_0 - |\xi_i - \xi_k|\right) \theta\left(\varepsilon_0 - |\xi_j - \xi_l|\right) + \theta\left(\varepsilon_0 - |\xi_j - \xi_k|\right) \theta\left(\varepsilon_0 - |\xi_i - \xi_l|\right), \tag{3}$$

where θ is the Heaviside θ-function (or any other reasonable cutoff function, for that matter). The scale ε_0 plays a role of a generalized Debye frequency that limits the maximum energy exchange. Optionally, one can also take into account Hartree-type terms.

One is then interested in the case when the single-particle Hamiltonian features Anderson localization phenomena, such as localized wave-functions with complex structure and nontrivial level statistics. Although the general understanding of the model (2) for this situations is, to say the least, far from complete, a lot of qualitative and, in some cases, quantitative knowledge is obtained in [7]. In what follows we will rely on some qualitative findings of that paper.

Now we are going to discuss some phenomenological facts that allow us to qualitatively justify a simplified version of the BCS Hamiltonian (2). Unfortunately, there is no rigorous derivation in place, but rather a set of compelling arguments that might later be checked by successive relaxation of the suggested approximations. To a great extent, we rely on a purely qualitative picture extracted from various experimental data on the topic.

3.1 Preformation of Cooper pairs

In a number of studies [8, 9, 10] for the most of the materials of our interest it was observed that prior to an establishment of a global superconducting order one observes formation of local superconducting structures. These structures are not capable of creating global superconducting phase coherence, while, nevertheless, they cause an opening of a hard gap in single-particle spectrum. A snippet of experimental results demonstrating this behavior is illustrated and commented on Figure 3.

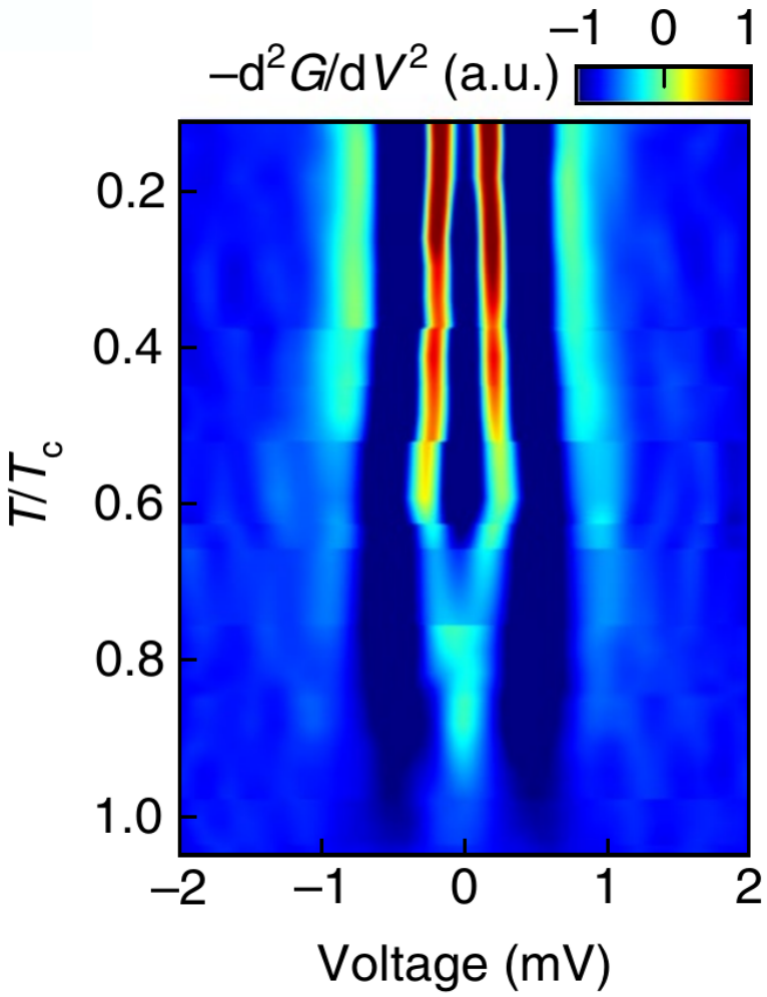


Figure 3: Color map of the second derivative of conductance $-d^2G/dV^2$ (arbitrary units) versus bias voltage V and reduced temperature T/T_c, as obtained by means of Andreev spectroscopy in superconducting films of amorphous indium oxide InO, with T_c extracted from resistance measurements. The plot is quoted from [8]. One can observe two coherence peaks corresponding to two pairs of vertical bright lines. The narrow set of lines merges at $T = T_c$, indicating the superconducting transition, while the broad one persists even at higher temperatures, albeit experiencing smearing.

The phenomenon is now referred to as "Cooper pair preformation" and has already obtained theoretical explanation [7, sec. 2.3]. For the current research it is only important that even at temperatures well above the temperature of the true superconducting transition T_c electrons are already paired, and the typical energy Δ_p required to break these pairs greatly exceeds any typical superconducting scales.

Specifically for our model it is also safe to say [7, sec. 4.3] that each single-particle state is either populated with two electrons with opposite spins or contains no electrons at all. This can be expressed as simplifications to the structure of the matrix elements of the Cooper interaction, where we can now keep only the terms diagonal in both pairs of indices:

$$V_{ijkl} = V_{iikk} \cdot \delta_{ij}\delta_{kl} \tag{4}$$

3.2 Single-particle localization of electron states

For most of the samples that demonstrate the presence of low-energy excitations we are interested in, it is also true that at temperatures above the superconducting transition threshold there exists insulating activation behavior of the resistance $R(T) \propto \exp\{T/T_0\}$, which only restores to conventional variable-range hopping behavior at higher temperatures [7, sec. 1.2 and ref. therein]. Together with other experimental facts, such as negative magnetoresistance, this serves as compelling evidence for well developed Anderson localization in those materials. Out of all the effects known to take place in Anderson insulators we will exploit only several qualitative aspects that allow us to simplify the model.

3.2.1 Large localization volume

Anderson localization implies that at the relevant energy band close to Fermi level every eigenfunction of the single-particle Hamiltonian is substantially present only within a finite localization volume V_{loc} concentrated around a particular point in space, as opposed to the regime of diffusive transport when eigenfunctions are extended over the whole system. Conversely, the very presence of superconductivity for a given material implies that the level spacing δ_L induced by a finite localization volume is much smaller than the typical scale ε_0 of the Cooper attraction. Because $\delta_L \sim V_{\text{loc}}^{-D}$,where D is the dimensionality of the problem, this consideration bounds the localization volume from below [7, sec. 1.1.3]. Therefore one might expect a finite range of parameters where the localization volume is ample to provide substantial overlap with a large number of other localized eigenfunctions, i.e., $nV_{\text{loc}} \gg 1$, where n is the electron concentration close to the Fermi level. We underline, nevertheless, that such a condition is mostly a compelling assumption.

3.2.2 Crude model for matrix elements of Cooper interaction

As a yet another manifestation of the complex physics of Anderson localization, single-particle eigenfunctions are known to demonstrate a nontrivial distribution of amplitudes in real space. In this work we are not going to dive into this rather complicated topic and instead employ a fairly crude approximation for the matrix elements of a local interaction (which conventional Cooper attraction surely is). We will assume that such a perturbation arranges a constant matrix element from a given eigenstate to a finite fraction $k \in [0, 1]$ of all other $\sim nV_{\text{loc.}} \gg 1$ eigenstates substantially present within the localization volume V_{loc}. To simplify things even further, we will assume that the total number of eigenstates to which a given eigenstate can transit is identical for all eigenstates and equal to some large number $K = knV_{\text{loc}} \gg 1$.

An attentive reader can immediately notice that such an approach is completely ignorant of the high energy cutoff ε_0 of the Cooper channel attraction. Additionally, the introduced number K fluctuates across eigenvalues and its statistics and correlations are a subject for a separate study with abundant phenomenology, as it was shown in [7]. Unfortunately these are the approximations that we have to adopt in order to make the resulting model tractable analytically. One qualitative argument that suggests qualitative validity of our approach is the following: the choice of a particular structure of the Cooper channel attraction governs the magnitude of superconducting scales, and once they are set, all higher energy scales are expected to become completely irrelevant. As a result, altering the approximation of the matrix elements in (2) leads only to a different value of the overall superconducting scale, while preserving all conclusions for smaller energy scales. Within our analysis we observe that this is indeed the

case: the mathematical framework appears to be completely insensitive to all such details once the superconducting scale is set.

The full situation is sketched on Figure 4.

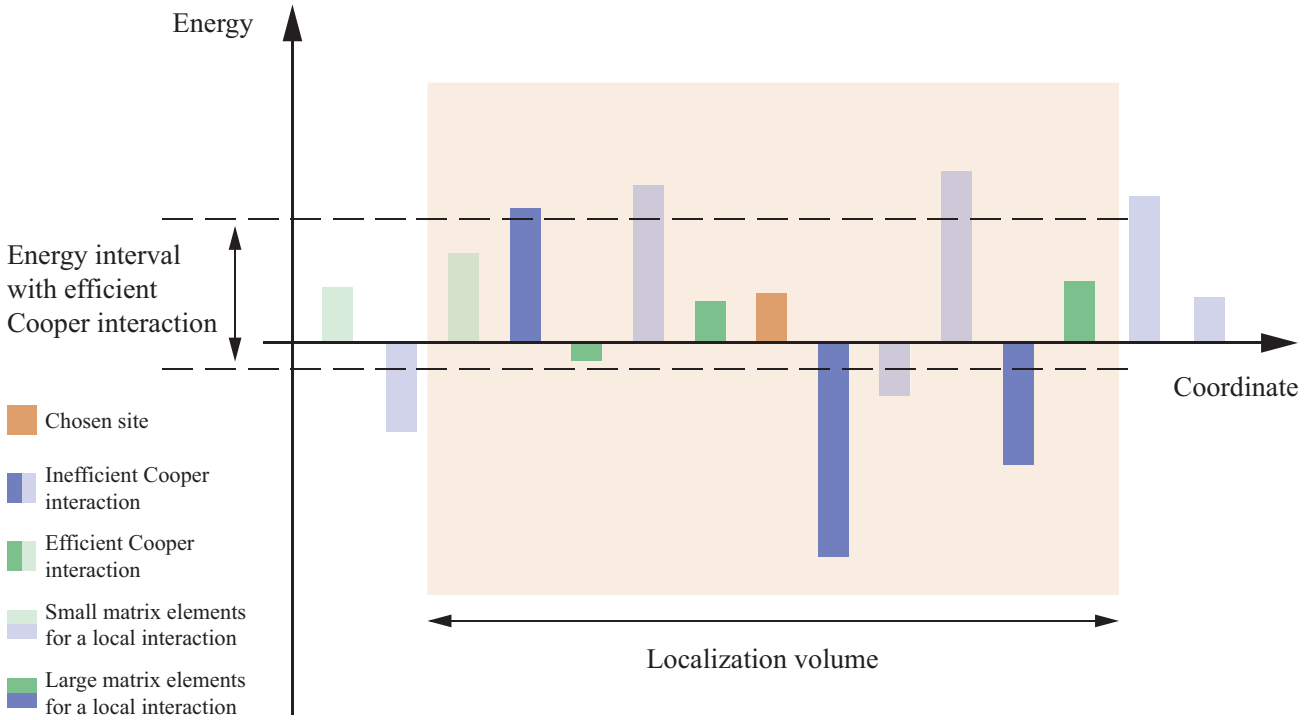


Figure 4: A chart for simplified matrix elements of the Cooper attractive interaction, illustrated in a 1D setting for clarity. Sites are arranged in a chain and are denoted by bars illustrating the onsite energy level ξ_j for each site. Let us consider a particular site highlighted in orange. The wave function of the eigenstate corresponding to this site occupies a large number of neighboring sites enclosed in the localization volume (orange rectangle). However, due to a fractal structure of the single-particle eigenfunction within the localization volume, a local interaction can only arrange hops to a random subset of all neighbors, that are denoted by opaque filling, while interaction with other sites (transparent filling) is weak. Finally, Cooper interaction itself is only efficient within a finite interval of width $2\varepsilon_0$ of energies relative to the onsite energy. Strictly speaking, only the green bars that fall within the corresponding energy band (denoted by dashed lines) can interact with the chosen orange site. The total neighbors that interact with the given site are opaque green. Yet in our model we neglect the presence of a finite Cooper interaction cutoff, so that in our model the orange site interacts with all neighbors with opaque filling, regardless of color.

3.3 Irrelevance of Coulomb repulsion

The last bit of phenomenology to be discussed is the role of Coulomb repulsion in the formation of superconducting behavior in our problem. Essentially, our further consideration completely ignores Coulomb repulsion occurring between electrons in the system. There are several qualitative arguments in favor of such an approximation:

- Most of the materials in question possess a relatively large dielectric constant [7], which suggests strong screening of Coulomb repulsion by electrons deep under the Fermi surface.

- In addition to a large dielectric constant, the very existence of SC in materials with such a large resistance is indicative of some intrinsic mechanism that suppresses Coulomb repulsion. Indeed, for a thick film of a regular disordered metal, large sheet resistance would imply such a strong disorder that the resulting effect of Coulomb repulsion would inevitably have driven the temperature of the superconducting transition down to zero. Yet, the experimentally detected superconductivity in disordered thick films suggests that the Coulomb repulsion is somehow suppressed at short scales, thus allowing the superconducting order to be established[1].

- Finally, strong inhomogeneity of the order parameter induced by the fractal structure of localized electron wavefunctions is expected to weaken the Coulomb interaction at large scales. Although no rigorous arguments can be presented to support this claim, such a simplification seems to be a reasonable initial approximation.

As a result, we will henceforth ignore the existence of Coulomb interaction between electrons.

4 Hamiltonian of Anderson pseudospins on a random regular graph

Based on a more comprehensive consideration of the phenomenology presented above, in [12] it was proposed to model the low-energy physics of highly disordered superconductors with a simple Hamiltonian that is assumed to capture main qualitative physics at low energies.

4.1 Hamiltonian

Guided by the phenomenology of the previous section, we adopt the following qualitative simplifications to the model:

- At low temperatures of interest, electrons are all paired in Cooper pairs with zero total spin, so that the dynamics of the system consist of Cooper pair hops, while the process of Cooper pair breaking is assumed to require energies much larger than typical superconducting energy scales, and thus it is neglected.

- Cooper attraction between electrons is described by the interaction between the eigenstates of the single-particle Hamiltonian and characterized by a binary distribution of matrix elements: between two selected eigenstates the attraction either does not occur at all or has a constant matrix element that we denote by g. In what follows we traditionally assume the Cooper attraction to be a weak process, i.e., $g \ll 1$. Additionally, we note that the energy cutoff in Cooper channels is downright ignored. We assume that taking into account energy dependence of Cooper attraction will only alter the overall energy scale of superconducting correlations, while preserving major qualitative results. Our further analysis is supportive of that claim.

[1]For instance, [10] uses disordered InO films with the thickness of 30 nanometers, which is equivalent to approximately 10^2 atomic layers. For all intents and purposes this can be considered to be a fully 3D system. On the other hand, the sheet resistance of these films is measured to be in the range of several $k\Omega$ per square, which evaluates to 3D resistivity of order $\rho \sim 10^1\, m\Omega \cdot \mathrm{cm}$, being well beyond the range of $\sim 10^0\, m\Omega \cdot \mathrm{cm}$ where the superconductivity is still observed for typical disordered metals [11, see Fig. 1].

- Energies ξ_j of free Cooper pairs are independent random variables with a broad distribution of typical width W which is assumed to be larger than all superconducting scales in the problem. The scale W sets the typical value of the density of states in normal state at the Fermi level, so that it is quite natural to take it as a default energy unit.

- Each Cooper pair is able to hop to a fixed large number of states $Z = K + 1 \gg 1$, all of which are chosen independently at random. In other words, the hopping matrix represents an instance of random $K+1$-regular graph with N vertices, with each vertex having exactly $K + 1$ neighbors. Strictly speaking, such a design choice is fairly artificial and is applied only to render the resulting model solvable. In particular, our analytical solution relies on the absence of short loops in random regular graphs, as discussed later in Section 5, while the real system is not by any means protected from presence of short loops. However, the analysis we will present as the main part of this work suggests that upon fixing the typical superconducting energy scales the exact structure of the matrix elements is mostly irrelevant, so that the proposed simple approximation is likely to reproduce the actual physics.

- Finally, Coulomb repulsion is completely neglected. As already discussed before, the validity of such an approximation is an open question, yet it seems to be an appropriate design choice to start with.

As a result, the Hamiltonian reads:

$$H_{\text{model}} = \sum_j \xi_j \cdot \sum_\sigma a^\dagger_{j\sigma} a_{j\sigma} - \frac{2g}{K} \sum_{\langle i,j \rangle} \left(a^\dagger_{i\uparrow} a^\dagger_{i\downarrow} a_{j\downarrow} a_{j\uparrow} + \text{h.c.} \right), \tag{5}$$

where i enumerates vertices of a random $K + 1$-regular graph with a thermodynamically large total number of vertices N; ξ_j are random energies of the Cooper pair states, distributed independently with a some distribution of characteristic width $W \sim 1$; g plays a role of the dimensionless Cooper constant; and the summation in the second term is performed over all pairs $\langle i, j \rangle$ that are assumed to be connected by interaction, i.e., in the given realization of a random regular graph.

It is also convenient to rewrite the Hamiltonian in terms of Anderson pseudospins [13]:

$$2s^z_j = \sum_\sigma a^\dagger_j a_j - 1, \quad s^+_j = s^x + is^y := a^\dagger_{j\uparrow} a^\dagger_{j\downarrow}, \quad s^-_j = s^x - is^y = a_{j\downarrow} a_{j\uparrow}, \tag{6}$$

which then appear to form a proper spin 1/2 algebra:

$$\left[s^\alpha_i, s^\beta_j \right] = \delta_{ij} \cdot \varepsilon^{\alpha\beta\gamma} s^\gamma_j, \quad s^\alpha_i s^\alpha_i = 3/4, \tag{7}$$

where $\alpha, \beta, \gamma \in \{x, y, z\}$ with Einstein summation convention for Greek indices implied, and $\varepsilon^{\alpha\beta\gamma}$ is the fully antisymmetric tensor. In these terms, the Hamiltonian (5) can be rewritten as the so called pseudospin Hamiltonian [13]:

$$H_{\text{PS}} = \sum_j 2\xi_j \cdot s^z_j - \frac{2g}{K} \sum_{\langle i,j \rangle} \left(s^+_i s^-_j + \text{h.c.} \right) \tag{8}$$

At this point it is worth mentioning some prior works within the model (5) or its equivalent (8). The work [12] presents a comprehensive study of the mean-field approximation as well

as its refined version. Moving closer to the topic of this research, [14] uses the model (8) to describe paraconductivity contributions to the insulating phase of the model (i.e. above the superconducting transition), and [15] discusses ultrasound attenuation caused by order parameter fluctuations. Finally, the paper [1] draws some predictions on the spectrum of transverse fluctuations of the order parameter, which is the topic of this research. The aforementioned predictions are discussed in the corresponding Part IV.

4.2 Representation in terms of a functional integral

The field theory approach is typically used to establish connections to the conventional phenomenology of superconductors [16, ch. 1]. However, our case presents a complication in a form of preformed Cooper pairs, thus requiring a deviation from the standard approach in terms of a fermionic functional integral [17]. Namely, we are going to use the semionic functional integral [18] as a way to deal with nontrivial pseudospin algebra (7). A concise introduction to this approach applied to the pseudospin Hamiltonian (8) is presented below, while a more detailed review is done in the previous iteration of this research [19].

4.2.1 Semionic functional integral

The semionic approach effectively establishes correspondence between thermal equilibrium partition functions and generating functionals for the pseudospin Hamiltonian (8) and the following fermionic Hamiltonian with imaginary chemical potential:

$$H_{\text{sem}} = \sum_j 2\xi_j \cdot c^\dagger_{j\alpha} s^z_{\alpha\beta} c_{j\beta} - \frac{2g}{K} \sum_{\langle i,j \rangle} \left(s^+_{\alpha\beta} s^-_{\rho\sigma} \cdot c^\dagger_{i\alpha} c_{i\beta} c^\dagger_{j\rho} c_{j\sigma} + \text{h.c.} \right) \tag{9}$$

$$\beta\mu = -\frac{i\pi}{2} \tag{10}$$

where $c^\dagger, c$ are fermionic creation and annihilation operators that operate on a set of single-particle states of true electrons, but endowed with an additional spinor structure with two components enumerated by Greek indices $\alpha \in \{1, 2\}$, and spin 1/2 matrices s^μ now operate on this auxiliary vector space. For convenience we will further denote

$$J_{ij} := \frac{2g}{K} A_{ij}, \tag{11}$$

where A_{ij} is the adjacency matrix of the base random regular graph. This also allows us to rewrite the sums of symmetric functions over edges as:

$$\frac{2g}{K} \sum_{\langle i,j \rangle} \left[f\left(i,j\right) + f\left(j,i\right) \right] = \sum_{i,j} J_{ij} \cdot f\left(i,j\right) \tag{12}$$

The aforementioned imaginary chemical potential is understood as a formal weight in the grand partition function:

$$\mathcal{Z} = \text{Tr} \left\{ \exp\left\{ -\beta \left(H_{\text{sem}} + \mu N_{\text{sem}} \right) \right\} \right\}, \tag{13}$$

$$N_{\text{sem}} = \sum_{j\alpha} c^\dagger_{i\alpha} c_{j\alpha}, \tag{14}$$

where the trace is now taken over all states of usual spinless fermions. The key statement is that all thermodynamical averages with the original pseudospin Hamiltonian (8) are obtained by formal replacement of each spin operator with its semionic representation:

$$s_i^\mu \to c_{i\alpha}^\dagger s_{\alpha\beta}^\mu c_{i\beta} \tag{15}$$

and then evaluating the resulting partition functions as those for usual spinless fermions with Hamiltonian (9) and chemical potential (10). In particular, the latter will reveal itself in the behavior of the Matsubara frequencies should one employ the corresponding representation [18]:

$$\omega_n = \pi T \left(2n + 1 - \frac{1}{2} \right) \tag{16}$$

As a result, one can use the standard fermionic functional integral and employ the whole variety of field theory methods. A self-contained discussion of fermionic functional integration technique can be found in [17, ch. 4], while here we will only present the results for the partition function in terms of an imaginary time functional integral:

$$S\left[\overline{\psi}, \psi\right] = \int_0^\beta d\tau \cdot \sum_i \overline{\psi}_i^\alpha \left(\partial_\tau - \mu\right) \psi_i^\alpha - H_{\text{sem}}\left(\overline{\psi}, \psi\right), \tag{17}$$

$$H_{\text{sem}}\left(\overline{\psi}, \psi\right) = \sum_j 2\xi_j \cdot \overline{\psi}_j^\alpha s_{\alpha\beta}^z \psi_j^\beta - \frac{1}{4} \sum_{ij} J_{ij} \left(s_{\alpha\beta}^x s_{\rho\sigma}^x + s_{\alpha\beta}^y s_{\rho\sigma}^y \right) \overline{\psi}_i^\alpha \psi_i^\beta \overline{\psi}_j^\rho \psi_j^\sigma, \tag{18}$$

$$\begin{aligned} \mathcal{Z} = \int & D\left[\overline{\psi}, \psi\right] \quad \exp\left\{-S\left[\overline{\psi}, \psi\right]\right\}, \\ & \psi(0) = -\psi(\beta) \\ & \overline{\psi}(0) = -\overline{\psi}(\beta) \end{aligned} \tag{19}$$

where $\tau \in [0, \beta]$ is the imaginary time; S is the action of the functional integration, and the integration is done over standard fermionic fields, i.e., over all τ-dependent Grassman anti-commuting fields $\overline{\psi}, \psi$ with odd boundary conditions. Each field has two components with upper indices $\alpha \in \{1, 2\}$ on each site i of the graph, with spin matrices $s^\alpha, \alpha \in \{x, y, z\}$ acting in the space of these components. All sorts of observables are then obtained either by explicitly plugging the corresponding function of ψ fields in the pre-exponent of (19), or by method of generating functionals, i.e., by adding external source terms to the action (17) and the subsequent differentiating w.r.t source fields. For further details, one can consult [17, ch. 5] or [19, sec. 2.2].

4.2.2 Functional integral for the order parameter field

Moving further with the functional integral formalism, one can decouple the interaction term in (19) by means of Hubbard-Stratanovich integral and then switch to the functional integral of the order parameter field. Without going into further detail, otherwise presented in [19, sec. 2.2], the result then reads

$$S\left[\Delta\right] = \int_0^\beta d\tau \cdot \sum_{ij} J_{ij}^{-1} \Delta_i^\alpha(\tau) \Delta_j^\beta(\tau) + S_{\text{sem}}\left[\Delta\right], \tag{20}$$

$$
\begin{aligned}
&\exp\left\{-S_{\text{sem}}\left[\Delta\right]\right\} \\
&= \int\limits_{\substack{\psi(0)=-\psi(\beta) \\ \overline{\psi}(0)=-\overline{\psi}(\beta)}} D\left[\overline{\psi},\psi\right] \quad \exp\left\{-\int\limits_0^\beta d\tau\cdot\left[\sum_i \overline{\psi}_i^\alpha\left(\partial_\tau-\mu\right)\psi_i^\alpha-\sum_j \overline{\psi}_j^\alpha\left(2\xi_j s_{\alpha\beta}^z-2\Delta_j^\gamma\left(\tau\right)s_{\alpha\beta}^\gamma\right)\psi_j^\beta\right]\right\} \\
&\equiv \operatorname{Tr}\left[T_\tau\exp\left\{-\beta\sum_j\left(2\xi_j s^z-2\Delta_j^\gamma\left(\tau\right)\cdot s^\gamma\right)\right\}\right], \qquad (21)
\end{aligned}
$$

$$
\mathcal{Z}=\int\limits_{\Delta(0)=\Delta(\beta)} D\Delta\exp\left\{-S\left[\Delta\right]\right\}, \qquad (22)
$$

where the field Δ is now a bosonic field with symmetric boundary conditions and two components indexed by $\gamma\in\{x,y\}$, so that the summation over γ in (21) is done only over $\gamma=x,y$. Note that this time the indices γ are contracted with the upper indices of spin operators, i.e., Δ is not acted upon by spin operators, but rather serves as a conjugated field. The term S_{sem} in the bosonic action (20) results from integrating out the semionic degrees of freedom, as expressed in the second line of (21). It can also be written as the time-ordered exponent $T_\tau\exp$ for the original spin variables interacting now only with the dynamic order parameter field, as done in the last line of (21). There is also a technical issue with the inverse of J which we will discuss in Section 5.

Starting from this functional representation, one can derive all classical results for superconductivity and draw corresponding connections to textbook physics [16, ch. 1]. For instance, the vector field Δ is nothing but the order parameter field with components representing real and imaginary part, respectively. The saddle-point of the functional integral (22) corresponds to the self-consistency equation. The value of the semionic Green function is given by a counterpart of Gor'kov's equations. By expanding quadratic fluctuations of the order parameter field around its saddle-point value one can analyze fluctuations corrections. Adding quadratic terms in expansion of action to the analysis, one can derive the Ginzburg-Landau functional for the problem.

4.2.3 Manifestations of spin algebra in the symmetries of the model

The functional integral (22) carries all information about the original spin algebra (7) in the form of a gauge transformation. Alternatively, one can think of the gauge transformation of the integration fields Δ that can be compensated either by the integration measure or by redefinition of various objects in the theory. Below we briefly review the structure of these transformations and discuss their physical meaning. A derivation of these transformations is a more of a technical task and will be presented elsewhere.

- Commutation relations $[s^\pm,s^z]=\mp s^\pm$ lead to phase gauge transformations of the J matrix:

$$
J_{ij}^{-1}\mapsto\exp\left\{+2i\sigma^y\alpha_i\right\}J_{ij}^{-1}\exp\left\{-2i\sigma^y\alpha_j\right\}\Longleftrightarrow J_{ij}\mapsto\exp\left\{+2i\sigma^y\alpha_i\right\}J_{ij}\exp\left\{-2i\sigma^y\alpha_j\right\}, \qquad (23)
$$

  where α is the scalar gauge field parameterizing the transformation. The presence of σ^y matrix in the definition of the transformation means that the resulting matrix starts to act nontrivially in the space of the order parameter.

- Adding imaginary time dependence to the generating field α allows one to carry out gauge transforms of ξ fields as well:
$$\xi_i \mapsto \xi_i - 2i\alpha_i'(\tau) \tag{24}$$
- Because we are dealing with spin 1/2 operators, each of the 3 identities $[s^\alpha]^2 = 1/4, \alpha \in \{x, y, z\}$ reflect themselves as a yet another gauge transformation of the J matrix:
$$J_{ij} \mapsto J_{ij} + (\phi_i \cdot 1 + \eta_i \sigma^z + i\theta_i \sigma^y)\,\delta_{ij}, \tag{25}$$
where θ, ϕ, η are the 3 scalar fields generating the transformation, and matrices $1, \sigma^z, \sigma^y$ act in the subspace of order parameter components. This transformation leads to a change in action
$$\delta S = \frac{1}{2}\int d\tau \cdot \sum_i \phi_i, \tag{26}$$
which does not contain dynamics fields being integrated over and thus is absent in any physical answers.

There exist nontrivial commutation relations between these symmetries, that are intimately connected with the structure of the original spin algebra (7), but this is a topic for a separate discussion. The discussed symmetries pave the way for deriving physical observables, such as current.

4.3 Description of the current-current response function

We are now in position to discuss how our somewhat artificial model maps to the physical world.

4.3.1 Lost information about the real space structure

One immediate consequence of our simplification is that the base random regular graph that governs all topology of the model bears no information about structure or mutual disposition of the electron states in real space. Therefore, any result calculated within the simple Hamiltonian 8 should be supplied with some information about the structure in real space. This presents a qualitative complication to the interpretation of the analytical results of this model, and we are yet to come up with a proper scheme for calculating real space quantities.

4.3.2 Nontrivial physical observables: current operators

As it always happens in field theory, purely from phase gauge symmetry one can infer the full form of the current and charge operators, that are perfectly consistent with what one would intuitively expect from the model. The complete derivation of the current and charge operators will be presented elsewhere, while here we present a physics way of understanding the key properties of the current operator.

Within our model, the process of charge transfer is realized by hops of Cooper pairs from a given site to one of its nearest neighbors. One can then easily restore the charge operator
$$Q_i = -e \cdot \sum_\sigma a_{i\sigma}^\dagger a_{i\sigma} \mapsto -e \cdot (1 + 2s_i^z) \tag{27}$$
where the second expression represents the corresponding operator in pseudospin representation. The current flow is characterized by the direction of hopping and the rate at which such processes

occur along each oriented edge of the base graph. The orientation is required, as one obviously expects the current $I_{i\to j}$ from site i to a neighboring site j to be equal to the opposite of current $I_{j\to i}$ from j to i, i.e.,

$$I_{i\to j} = -I_{j\to i} \tag{28}$$

Therefore, the current in the system is described by a function of the oriented edges of the graph, i.e., it assigns a value to each oriented edge of the graph and has the antisymmetry property (28). Alternatively, current can be regarded as an antisymmetric matrix in the site space that has nonzero entries only when the sites i and j are connected by an edge, i.e.,

$$I_{i\to j} \propto \pm A_{ij} \tag{29}$$

where A is the adjacency matrix of the graph.

This answer hints at the analogy of traditional vector fields: what was a geometrical 1-form in standard 3D electrodynamics becomes an antisymmetric matrix in a discrete setup. In particular, this is true about the vector potential $\mathcal{A}$ in the system, which in our problem also takes the form of an antisymmetric matrix. It can be shown in a more rigorous manner, that the electrodynamic potentials $\mathcal{A}$ and Φ (scalar electrostatic potential, which is now represented by a scalar function on the graph vertices) can be introduced to the problem in a standard way, safe for the imaginary time formulation:

$$\delta S\left[\delta\mathcal{A}, \delta\Phi\right] = \int d\tau \cdot \frac{1}{2}\sum_i \sum_{j\in\partial i} \delta\mathcal{A}_{i\to j} \cdot I_{i\to j} - \int d\tau \cdot \sum_i \delta\Phi_i Q_i \tag{30}$$

Moving further, one can show that the gauge transformations are intimately related to the current flow and allow one to restore the form of the current operator:

$$I_{i\to j} = (-2e)\cdot 2m_i^\alpha \cdot \left(i\sigma^y\right)_{\alpha\beta} J_{ij} \cdot m_j^\beta, \;\; m^\alpha = \sum_n \left(J^{-1}\right)_{in} \Delta_n^\alpha \tag{31}$$

written here for the case without external vector potential. Note that it reproduces both the antisymmetry and locality properties mentioned before, in return for being heavily non-local w.r.t the order parameter, as the latter enter the expression via a field m involving nonlocal transformation J^{-1}. A special case of this equation when J is assumed to retain some residual structure of the real space is discussed in [14].

Going from here, one can derive all standard physical objects, such as the current-current response function describing the conductivity tensor, which is now going to represent a matrix in the space of all edges of the graph, thus once again demonstrating the enormous complications arising from our ignorance of real space structure. The resulting expressions are quite cumbersome, while here we only discuss the qualitative structure that is important for understanding the results of our research from the physical point of view.

Note also that one still has to restore real space geometry in an independent way, as already discussed before. For instance, out of $K \gg 1$ components of current I_{ij} on each site there are only $D \sim 1$ independent ones (D is the dimensionality of the initial problem).

The expression for the current allows one to calculate physically observable current-current response function χ, which in this system described the dynamic response of current along a directed edge $i \to j$ to application of vector-potential along a directed edge $m \to n$:

$$\begin{aligned}
&\chi\left(i\to j, \tau | m\to n, \tau'\right) := \left\langle \mathrm{T}_\tau \left\{ I_{i\to j}\left(\tau\right) \cdot I_{m\to n}\left(\tau'\right)\right\}\right\rangle \\
&= (-2e)^2 \cdot \left[2i\sigma^y J\right]_{ij}^{\alpha\beta} \left[2i\sigma^y J\right]_{mn}^{\mu\nu} \cdot \left\langle T_\tau \left\{ m_i^\alpha\left(\tau\right) m_j^\beta\left(\tau+0\right) m_m^\mu\left(\tau'\right) m_n^\nu\left(\tau'+0\right)\right\}\right\rangle,
\end{aligned} \tag{32}$$

where T_τ stands for imaginary time ordering, and m fields are the fields introduced in (31).

Outside the region of strong dynamical fluctuations of the order parameter, one can expand the four-point correlator using the Wick's theorem and obtain that the time-dependence of the response function is expressed via that of the 2-point correlator of fluctuations of the order parameter:

$$G_{ij}^{\alpha\beta}(\tau,\tau') := \left\langle\left\langle \Delta_i^\alpha(\tau)\cdot\Delta_j^\beta(\tau')\right\rangle\right\rangle \tag{33}$$

The full expression for χ after applying the Wick's theorem contains two types of contributions:

- product of two correlators (33), that take the form of a convolution over Matsubara frequencies in Fourier space:

$$\chi_1(i\omega_n) \propto [\sigma^y]^{\alpha\beta}[i\sigma^y]^{\mu\nu}\cdot T\sum_{\varepsilon_m}\left[G^{\alpha\mu}(i\omega_n+i\varepsilon_m)\,G^{\beta\nu}(i\varepsilon_m)\right], \tag{34}$$

  where we omit space indices and the corresponding nonlocal prefactor for brevity. This contribution is showed to produce paraconductivity in normal state [14] above the superconducting transition temperature.

- a term linear in the correlators of the fluctuations and the average (time-independent) value of the order parameter field:

$$\chi_2(i\omega_n) \propto [i\sigma^y]^{\alpha\beta}[i\sigma^y]^{\mu\nu}\cdot\langle\Delta^\alpha\rangle\langle\Delta^\mu\rangle\,G^{\beta\nu}(i\omega_n), \tag{35}$$

  where we have again dropped the spatial indices (reflecting them inflates to 4 terms of the form above). The most important observation is that only the fluctuations perpendicular to the direction of the order parameter contribute to the current response, which is secured by the $i\sigma^y$ term.

From this result we can infer that the frequency dependence of the response function is completely governed by the fluctuations of the order parameter. In particular, real part of physical conductivity at a given frequency ω, responsible for dissipation, is nonzero when the imaginary part of the corresponding matrix element of G is nonzero, which, in turn is only possible when the correlator G has eigenmodes on the given frequency. This understanding is the main qualitative motivation to analyze the structure of transverse fluctuations of the order parameter.

Although we are not even anywhere near a position to claim this as a quantitative correspondence between physical quantities, our analysis indicates that the correlator of transverse fluctuations does indeed contain excitations with energy significantly lower than the superconducting gap, which then might find their way to show up in the electromagnetic response measurements.

5 Properties of Random Regular Graphs

In this section we will review some of the most important properties of random regular graphs. Later we will find these properties to be of great use on our way to analytical results.

5.1 Local tree-like structure and translational invariance

The first property to discuss is the celebrated local tree-like structure of the graph. Informally, it can be stated as following:

> Physical Principle. Any quantity characterized as containing a finite spatial scale d cannot discriminate between a random $K+1$-regular graph of size $N \gg K^{d+1}$ and an infinite Bethe lattice, i.e., a tree graph with constant number of descendants equal to $K+1$ for the root vertex and K for all other sites.

The phrase "containing a finite spatial scale" describes properties that due to some physical mechanisms are only sensitive to a finite local subset of the graph, i.e., are characterized by effects that decay with distance fast enough. To this class one can attribute the order parameter field itself, as it will be demonstrated later, as well as practically any dynamic quantities probed at a finite frequency that sets the decay scale [20].

Speaking formally, the following theorem holds [21]:

> Theorem. Let $R(N,K)$ be the equiprobable ensemble of all random $K+1$-regular graphs on N vertices and let $n_l(G)$ be the number of cycles of length $l \in \mathbb{N}, l \geq 2$ in a given graph G. Then $n_l(G)$ has a well defined average over ensemble $R(N,K)$ with the following behavior in the thermodynamical limit:

$$\langle n_l(G) \rangle_{G \in R(N,K)} := \frac{1}{|R(N,K)|} \sum_{G \in R(N,K)} n_l(G) = \frac{K^l}{2l} \left(1 + o(1)\right), N \to \infty \qquad (36)$$

The Physical Principle stated above then follows from this statement by noting that the total number of sites containing all cycles of length l is at most $l \cdot n_l$ and thus approaches a finite value as N approaches infinity. As a result, the probability of a given site—or a finite subset of sites, for that matter—to be present in a cycle of any given length eventually vanishes in the thermodynamical limit. Thus any quantity that does not track information beyond a finite distance on a graph will eventually observe only a loopless fragment of Bethe lattice.

Regardless of the formulation one prefers, the idea of local tree-like structure also provides a natural scale of finite size effects in the system resting on presence of large cycles. Namely, one expects the size of the graph to influence the physical observables when it becomes comparable to $N_f \sim K^{2d}$, where d is the typical range of correlations of the quantity in question. This result can be demonstrated by the following qualitative argument. Assuming the local tree-like structure to take place, one conducts breadth-first search[2] on a graph starting from a given site i_0. As long as search depth remains shallow, the total number of discovered sites is negligible compared to the system size, so that the probability to hit an already discovered site is negligible, because neighbors are chosen almost at random at each step. Therefore, each new level d of depth will initially reveal a new set of sites of total count $(K+1) \cdot K^{d-1}$ per level. However, the total number of discovered sites grows exponentially with depth d, so that at a depth of order $d_N \sim \ln N / \ln K$ one will inevitably cover all sites of the graph, meaning that there is a path from the starting site to any other site with length $\sim d_N$. Repeating the breadth-first search from any of the sites discovered at the latest stages is likely to provide a yet another path to the original site i_0 of the same length $\sim d_N$. Therefore, after this procedure one is left with a

[2]Breadth-first search is a process of traversing vertices of a graph. It starts from a chosen node and explores all neighbors in the "closest first" order.

cycle of total length $\sim 2d_N$ with probability close to 1. Albeit being nowhere near rigorous, this consideration is still quite useful to understand the results that are about to be presented below.

Another important property that should be mentioned as a consequence of tree-like structure is the translational invariance. This property manifests itself in the independence of local physical properties from the absolute position in a large random regular graph. The existence of this property on an RRG follows directly from the fact that it is trivially present on an infinite Bethe lattice, while any local quantity cannot differentiate between the 2 structures.

5.2 Spectral expansion of operators on a graph

There is a good portion of prior knowledge about the spectral properties of random regular graphs. The basic intuition behind all of the results above is intimately related to the local tree-like structure of RRGs, yet we will omit the corresponding discussion for now and go on with merely postulating some important theoretical results.

In what follows, we will use a concept of functions of matrices, although we will not dive deep into the corresponding branch of measure theory. For an analytical function f, one can define the value of a function f on a matrix A as a convergent series

$$f(A) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} A^k \tag{37}$$

extending the definition as necessary for other types of functions.

5.2.1 General structure of the spectrum

The spectrum of adjacency matrix A of an RRG consists of a continuous spectrum in the interval $\lambda \in \left[-2\sqrt{K}, 2\sqrt{K}\right]$ and a solitary eigenvalue $+(K+1)$. Interestingly enough, the situation is slightly different on an infinite Bethe lattice, where there also exists another solitary eigenvalue $-(K+1)$. Qualitatively this can be understood as a consequence of an inevitable existence of large loops in any finite random regular graph, as opposed to a purely loopless structure of a Bethe lattice.

5.2.2 Density of states

The reader might also be familiar with the result by McKay [21] that describes the density of eigenvalues of the adjacency matrix in the thermodynamical limit:

$$\rho_0(E) = \left\langle \frac{1}{N} \sum_{\lambda} \delta(E-\lambda) \right\rangle_{R(N,K)} = \theta\left(2\sqrt{K} - |E|\right) \cdot \frac{K+1}{2\pi} \cdot \frac{\sqrt{E^2-4K}}{(K+1)^2 - E^2}, \tag{38}$$

where the summation is done over all eigenvalues of a random regular graph, and the average is taken over the ensemble of all RRGs of a given size. Upon a proper normalization of the form $E = \varepsilon\sqrt{2K}$, one recognizes the large K limit of this expression to be the celebrated Wigner-Dyson semicircle. Additionally, knowledge of the density of states allows one to calculate traces of arbitrary matrix functions of the adjacency matrix, including powers of A. The latter are known to deliver the number of closed walks on a graph, so that one can reproduce the result (36) about the number of cycles.

At this point it becomes clear why one should exercise carefulness while using the inverse of A. Indeed, nonzero density of states at zero eigenvalue suggests that a large graph might have zero eigenvalue rendering its adjacency matrix degenerate. However, the existence of zero eigenvalue is not protected by any kind of symmetry and thus can occur in any finite RRG only by accident, with the majority of RRGs of the given size having eigenvalues close to but no equal to zero (as physicists, we simply checked this claim numerically, being unaware of any rigorous arguments on the topic). In other words, in the thermodynamical limit the set of all RRGs with degenerate adjacency matrix has measure zero, if it is not empty at all. Because we expect any physical answers to be completely insensitive to configurations of probability zero, it is safe to assume that despite having nonzero density of states at zero eigenvalue, the adjacency matrix of an RRG is still invertible. Mathematically speaking, it means that one expects no divergences to occur near zero eigenvalue for any spectral integrals expressing physical quantities.

5.2.3 Green function

Let us consider the Green function of the graph defined as

$$G(z) := (z - A)^{-1}, z \in \mathbb{C}, \tag{39}$$

where A is the adjacency matrix of the graph. By definition, a solution to all sorts of diffusion problems on a graph are expressed via the Fourier image of the Green function over its spectral variable z. Therefore, in the limit of infinite system size one expects the matrix elements $G_{ij}(z)$ of the Green function between the two sites i and j to be dependent only on the distance $d = |i - j|$ between the sites. As it is shown in [22], the matrix elements G_{ij} for a random regular graph demonstrate a well-defined limit as the total number of sites N approaches infinity:

$$G_d(z) = \frac{2}{\pi K^{d/2-1}} \int_0^\pi \frac{d\theta}{z - 2\sqrt{K}\cos\theta} \cdot \frac{\sin\theta \cdot [K\sin(d+1)\theta - \sin(d-1)\theta]}{(K+1)^2 - 4K\cos^2\theta} \tag{40}$$

From this result one can immediately restore McKay's result (38) by standard means:

$$\rho_0 = -\frac{1}{N\pi}\mathrm{Im}\{\mathrm{Tr}G(E + i0)\} \tag{41}$$

5.2.4 Fourier analysis on a random regular graph

Nevertheless, we can extract much more information out of the expression for the Green function. Based on expression (40), we can build a complete analog of Fourier analysis on finite-dimensional lattices, which proves to be an extremely useful tool for all practical calculations.

The key motivation for this part is the following observation: similarly to problems on finite-dimensional lattices, the translational invariance is restored upon averaging over disorder fields ξ. For our problem it means that any matrix necessarily expands in powers of adjacency matrix, because this is the only transitionally invariant matrix left in the problem. This also implies, that all matrices of the problem are diagonal in the same eigenbasis, again, in full alignment with the idea of plain waves for usual lattices. Our task is just to restore this expansion in powers of J.

Being a real symmetric matrix, any adjacency matrix can be expanded in projectors:

$$A = \sum_\lambda \lambda \cdot P^{(\lambda)},$$

where $P^{(\lambda)}$ is the projector to subspace of eigenvectors corresponding to eigenvalue λ. Same holds for any function of A:

$$f(A) = \sum_{\lambda} f(\lambda) \cdot P^{(\lambda)} \tag{42}$$

The Green function (39) is itself a yet another function of adjacency matrix and is thus no exception to this expansion. In particular, so is the difference between retarded and advanced functions of real spectral argument:

$$\frac{G(E-i0) - G(E+i0)}{2\pi i} = \int d\lambda \cdot \frac{1}{2\pi i} \left\{ \frac{P^{(\lambda)}}{E+i0-\lambda} - \frac{P^{(\lambda)}}{E-i0-\lambda} \right\} = \int d\lambda \cdot P^{(\lambda)} \cdot \delta(E-\lambda) \equiv P^{(\lambda)}, \tag{43}$$

where we have replaced summation over eigenvalues with integration applicable in the thermodynamical limit. Therefore, availability of the explicit expression (40) for G allows us to restore the form of P_E, which now reads:

$$P_d^{(E)} = \rho_0(E) \cdot \frac{K U_d\left(\frac{E}{2\sqrt{K}}\right) - U_{d-2}\left(\frac{E}{2\sqrt{K}}\right)}{K+1} \cdot \frac{1}{K^{d/2}}, \tag{44}$$

where $\rho_0(E)$ is McKay's density of states (38), and

$$U_n(\cos x) = \frac{\sin(n+1)x}{\sin x}$$

are Chebyshev polynomials of the second kind. This result is already quite useful as it allows one to restore all matrix elements of arbitrary functions of adjacency matrix via the spectral decomposition (42):

$$[f(A)]_d = \int\limits_{2\sqrt{K}}^{2\sqrt{K}} f(E) \cdot P_d(E) \cdot dE \tag{45}$$

It is convenient to get back from spectral parameterization E to angle parameterization θ:

$$[f(A)]_d = \int\limits_0^{\pi} f\left(2\sqrt{K}\cos\theta\right) \cdot P_d(\theta) \cdot \frac{d\theta}{\pi}, \tag{46}$$

where the projector $P_d(\theta)$ takes the following form after the substitution:

$$\begin{aligned} P_d(\theta) &= \frac{K\sin(d+1)\theta - \sin(d-1)\theta}{(K+1)^2 - 4K\cos^2\theta} \cdot \frac{2}{K^{d/2}} \\ &= \frac{(-i)K}{(K+1)^2 - 4K\cos^2\theta} \cdot \left[\frac{e^{i\theta}}{\sqrt{K}}\right]^d \left(Ke^{i\theta} - e^{-i\theta}\right)\sin\theta + (\theta \mapsto -\theta), \end{aligned} \tag{47}$$

where the second line is yielded after some algebraic transformation, and the second term is equal to the first one with an inverted value of θ. Because the function $f\left(2\sqrt{K}\cos\theta\right)$ is an even one under θ inversion, we can rewrite the angle integral (46) as

$$[f(A)]_d = \int\limits_{-\pi}^{\pi} f\left(2\sqrt{K}\cos\theta\right) \cdot \Pi(\theta) \cdot \left[\frac{e^{i\theta}}{\sqrt{K}}\right]^d \cdot \frac{d\theta}{2\pi} \tag{48}$$

$$\Pi(\theta) = \frac{-2iK\sin\theta}{Ke^{-i\theta} - e^{i\theta}} \tag{49}$$

In this form the expansion resembles the traditional diverging and converging waves expansion for usual lattices, with θ playing a role of the momentum. Contrary to power law decay of waves in a usual lattice setting, we can observe exponential decay $\propto K^{-d/2}$ that reflects exponential growth of a sphere on a random regular graph with the radius of that sphere. This also demonstrates that all proper functions of adjacency matrix of RRG have matrix elements that naturally decay with distance at least as $K^{-d/2}$.

We can compete our analysis by providing an analog of inverse Fourier transform. This is achieved in a standard way by exploiting the orthogonality of projectors $P^{(E)}$:

$$P^{(E)}P^{(E')} = \delta(E - E') \cdot P^{(E)} \tag{50}$$

However, to make the basis complete, one has to supply the set of projectors $P^{(K)}$ with a projector on the solitary eigenvalue $K+1$, which is known to be always present in the spectrum of an RRG. Luckily, the eigenfunction for this eigenvalue is just a uniform vector, so that the projector reads

$$\left[P^{(K+1)}\right]_{ij} = \frac{1}{N} \tag{51}$$

The full set of projectors can be shown to be from the complete decomposition of identity:

$$1 = \sum_{\lambda \in \left[-2\sqrt{K}, 2\sqrt{K}\right]} P^{(\lambda)} + P^{(K+1)}$$

Omitting some algebra, one then arrives to the following result: given an operator X that is translationally invariant on an RRG, one can restore its spectral decomposition with the following set of formulae:

$$X^{(K)} := \frac{1}{N}\mathrm{Tr}\left\{X \cdot P^{(K)}\right\} = \frac{1}{N^2} \cdot \sum_{ij} X_{ij}, \tag{52}$$

$$x(\theta) = \frac{1}{N}\mathrm{Tr}\left\{\left[X - X^{(K)} \cdot P^{(K+1)}\right] \cdot [\theta]_{ij}\right\}, \tag{53}$$

$$X(\theta) = \frac{\Pi(\theta)\, x(\theta) + \Pi(-\theta)\, x(-\theta)}{\Pi(-\theta) + \Pi(\theta)}, \tag{54}$$

where $\Pi(\theta)$ is the angle weight function (49), and $[\theta]$ is a shorthand for the spherical wave:

$$[\theta]_{ij} = \left[\frac{e^{i\theta}}{\sqrt{K}}\right]^{|i-j|} \tag{55}$$

Finally, because we expect the expansion (42) to be applied to translationally invariant operators in the first place, it is worth rewriting the trace of a translationally invariant operator (which the traced expression in (53) is) as a sum over distances:

$$\frac{1}{N}\mathrm{Tr}\{X\} = X_0 + \sum_{d=1}^{\infty} (K+1)\, K^{d-1} \cdot X_d, \tag{56}$$

where the coefficient $(K+1)\,K^{d-1}$ reflects the area of a sphere of radius d on a K-regular tree. Strictly speaking, one can either apply the expression (56) for an infinite Bethe lattice or assume

that the terms of the series decay at sufficient rate as d grows, so that the whole series gains its value on a finite set of terms and thus is ignorant about the difference between the Bethe lattice and a large RRG. The expression (48), however, suggests that exponential decay of matrix elements with distance is always the case for all functions of adjacency matrix. For physics level of rigorousness this is surely enough.

In this way we are able to build a counterpart of Fourier analysis on an RRG, with (48) and (52-54) playing the role of direct and inverse expansion over spherical waves, respectively. It is hard to underestimate the usefulness of this result for further calculations, so we will freely turn to these results in what follows without further explanations.

Part III
Saddle point equation

6 Main features of the saddle-point equation

6.1 Derivation and free energy

One typically expects mean-field analysis to be perfectly applicable for a well developed superconductivity outside the fluctuation region [16]. In the language of functional integration this corresponds to a saddle-point approximation for the value of the functional integral. The corresponding saddle-point is sought among τ-independent field configurations. In this case, the action (20) reads:

$$S\left[\Delta\right]=\frac{\beta}{2}\cdot\Delta_i^{\alpha}\left(\frac{J}{2}\right)_{ij}^{-1}\Delta_j^{\beta}-\sum_i\ln\left[\frac{\cosh\left\{\frac{\beta}{2}\sqrt{\xi_i^2+\left|\Delta_i\right|^2}\right\}}{\cosh\left\{\frac{\beta}{2}\left|\xi_i\right|\right\}}\right] \tag{57}$$

The explicit evaluation of the functional determinant (21) is possible for configurations with constant in time Δ because the associated dynamical problem for the semion subsystem (or spin subsystem, for that matter) is easily solvable. The denominator $\cosh\left\{\beta\left|\xi_i\right|/2\right\}$ represents the contribution from the normal phase, so that the resulting action can be regarded as the free energy difference between normal and superconducting phases.

By direct variation of the free energy (57) w.r.t the order parameter field Δ_i^{α} we arrive at the following saddle-point equation:

$$\Delta_i^{\alpha}=\sum_j J_{ij}\cdot\frac{\Delta_j^{\alpha}}{\sqrt{\xi_j^2+\left|\Delta_j\right|^2}}\frac{\tanh\left\{\beta\sqrt{\xi_j^2+\left|\Delta_j\right|^2}\right\}}{2} \tag{58}$$

Given a particular RRG and a realization of fields ξ_i, one has to solve the set of equations (58) for the values of order parameter Δ_i on each site. One is then interested in the statistics of various functions of Δ_i and ξ_i with respect to the ensemble of ξ's.

Provided that current and/or magnetic field is absent in the system, the (complex) order parameter can naturally be chosen to be real across the whole system, which corresponds to

the choice $\Delta_i^\alpha = \Delta_i \cdot \delta^{\alpha,1}$ in vector notation, so that one arrives to the equation on the absolute value of the order parameter:

$$\Delta_i = \sum_j J_{ij} \cdot \frac{\Delta_j}{\sqrt{\xi_j^2 + \Delta_j^2}} \frac{\tanh\left\{\beta\sqrt{\xi_j^2 + \Delta_j^2}\right\}}{2} \tag{59}$$

The equation has a well-defined zero-temperature limit:

$$\beta \to \infty : \Delta_i = \sum_j \frac{J_{ij}}{2} \cdot \frac{\Delta_j}{\sqrt{\xi_j^2 + \Delta_j^2}} \tag{60}$$

Below we investigate various statistical properties of the solution to this equation in different disorder realizations, starting with some crude approximation revealing the key qualitative features of the equation. Our study is focused on zero temperature limit $T = 0$, although we will refer to some previous studies carried out for the system at finite temperature.

6.2 Infinite K approximation and connections with the conventional theory of superconductivity

The physical sense of the self-consistency equation becomes clear if one formally considers the case of arbitrarily large K. In this case, the r.h.s of the equation represents a sum of large number of terms and thus can be replaced by its mean value w.r.t the sought distribution:

$$\Delta_i = g \cdot \left\langle \frac{\Delta}{\sqrt{\xi^2 + \Delta^2}} \tanh\left\{\beta\sqrt{\xi^2 + \Delta^2}\right\} \right\rangle_{\xi,\Delta} \tag{61}$$

where the average should be done over the joint distribution of ξ and Δ on the given site. One can now observe that the r.h.s becomes site-independent, thus suggesting a constant value of the order parameter across the whole system. The equation then takes a well-recognized form of the self-consistency equation on the homogeneous super-conducting order parameter [REF_HERE]:

$$\Delta_0 = g \cdot \left\langle \frac{\Delta_0}{\sqrt{\xi^2 + \Delta_0^2}} \tanh\left\{\beta\sqrt{\xi^2 + \Delta_0^2}\right\} \right\rangle_\xi = g \cdot \int P(\xi)\, d\xi \cdot \frac{\Delta_0}{\sqrt{\xi^2 + \Delta_0^2}} \tanh\left\{\beta\sqrt{\xi^2 + \Delta_0^2}\right\} \tag{62}$$

thus also confirming the role of the disorder distribution $P(\xi)$ as the single-particle density of states. There are 2 particularly instructive special cases:

- upon setting $\Delta_0 = 0$, one obtains a standard equation for the transition temperature T_c:

$$1 = g \cdot \int P(\xi)\, d\xi \cdot \frac{\tanh\{|\xi|/T_c\}}{|\xi|} \tag{63}$$

- whenever $g \ll 1$, the solution to the self-consistency equation (62) is exponentially small w.r.t g. Let us consider $T = 0$ and a box-shaped disorder distribution $P(\xi) = \frac{1}{2}\theta(1 - |\xi|)$ for simplicity. Assuming Δ_0 to be small, one can evaluate the integral approximately and solve for the value of Δ_0, obtaining:

$$g \ll 1, P(\xi) = \frac{1}{2}\theta(1 - |\xi|) : \Delta_0 \approx 2\exp\left\{-\frac{1}{g}\right\} \ll 1 \tag{64}$$

Because the integral of ξ is logarithmic, the exact form of the disorder distribution will only alter the numerical coefficients in front of $1/g$ exponential dependence and in the prefactor, whereas the exponential dependence itself is quite a robust feature of the model[3].

In what follows, $\Delta_0 \ll 1$ is assumed to be among the smallest energy scales of the problem, owing to its exponential smallness.

6.3 Previous studies: vicinity of the transition point $T = T_c$

At large but finite values of K, one expects the transition temperature to fluctuate from one realization to another with both mean value and characteristic scale of fluctuations determined by the relation between the parameters of the model. The equation (59) has been studied in great detail in the vicinity of the transition point where the saddle-point equation can be linearized w.r.t small value of the order parameter [12]. Assuming $g \ll 1$ and $P(0) = 1$, the key findings relevant to our problem can be summarized in the following fashion:

- The linearized counterpart of the saddle point equation (59) is rendered inapplicable in the region $K \lesssim K_1$, where K_1 is given by:

$$K_1 = g \exp\left\{\frac{1}{2g}\right\} \gg 1 \tag{65}$$

  Note that this scale is still a large quantity. Yet, when K is as small as K_1, slow fluctuations are known to destroy the saddle-point approximation. The associated distribution of the true order parameter (which is not given by (58) anymore) demonstrates fat tails and lack of any kind of self-averaging.

- When the parameter K is greater than some scale K^*, one expects the equation to exhibit self-averaging behavior similar to the one described by the mean field approximation in the previous section. The theory predicts K^* to be given by:

$$K^* = \frac{\pi}{4} e^{-\gamma} \cdot g \exp\left\{\frac{1}{g}\right\}, \tag{66}$$

  where $\gamma \approx 0.577$ is the Euler–Mascheroni constant. Above $K = K^*$ the fluctuations are suppressed by virtue of the central limit theorem as previously discussed. Note that $K^* \gg K_1$, thus leaving an exponentially large range $K_1 \ll K \ll K^*$ with nontrivial behavior of the saddle-point equation (59), where one expects the order parameter distribution to be nontrivial.

Although the analysis of [12] is only valid near $T = T_c(K, g)$, one expects the implications to be qualitatively correct at any T. In other words, the fluctuations and self-interaction play important role only at $\ln K \lesssim \ln K_1 \sim 1/2g$. In this work, we are going to study the regime $K \gg K_1$, when the order parameter is well described by its saddle-point value found by solving (59). Note that because $g \ll 1$, this also implies that K are themselves exponentially large. And even when K is a large quantity, there is still a room for nontrivial behavior in the region $K \lesssim K^*$ because of strong fluctuations of the r.h.s of the saddle point equation (59).

[3]For instance, a Gaussian model of disorder gives

$$g \ll 1, P(\xi) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{\xi^2}{2}\right\} : \Delta_0 \approx 2\sqrt{2} e^{-\gamma/2} \cdot \exp\left\{-\sqrt{\frac{\pi}{2}}\frac{1}{g}\right\} \approx 2.12 \exp\left\{-\frac{1.25}{g}\right\}$$

with $\gamma \approx 0.577$ being the Euler–Mascheroni constant.

6.4 Numerical solution

The analytical study presented below is backed by numerical investigation of the problem. The corresponding numerical routine consists of the following steps:

- Generate a random regular graph with a total of N sites and $K+1$ neighbors each. Systems with up to $N = 2^{22} \approx 4.2 \cdot 10^6$ sites were produced.
- Generate M realizations of disorder field ξ_i. The number of realization is generally chosen based on the system size, as larger systems require more computational time for their solution in return for better statistical self-averaging within a single sample.
- For each of M disorder realizations, numerically solve the saddle-point equation (60) for the values of Δ_i.
- Average the required function of fields Δ, ξ over the obtained solutions \ calculate the required distribution functions.

For each instance of numerical results, the values of the parameters g, K, as well as the used system size N and the number of disorder realizations M are specified, so that one is able to estimate the effects of the finite size of the system as well as the quality of statistical averaging over different disorder realizations. The main qualitative aspects of the latter deserve a separate consideration:

- The key effect of a finite system size N is the distortion introduced to the local tree-like structure of a large RRG due to the existence of short cycles in the system. Indeed, given an RRG of degree $K+1$ and size N, the local tree-like structure for a given site can only develop up to a distance
$$d_c \sim \frac{\ln N}{\ln K}, \tag{67}$$
at which point it already includes a finite fraction of sites of the full system. Therefore, one expects that those physical properties resting on correlations at distances $d \gtrsim d_c$ will be significantly distorted. One, however, expects that the physical behavior is completely determined by local effects with bounded correlation length, so that one would observe saturation of all physical quantities as system size N grows.
- Secondly, the system size itself contributes to the quality of averaging of local quantities. Indeed, if a given effect rests on correlations at distances up to r, one can effectively split the whole system into a set of uncorrelated regions of size $V_r \sim K^r$, and the resulting number of regions plays a role similar to that of M.

One technical obstacle is that for very large values of $K \geq 200$ the procedure of sampling random regular graphs becomes computationally expensive, as the average time to generate a sample grows at least as NK^3 [23]. To circumvent this problem, we used the approximate routine that generates nearly regular graphs by randomly partitioning K copies of N vertices into pairs corresponding to graph edges and then removing loops and double edges. Such an approximation leads to a finite variance of vertex degrees in the system. This approximate algorithm was used only for sufficiently large K, when any correlations are well suppressed and the Gaussian regime dominates, so that the exact structure of the graph is rendered irrelevant. In what follows, we explicitly indicate the usage of this approximation in the numerical calculations.

7 Intuitive approximations

Before diving into the full analysis of the self-consistency equation, let us first review some approximate yet instructive schemes of analysis that provide important intuition and technical aspects for the full solution. In particular, they serve as a good semi-quantitative introduction to the role of the key control parameters of the theory.

7.1 Large K approximation and structure of correlations

Let us start with the case when K is large enough to suppress the fluctuations on the r.h.s of the saddle point equation (60). In this case one expects the solution Δ_i to exhibit small fluctuations around its mean value. One can then expand the r.h.s of (60) w.r.t weak fluctuations around the mean value:

$$\delta_i = \Delta_i - \langle\Delta\rangle \ll \langle\Delta\rangle \tag{68}$$

$$\langle\Delta\rangle + \delta_i = \sum_j \frac{J_{ij}}{2} \cdot \left(\frac{\langle\Delta\rangle}{\sqrt{\xi_j^2 + \langle\Delta\rangle^2}} + \frac{\xi_j^2}{\left(\xi_j^2 + \langle\Delta\rangle^2\right)^{3/2}} \cdot \delta_j \right) \tag{69}$$

One can explicitly express the value of δ_i:

$$\delta = X^{-1} v, \tag{70}$$

$$X_{ij} = \delta_{ij} - \frac{J_{ij}}{2} \cdot \frac{\xi_j^2}{\left(\xi_j^2 + \langle\Delta\rangle^2\right)^{3/2}} \tag{71}$$

$$v_i = \left(\sum_j \frac{J_{ij}}{2} \cdot \frac{\langle\Delta\rangle}{\sqrt{\xi_j^2 + \langle\Delta\rangle^2}} \right) - \langle\Delta\rangle \tag{72}$$

As we can see, matrix X has δ_{ij} as a leading term with a small correction of order g/K. Therefore, the behavior of δ is qualitatively determined by the fluctuations of vector v. The latter represents a sum of a large number of independent fluctuating variables, thus suggesting applicability of the central limit theorem. Assuming Gaussian statistics for δ, one then arrives to the following correlator of the fluctuations:

$$\left\langle \delta\delta^T \right\rangle = \left\langle X^{-1} \cdot vv^T \cdot \left(X^{-1}\right)^T \right\rangle_\xi , \tag{73}$$

which should also be supplied with the equation on $\langle\Delta\rangle$ derived from its definition:

$$0 \equiv \langle \Delta - \langle\Delta\rangle \rangle \equiv \langle\delta\rangle = \left\langle X^{-1} v \right\rangle_\xi \tag{74}$$

Because both X and v contain the disorder field ξ, calculating both the correlator $\left\langle \delta\delta^T \right\rangle$ and the mean value $\langle\Delta\rangle$ requires employing a diagrammatic technique to deal with the inversion of a random matrix X. The associated theoretical analysis involves a simple self-consistent Born approximation and will be discussed elsewhere, while here only the most relevant results are presented, as the derivation technique is not of great importance on its own.

7.1.1 Simple approximation for onsite fluctuation

It is instructive to review the naive approximation that manages to correctly reproduce the leading terms for the mean value and the variance of the order parameter distribution. Let us approximate the value of the X matrix with only the leading term:

$$X^{-1} = 1 + O\left(\frac{g}{\sqrt{K}}\right), \tag{75}$$

where the K-dependence of the correction is obtained from the typical spectrum of a random regular graph:

$$J \sim g\lambda/K, \lambda \sim \sqrt{K} \Rightarrow J \sim g/\sqrt{K} \tag{76}$$

Upon ignoring this correction, the averaging is trivial and leads to

$$\langle \Delta \rangle = \Delta_0 + O\left(\frac{1}{K}\right), \tag{77}$$

$$\left\langle \delta\delta^T \right\rangle \approx \left(\frac{J}{2}\right)^2 \cdot \left\langle \left(\frac{\langle \Delta \rangle}{\sqrt{\xi^2 + \langle \Delta \rangle^2}} - \left\langle \frac{\langle \Delta \rangle}{\sqrt{\xi^2 + \langle \Delta \rangle^2}} \right\rangle \right)^2 \right\rangle_{\xi} \approx \left(\frac{J}{2}\right)^2 \cdot 2P(0) \cdot \frac{\pi \Delta_0}{2}, \tag{78}$$

where Δ_0 is the previously introduced mean field value of the order parameter (62). The second line is obtained as the leading term in the expansion of the exact average over ξ in powers of Δ_0. According to these predictions, one expects Δ to be distributed as a Gaussian random variable with the mean value of each component given by (77) and covariance matrix given by (78). The obtained results appear to be a correct answer up to the leading order in formal expansion in power of K.

In particular, the diagonal matrix element of the covariance matrix(78) represents the variance of the order parameter. It is convenient to compare it with the mean value. The result reads:

$$\frac{\langle \delta^2 \rangle}{\langle \Delta \rangle^2} = \left(\frac{J}{2}\right)^2_{ii} \cdot 2P(0) \cdot \frac{\pi}{2\Delta_0} = \frac{g^2}{K\Delta_0} \cdot 2P(0) \cdot \frac{\pi}{2} \tag{79}$$

As it is expected within the Gaussian regime, the relative fluctuations decay as K^{-1}. From this result we can infer the criteria of applicability as a consequence of assumed smallness of the relative fluctuations:

$$\frac{\langle \delta^2 \rangle}{\langle \Delta \rangle^2} \ll 1 \Rightarrow \frac{g^2}{K\Delta_0} \ll 1, \tag{80}$$

where we have dropped numerical prefactors of the order unity. One particular conclusion is that the described Gaussian regime only kicks in when K is as large as

$$K \gtrsim K_{\text{Gauss}} = g^2 \cdot \Delta_0^{-1} \approx 2g^2 \exp\left\{-\frac{1}{g}\right\}, \tag{81}$$

which is consistent with a more thorough analysis of [12], since the obtained scale coincides up to exponential precision with the value of K^* given by (66), i.e.,

$$\ln K_{\text{Gauss}} \approx \ln K^* \tag{82}$$

The presented analysis, however, fails to recover the behavior of the cross-site correlations of the order parameter, that are supposed to be given by the off-diagonal matrix elements of the covariance matrix (78). Within the answer obtained so far, the only nonzero matrix element is for distance $d = 2$ correlations and is given by

$$|i-j| = 2 : \langle \delta_i \delta_j \rangle = \left\langle \delta^2 \right\rangle / K, \tag{83}$$

while all other correlations are equal to zero:

$$|i-j| \neq 0, 2 : \langle \delta_i \delta_j \rangle = 0 \tag{84}$$

The result is obtained by using the fact that the matrix element of a power n of the adjacency matrix of the graph gives the number of distinct paths of length n connecting the 2 given sites. Given the local tree-like structure of the graph, one obtains:

$$(J/2)^2_{ij} = \frac{g^2 (K+1)}{K^2} \cdot \delta_{ij} + \frac{g^2}{K^2} \delta_{|i-j|,2} \tag{85}$$

which leads to answers presented above. In reality, however, the correlations between 2 arbitrary sites do exist, but these correlations are beyond the precision of the proposed simplistic approximation and represent subleading corrections in formal $1/K$ expansion. The results of taking into account the subleading terms of the X matrix are presented below and are found to accurately reproduce the entire structure of the correlations within the region of applicability given by (81).

7.1.2 Results of a more accurate approximation

Without going into much detail, we present here the answers of a more accurate computation of the covariance matrix. The simplistic treatment described above is improved by accounting for the influence of subleading corrections to the value of the average $\langle X \rangle$ of the X matrix. First of all, one can compute the $1/K$ correction to the mean value $\langle \Delta \rangle$ of the order parameter:

$$\langle \Delta \rangle \approx \Delta_0 \left(1 + \frac{4}{3} \cdot \frac{K_1}{K} \cdot \frac{1-g}{g} \right) \tag{86}$$

with K_1 being the native scale for the corrections:

$$K_1 = \frac{3\pi}{16} \cdot \frac{g^2}{\langle \Delta \rangle} \approx \frac{3\pi}{16} \cdot \frac{g^2}{\Delta_0} \sim K_{\text{Gauss}}, K^* \tag{87}$$

Secondly, one finds the following expression for the covariance matrix:

$$\left\langle \delta \delta^T \right\rangle = \langle \Delta \rangle^2 \cdot \left[\frac{(J/2)}{1 - \left(\frac{1}{g} - 1 \right) (J/2)} \right]^2 \tag{88}$$

One can calculate the matrix elements explicitly, obtaining:

$$|i-j| = d : \langle \delta_i \delta_j \rangle = \begin{cases} \frac{g^2}{K \langle \Delta \rangle} \left\langle \left(\frac{\langle \Delta \rangle}{\sqrt{\xi^2 + \langle \Delta \rangle^2}} - \left\langle \frac{\langle \Delta \rangle}{\sqrt{\xi^2 + \langle \Delta \rangle^2}} \right\rangle \right)^2 \right\rangle_\xi, & d = 0 \\ \langle \delta^2 \rangle \cdot \frac{4(1-g)}{K}, & d = 1 \\ \langle \delta^2 \rangle \cdot (d-1) \frac{(1-g)^{d-2}}{K^{d-1}}, & d \geq 2 \end{cases} \tag{89}$$

One can see that the simplistic approximation presented above does indeed correctly reproduce the leading order in the formal $1/K$ expansion.

The results are showcased on Figures 5 through 7. One can observe that the presented theory is in excellent agreement with the numerical data within the region of applicability $K \gg K_1$ and in some cases even up to $K \approx K_1$, as seen on the plot of the mean value on Figure 6. There is, however, a resolvable skewness of the true distribution seen in the discrepancy with the Gaussian fit of the data on Figure 5. This effect is discussed in the next section as it is beyond the Gaussian approximation.

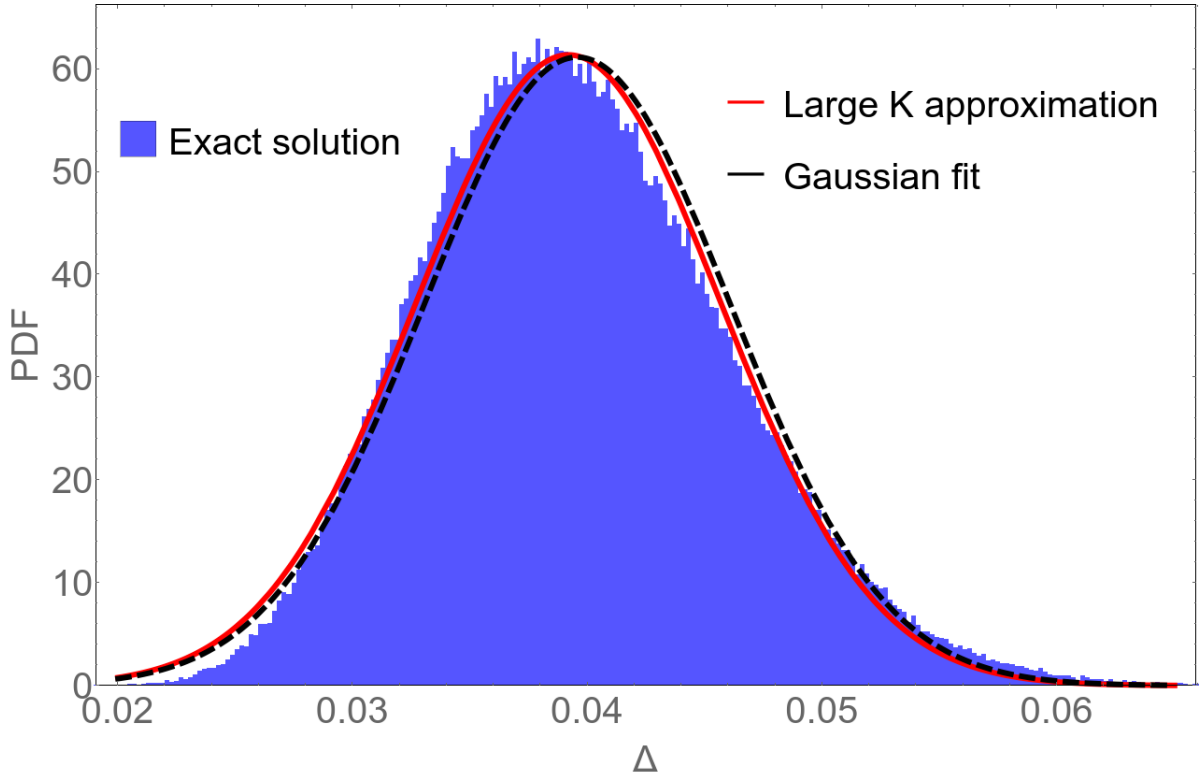


Figure 5: The onsite distribution of the order parameter Δ. Parameters of the model: P_ξ is a box-shaped distribution in the domain $[-1, 1]$, $K = 58$, and $g = 0.25$. The blue histogram represents the statistics of the exact numerical solution of the self-consistency equation on a random regular graph with $N = 2^{18} \approx 2.6 \cdot 10^5$. The dashed line is a Gaussian distribution with $\langle\Delta\rangle = 3.96 \cdot 10^{-2}$, $\sigma_\Delta = 6.52 \cdot 10^{-3}$, with both parameters calculated directly from the numerical data. The red solid line represents the Gaussian distribution with theoretically predicted parameters $\langle\Delta\rangle = 3.92 \cdot 10^{-2}$, $\sigma_\Delta = 6.50 \cdot 10^{-3}$, as given by (86) and (89).

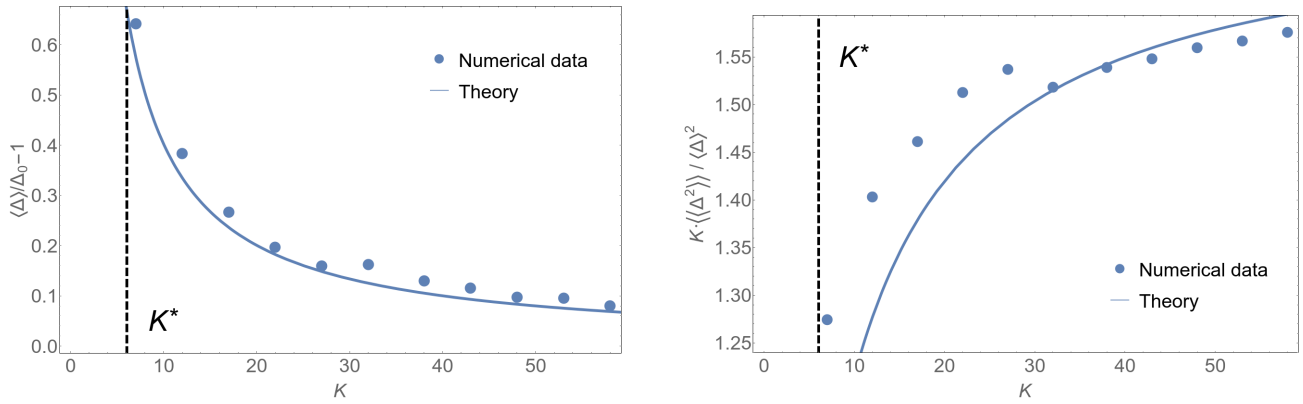


Figure 6: Dependencies of the mean value (left) and variance (right) of Δ on K, normalized to native units: the left plot shows the relative difference to the mean field value Δ_0, while the right plot represents the variance normalized to the squared mean value $\langle\Delta\rangle$ and native factor of K. The parameters of the model, other than K, are identical to those on Figure 5. On both plots, points represent the data extracted from exact numerical solution with the same parameters as those of Figure 5, while solid lines demonstrate the theoretical predictions given by (86) - (89). The vertical dashed line represents the value of K^* where one expects the Gaussian model to break down.

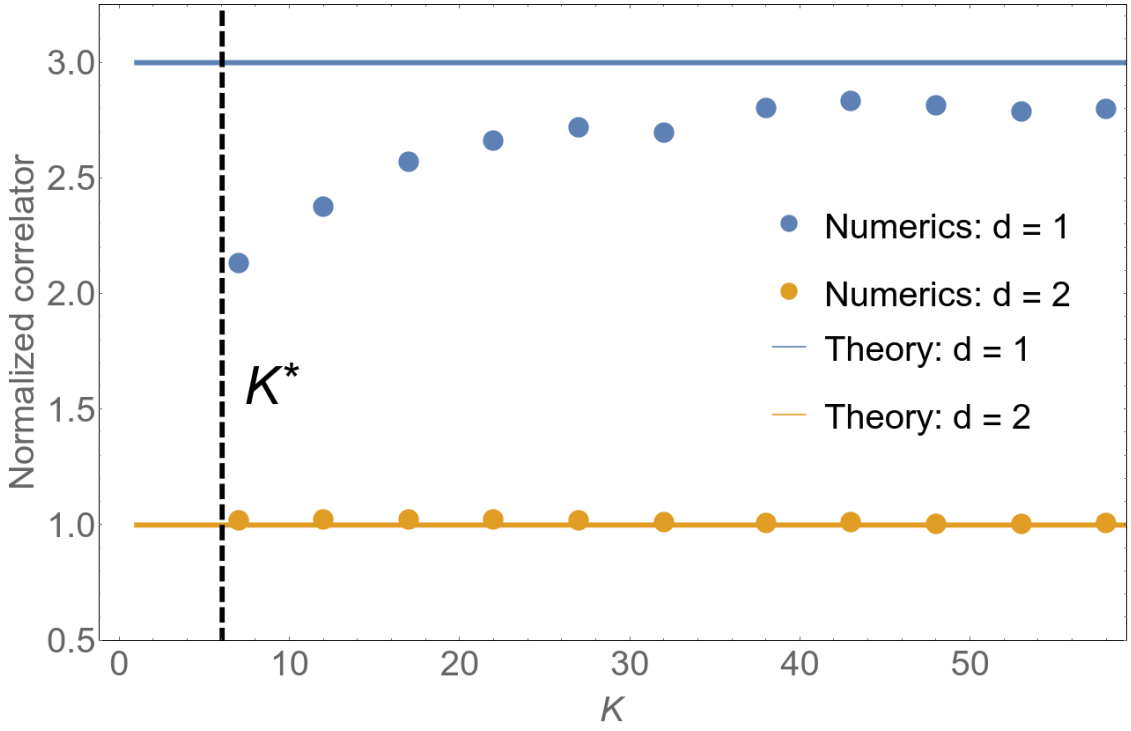


Figure 7: Dependence of the off-diagonal matrix elements of the covariance matrix $\left\langle\left\langle \delta\delta^T \right\rangle\right\rangle$ of the order parameter on K for 2 distances $d = |i - j| = 1$ (data in blue) and $d = |i - j| = 2$ (data in orange). The parameters of the model are identical to those on Figure 5. The values of the correllators are normalized to $\langle\delta^2\rangle / K$. Parameters for both the data from exact numerical solution (points) and theoretical predictions (solid lines) are the same as on Figure 6. The vertical dashed line represents the value of K^* where one expects the Gaussian model to break down.

The main qualitative conclusions to be drawn from these results about the large K-regime are the following:

- the order parameter is distributed with a correlated Gaussian distribution with the mean value given by (86) and covariance matrix given by (88). The characteristic scale for both the average order parameter and its fluctuations is given by the mean field value Δ_0, determined from the solution to the mean-field self-consistency equation.

- Correlations in the system, characterized by off-diagonal matrix elements of the covariance matrix, decay exponentially with distance and can be estimated as

$$\langle \delta_i \delta_j \rangle \sim \frac{\langle \delta \rangle^2}{K^{|i-j|-1}}, |i-j| \geq 2 \tag{90}$$

- The self-averaging regime of the zero-temperature saddle-point approximation is applicable at

$$K \gtrsim \frac{g}{\Delta_0} \sim g \exp\left\{\frac{1}{g}\right\}, \tag{91}$$

  which is consistent with previous studies for $T \approx T_c$ case.

7.2 Approximation of uncorrelated sites and non-Gaussian regime

As we have seen in the previous section, the self-averaging behavior of the r.h.s of the saddle point equation (59) is only guaranteed at sufficiently large $K \gtrsim g \cdot \exp\{1/g\}$. Below this limit one can expect a nontrivial interplay between the tendency to the self-averaging due to a large number of terms in the r.h.s. and the strong fluctuations of each term due to exponential smallness of the typical value of the order parameter Δ_0. As it appears, this interplay can be qualitatively reproduced by a simple model that we are going to present below. Although this model does not have an explicit control parameter that measures the similarity to the actual saddle-point equation, this model reveals some technical aspects essential to the derivation of the full theory.

The model rests on the fact that correlations in the order parameter decay exponentially with distance, as it was observed in the previous section. Therefore, a qualitatively correct approximation would be to neglect the correlations completely and treat the Δ field as a set of independent random variables on each site of the system. The problem is then reduced to finding the corresponding onsite distribution of the order parameter $P(\Delta)$.

Subsections 7.2.1 to 7.2.4 present a detailed derivation of the main equations, that are, in turn, summarized in Subsection 7.2.5. Illustrations of the results as well as discussion of main qualitative features can be found in Subsection 7.2.6. Although they reveal some uncontrolled discrepancy with the true Δ distribution originating from the ignored inter-site correlations, the technical aspects of the solution to this model will be of crucial importance for the full theory.

7.2.1 Equation on the cumulant generating function

One starts by explicitly averaging the r.h.s of the saddle-point equation:

$$P(\Delta) := \langle \delta(\Delta - \Delta_i) \rangle_{\Delta_i, \xi} = \left\langle \delta\left(\Delta - \sum_j \frac{J_{ij}}{2} \cdot \frac{\Delta_j}{\sqrt{\xi_j^2 + \Delta_j^2}}\right) \right\rangle_{\xi, \Delta}, \tag{92}$$

where the average is performed over the onsite distributions of ξ and Δ. The next step is to rewrite the δ-function as a Fourier integral:

$$
\begin{aligned}
... &= \left\langle \int\limits_{\mathbb{R}} \frac{d\omega}{2\pi} \cdot \exp\left\{ -i\omega \left(\Delta - \sum_j \frac{J_{ij}}{2} \cdot \frac{\Delta_j}{\sqrt{\xi_j^2 + \Delta_j^2}} \right) \right\} \right\rangle_{\xi,\Delta} \\
&= \int\limits_{\mathbb{R}} \frac{d\omega}{2\pi} \cdot e^{-i\omega\Delta} \cdot \left\langle \prod_j \exp\left\{ i\omega \cdot \frac{J_{ij}}{2} \cdot \frac{\Delta_j}{\sqrt{\xi_j^2 + \Delta_j^2}} \right\} \right\rangle_{\xi,\Delta}
\end{aligned}
$$

Then we make use of the independence assumption, so that average of a product of local variables factorizes into a product of averages:

$$
\begin{aligned}
... = P(\Delta) &= \int\limits_{\mathbb{R}} \frac{d\omega}{2\pi} \cdot e^{-i\omega\Delta} \cdot \prod_j \left\langle \exp\left\{ i\omega \cdot \frac{J_{ij}}{2} \cdot \frac{\Delta_j}{\sqrt{\xi_j^2 + \Delta_j^2}} \right\} \right\rangle_{\xi,\Delta} \\
&= \int\limits_{\mathbb{R}} \frac{d\omega}{2\pi} \cdot e^{-i\omega\Delta} \cdot \left\langle \exp\left\{ i\omega \cdot \frac{g}{K} \cdot \frac{\Delta}{\sqrt{\xi^2 + \Delta^2}} \right\} \right\rangle_{\xi,\Delta}^K \\
&= \int\limits_{\mathbb{R}} \frac{d\omega}{2\pi} \cdot e^{-i\omega\Delta} \cdot \left[\int\limits_{\mathbb{R}} d\xi P(\xi) \int\limits_{\mathbb{R}} d\Delta P(\Delta) \exp\left\{ i\omega \cdot \frac{g}{K} \cdot \frac{\Delta}{\sqrt{\xi^2 + \Delta^2}} \right\} \right]^K ,
\end{aligned} \tag{93}
$$

where we have plugged in the explicit expression for J_{ij}. The resulting equation is an integral equation on the distribution function $P(\Delta)$.

Let us make some convenient variable substitutions:

- Rescaling the fluctuating fields to their native scale for this problem is convenient:

$$
x = \frac{\xi}{\Delta_0}, y = \frac{\Delta_i}{\Delta_0}, s = \Delta_0 \omega \tag{94}
$$

  where Δ_0 is the mean-field value of the order parameter. At this point, it serves purely as a characteristic scale, the exact value of which will be manually fixed later. To keep it simple, one can think of Δ_0 as the value given by (62).

- We will make use of the cumulant generating function for Δ fields:

$$
R(s) := \ln\left(\langle \exp\{isy\} \rangle_{y=\frac{\Delta}{\Delta_0}} \right) \equiv \ln\left(\int\limits_0^\infty d\Delta P(\Delta) \exp\left\{ is \cdot \frac{\Delta}{\Delta_0} \right\} \right) \tag{95}
$$

  Let us highlight the following important properties of R:

  - because $\Delta > 0$ (as it is reflected in the integration limits of the Δ integral), $\exp\{R(s)\}$ is an analytical function in the upper half-plane of the complex variable s.
  - Because the value of Δ is trivially bounded from above as $\Delta \leq g$, the function $\exp\{R(s)\}$ decays exponentially when $\mathrm{Im} s \gg 1$. One can expect this property to kick in at scales $\mathrm{Im} s \gtrsim g^{-1}\Delta_0^{-1} \gg 1$.

– Because Δ is a real quantity, complex conjugation can be factored out:

$$\exp\left\{R\left(\overline{s}\right)\right\} = \overline{\exp\left\{R\left(s\right)\right\}} \tag{96}$$

– the normalization of probability density P_Δ is translated in the exact identity

$$\int\limits_{\mathbb{R}} d\Delta P_\Delta\left(\Delta\right) = 1 \Rightarrow R\left(0\right) = \ln 1 = 0 \tag{97}$$

– The value of the probability function P_Δ is restored by inverse Fourier transform:

$$P_\Delta\left(\Delta\right) = \frac{1}{\Delta_0}\int\limits_{\mathbb{R}} \frac{ds}{2\pi} \exp\left\{-is\frac{\Delta}{\Delta_0}\right\} \cdot \exp\left\{R\left(s\right)\right\} \tag{98}$$

The integral is convergent because the real part of R is negative and can be shown to grow at least linearly with $|s|$ at large real s (as we will see later). The reality condition for R guarantees a real value of P.

- Let us also introduce a dimensionless parameter characterizing the relation between the energy scale of the order parameter and the number of neighbors K:

$$v := K\Delta_0 \tag{99}$$

The equation can be then rewritten in terms of the R function:

$$R\left(s\right) = \ln\left(\left[\int\limits_{\mathbb{R}} d\xi P_\xi\left(\xi\right)\int\limits_{\mathbb{R}} d\Delta P_\Delta\left(\Delta\right)\exp\left\{is\cdot\frac{g}{v}\frac{\Delta}{\sqrt{\xi^2+\Delta^2}}\right\}\right]^K\right) \tag{100}$$

7.2.2 Expansion in powers of small Δ_0

In our case we are able to evaluate the r.h.s of the equation on R in the limit of large K, small Δ_0 and constant v. First step is to use the fact that most integrals over both Δ and ξ naturally converge in a small region of size $\Delta, \xi \sim \Delta_0$ and thus contain an associated smallness of the order Δ_0. In particular, this is true for integrals of all functions that decay as ξ^{-2} or faster at large ξ. One can then carry out the following calculation:

$$\begin{aligned}
&\int\limits_{\mathbb{R}} d\xi P_\xi\left(\xi\right)\int\limits_{\mathbb{R}} d\Delta P_\Delta\left(\Delta\right)\exp\left\{is\cdot\frac{g}{v}\cdot\frac{\Delta}{\sqrt{\xi^2+\Delta^2}}\right\}\\
&= \int\limits_{\mathbb{R}} d\xi P_\xi\left(\xi\right)\int\limits_{\mathbb{R}} d\Delta P_\Delta\left(\Delta\right)\cdot 1\\
&+ \left(is\cdot\frac{g}{v}\right)\cdot\int\limits_{\mathbb{R}} d\xi P_\xi\left(\xi\right)\int\limits_{\mathbb{R}} d\Delta P_\Delta\left(\Delta\right)\cdot\frac{\Delta}{\sqrt{\xi^2+\Delta^2}}\\
&+ \int\limits_{\mathbb{R}} d\xi P_\xi\left(\xi\right)\int\limits_{\mathbb{R}} d\Delta P_\Delta\left(\Delta\right)\left(\exp\left\{is\cdot\frac{g}{v}\frac{\Delta}{\sqrt{\xi^2+\Delta^2}}\right\} - \left\{1+is\cdot\frac{g}{v}\frac{\Delta}{\sqrt{\xi^2+\Delta^2}}\right\}\right)
\end{aligned}$$

Due to the normalization of probabilities P_ξ, P_Δ, the first integral is equal to 1, and the integrand in the 3rd term decays with ξ as ξ^{-2} and thus converges in a small region of $\xi \sim \Delta_0 \ll 1$, so that one can completely neglect the ξ-dependence of P_ξ as it has a typical scale of $\xi \sim 1$. One then proceeds with

$$\begin{aligned}
&... = 1 + \left(is \cdot \frac{g}{v}\right) \cdot \Delta_0 \cdot C_1 + \\
&+ P_\xi(0) \int\limits_{\mathbb{R}} d\Delta P_\Delta(\Delta) \int\limits_{\mathbb{R}} d\xi \left(\exp\left\{ is \cdot \frac{g}{v} \frac{\Delta}{\sqrt{\xi^2+\Delta^2}} \right\} - \left\{ 1 + is \cdot \frac{g}{v} \frac{\Delta}{\sqrt{\xi^2+\Delta^2}} \right\} \right) \\
&+ O\left(\Delta_0^2\right),
\end{aligned}$$

where the last correction term contains all sorts of corrections from actual dependence of P on ξ, and by C_1 we have denoted the following number:

$$C_1 = \frac{1}{\Delta_0} \int\limits_{\mathbb{R}} d\xi P_\xi(\xi) \int\limits_{\mathbb{R}} d\Delta P_\Delta(\Delta) \cdot \frac{\Delta}{\sqrt{\xi^2+\Delta^2}}, \tag{101}$$

where the integrand decays only as ξ^{-1} with large ξ, so that one must take into account the ξ-dependence of $P(\xi)$ with logarithmic precision. However, the numerical value of C_1 is of the order $1/g$, as it is for a similar integral for the mean field value Δ_0, given by (62).

The final step is to exploit the special homogeneous form of the saddle-point equation:

$$\frac{\Delta}{\sqrt{\xi^2+\Delta^2}} = \frac{1}{\sqrt{1+(x/y)^2}}, \tag{102}$$

so that the only remaining integral can be evaluated by a simple change of variables $z = \xi/\Delta$:

$$\begin{aligned}
... &= 1 + \left(is \cdot \frac{g}{v}\right) \cdot \Delta_0 \cdot C_1 + 2P_\xi(0) \cdot \Delta_0 C_2 \cdot F\left(\frac{g}{v}s\right) + O\left(\Delta_0^2\right) \\
C_2 &= \frac{1}{\Delta_0} \int\limits_{\mathbb{R}} d\Delta \cdot P_\Delta(\Delta) \cdot \Delta
\end{aligned} \tag{103}$$

$$F(u) = \int\limits_0^\infty dz \left(\exp\left\{ iu \frac{1}{\sqrt{1+z^2}} \right\} - \left\{ 1 + iu \frac{1}{\sqrt{1+z^2}} \right\} \right) \tag{104}$$

Note that the function $F(u)$ is an entire function of complex variable s, as the integral is convergent as written for any $u \in \mathbb{C}$. Additionally, the numerical value of the constant C_2 is of the order unity, because it represents nothing but the mean value of the order parameter Δ under the given distribution P_Δ. Because we expect it to be still of the order of the mean field value Δ_0, one concludes that $C_2 \sim 1$.

By substituting $\Delta_0 = v/K$, we can finally take the limit of large K in the expression for R:

$$\begin{aligned}
&R(s) \\
&= \ln\left(\left[1 + \left(is \cdot \frac{g}{v}\right) \cdot \Delta_0 \cdot C_1 + 2P_\xi(0) \cdot \Delta_0 C_2 \cdot F\left(\frac{g}{v}s\right) + O\left(\Delta_0^2\right) \right]^K \right) \\
&= \ln\left(\exp\left\{ v \cdot \left[\left(is \cdot \frac{g}{v}\right) \cdot C_1 + 2P_\xi(0) \cdot C_2 \cdot F\left(\frac{g}{v}s\right) \right] + O(\Delta_0) \right\} \right) \\
&= v \cdot \left[\left(is \cdot \frac{g}{v}\right) \cdot C_1 + 2P_\xi(0) \cdot C_2 \cdot F\left(\frac{g}{v}s\right) \right] + O(\Delta_0)
\end{aligned} \tag{105}$$

so that we obtain the leading term of formal expansion of R in powers of Δ_0. Because Δ_0 is exponentially small and thus represents an excellent small parameter, we will henceforth disregard the correction as by no means can it deliver an important contribution to the behavior of R.

7.2.3 Analytical properties of the cumulant generating function and the probability function

First of all, let us examine the emerging expression for the probability P_Δ. The inverse Fourier transform of the result reads:

$$\begin{aligned} &P_\Delta(\Delta) \\ &= \frac{1}{\Delta_0}\int\limits_{\mathbb{R}} \frac{ds}{2\pi} \exp\left\{-is\frac{\Delta}{\Delta_0}\right\} \cdot \exp\{R(s)\} \\ &= \frac{1}{\Delta_0}\int\limits_{\mathbb{R}} \frac{ds}{2\pi} \exp\left\{-is\frac{\Delta}{\Delta_0} + v\cdot\left[\left(is\cdot\frac{g}{v}\right)\cdot C_1 + 2P_\xi(0)\cdot C_2\cdot F\left(\frac{g}{v}s\right)\right]\right\} \end{aligned} \tag{106}$$

The convergence of the integral is ensured by the asymptotic behavior of F at large real arguments:

$$F(\pm u \gg 1) = -\frac{\pi}{2}|u| - iu\cdot\{\ln 2|u| - 1 + \gamma\} \tag{107}$$

As we can see, the real part is always negative and decreases linearly, so that the integral converges at $g\cdot s \sim 1$. Moreover, it can be shown that the real part becomes negative at sufficiently large values of $|\mathrm{Re}u|$ for any positive $\mathrm{Im}u$ (see Figure 8):

$$u'' = \mathrm{Im}u > 0, u' = \pm\mathrm{Re}u \gg 1 \Rightarrow \mathrm{Re}\{F(\pm u' + iu'')\} \to -\infty \tag{108}$$

Therefore, one is free to shift the integration contour within this region. For future calculations it will be convenient to shift it to the lower half-plane:

$$P_\Delta(\Delta) = \frac{1}{\Delta_0}\int\limits_{\mathbb{R}-i0} \frac{ds}{2\pi} \exp\left\{-is\frac{\Delta}{\Delta_0}\right\} \cdot \exp\{R(s)\} \tag{109}$$

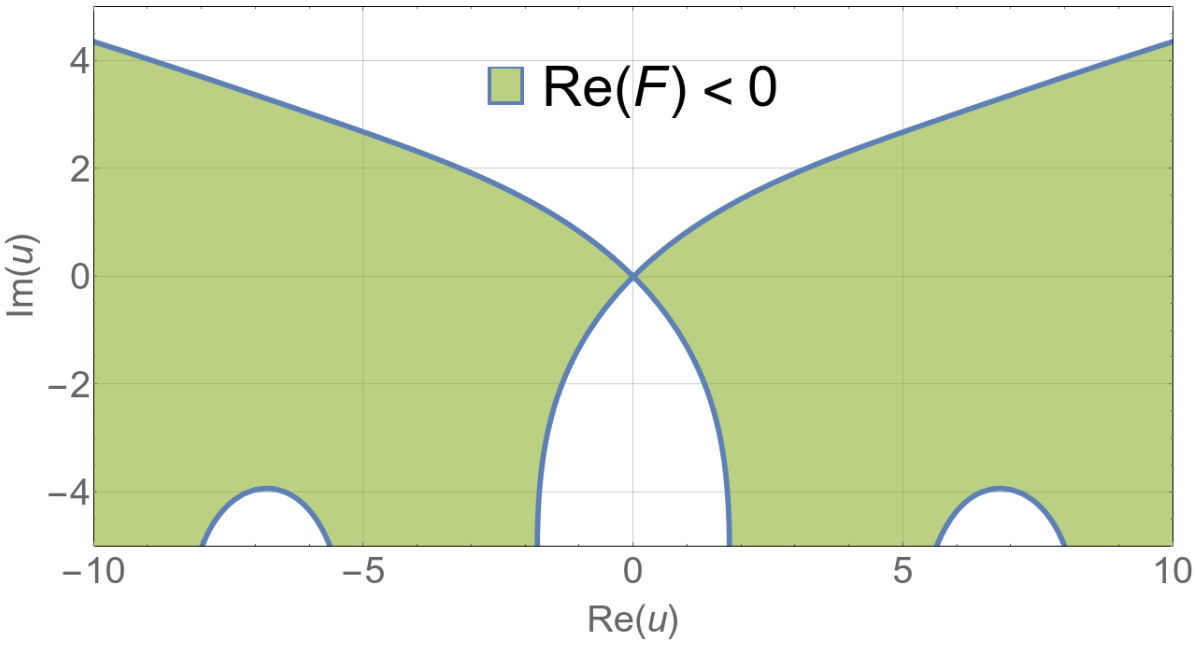


Figure 8: Visualization of the region with negative real part of $F(u)$. As long as the integration contour for an integral containing $\exp\{F(u)\}$ is in this green region, the integral is convergent in a finite neighborhood of $u = 0$.

At this point, we should yet try and restore the discussed analytical properties of $\exp\{R\}$, as they are going to be essential for further calculations. While the obtained expression represents an entire function in the complex plain of s, as it is expected from a Fourier image of a bounded probability distribution, one should consider the announced exponential decay of $\exp\{R\}$ for large s with positive imaginary part. Indeed, one can observe that the function $F(u)$ demonstrates a faster-than-linear growth for $\mathrm{Im} u \gg 1$:

$$F(u = ir, r \gg 1) \approx +r(\ln r + \gamma + \ln 2 - 1) + O\left(\frac{1}{r}\right), \tag{110}$$

so that the exponential decay of $\exp\{R\}$ provided by the C_1 term in (105) is overwhelmed, and $\exp\{R\}$ seems to be a growing function of $\mathrm{Im} s \gg 1$. The contradiction is resolved by the fact that the obtained results formally represent the limiting form of R for the case of infinitely small Δ_0. Because R is defined as the cumulant generating function for the quantity $y = \Delta/\Delta_0$, the distribution of y becomes unbounded in the considered limit. In reality, however, the value of Δ_0 is finite, and the discussed exponential decay is indeed observed at $\mathrm{Im} s \gtrsim s_0 = \Delta_0^{-1} g^{-1} \gg 1$, but the associated scale s_0 is exponentially large. The matching of our result to this asymptotic behavior occurs via the fact that at $|s| \gtrsim s_0$ our calculation is rendered inapplicable because for such large s the exact integrals over ξ converge at scale $(\ln s)^{-1}$ which now becomes of the order of unity, thus forbidding us to neglect the exact form of the P_ξ function. Therefore, we conclude that the obtained expression is valid for $|s| \lesssim s_0$, while at larger values of s it decays exponentially in the upper half-plane.

The most essential consequence of this result is that whenever one deals with integrals over s that do not involve functions which grow in the upper half-plane, one is allowed to close the integrating contour in the upper half-plane and thus evaluate the integral by means of the residue technique. This corresponds to the fact that such integrals converge at $s \sim 1$ and thus receive

no information about the region of large s. As an immediate demonstration of this idea, let us calculate the normalization of the probability distribution:

$$\int_0^\infty d\Delta \cdot P_\Delta(\Delta) = \int_0^\infty d\Delta \cdot \frac{1}{\Delta_0} \int_{\mathbb{R}-i0} \frac{ds}{2\pi} \exp\left\{-is\frac{\Delta}{\Delta_0}\right\} \cdot \exp\{R(s)\}$$

Because $\mathrm{Im} s < 0$ by our choice of the contour, the integral over Δ is trivially evaluated:

$$... = \int_{\mathbb{R}-i0} \frac{ds}{2\pi} \cdot \frac{1}{is} \cdot \exp\{R(s)\}$$

And because of the discussed property of the R function, we can close the integration in the upper half-plane, because the integrand $1/s$ decays at infinity. The result is then trivially evaluated as a residue at the only singularity of the integral:

$$... = \int_C \frac{ds}{2\pi} \cdot \frac{1}{is} \cdot \exp\{R(s)\} = \operatorname*{res}_{s=0} \frac{\exp\{R(s)\}}{s} = \exp\{R(0)\} = 1,$$

where the last equality follows from the normalization of R

7.2.4 Physical sense and self-consistent equations on the constants C_1 and C_2

Having the described properties of R in mind, we can calculate the value of C_2 explicitly from its definition.

$$\begin{aligned} C_2 &= \frac{1}{\Delta_0} \cdot \int_0^\infty d\Delta \cdot P(\Delta) \cdot \Delta \\ &= \frac{1}{\Delta_0^2} \cdot \int_0^\infty d\Delta \cdot \left[\int_{\mathbb{R}-i0} \frac{ds}{2\pi} \cdot \exp\left\{-is \cdot \frac{\Delta}{\Delta_0}\right\} \cdot \exp\{R(s)\} \right] \cdot \Delta \\ &= \frac{ds}{2\pi} \cdot \frac{\exp\{R(s)\}}{(is)^2} = i \cdot \operatorname*{res}_{s=0} \left\{ \frac{\exp\{R(s)\}}{(is)^2} \right\} = \left\{ \frac{dR(s)}{ids} \right\}_{s=0}, \end{aligned} \tag{111}$$

where we used the analytical properties of R to close the contour and normalization of R to calculate the residue. The final step is to note that the expansion of $F(u)$ in powers of u starts with $O(u^2)$ by construction, so that the derivative is given solely by the C_1 term of R:

$$C_2 = g \cdot C_1 \tag{112}$$

The result is of no surprise, as the constants $C_{1,2}$ represent the mean values of the corresponding sides of the saddle-point equation (up to obvious multipliers):

$$g \cdot C_1 = \frac{1}{\Delta_0} \cdot g \cdot \left\langle \frac{\Delta}{\sqrt{\Delta^2 + \xi^2}} \right\rangle = \frac{\langle \Delta \rangle}{\Delta_0} = C_2 \tag{113}$$

Yet the results represents a decent demonstration of the technique that is later going to be extensively used to evaluate similar integrals.

It remains to determine C_1, which is to be found self-consistently from its definition:

$$C_1 = \frac{1}{\Delta_0} \int\limits_{\mathbb{R}} d\xi P_\xi(\xi) \int\limits_{\mathbb{R}} d\Delta P_\Delta(\Delta) \cdot \frac{\Delta}{\sqrt{\xi^2 + \Delta^2}}$$

As we already know, the integral over ξ is logarithmic. To extract this logarithmic behavior, one adds and subtracts the value of the integrand for $\Delta = \Delta_0$:

$$\begin{aligned} ... &= \frac{1}{\Delta_0} \int\limits_{\mathbb{R}} d\xi P_\xi(\xi) \int\limits_{\mathbb{R}} d\Delta P_\Delta(\Delta) \cdot \frac{\Delta}{\sqrt{\xi^2 + \Delta_0^2}} \\ &+ \frac{1}{\Delta_0} \int\limits_{\mathbb{R}} d\xi P_\xi(\xi) \int\limits_{\mathbb{R}} d\Delta P_\Delta(\Delta) \cdot \left[\frac{\Delta}{\sqrt{\xi^2 + \Delta^2}} - \frac{\Delta}{\sqrt{\xi^2 + \Delta_0^2}} \right] \end{aligned}$$

Now the first integral can be recognized to contain the zero-temperature version of mean-field integral 62, and the second one gains at the small region of $\xi \sim \Delta$, so that the dependence of P_ξ can again be neglected:

$$\begin{aligned} ... &= C_2 \cdot \int\limits_{\mathbb{R}} d\xi P_\xi(\xi) \frac{1}{\sqrt{\xi^2 + \Delta_0^2}} + 2P_\xi(0) \cdot \frac{1}{\Delta_0} \int\limits_{\mathbb{R}} d\Delta P_\Delta(\Delta) \cdot \Delta \cdot \ln \frac{\Delta_0}{\Delta} \\ &= /C_2 = g \cdot C_1/ = gC_1 \int\limits_{\mathbb{R}} d\xi P_\xi(\xi) \frac{1}{\sqrt{\xi^2 + \Delta_0^2}} + 2P_\xi(0) \cdot \frac{1}{\Delta_0} \int\limits_{\mathbb{R}} d\Delta P_\Delta(\Delta) \cdot \Delta \cdot \ln \frac{\Delta_0}{\Delta} \end{aligned} \quad (114)$$

The resulting equation represents an equation on C_1 that enters the equation both explicitly and via the expression for P_Δ. One should treat this equation as a self-consistency condition to find the value of C_1. This can be done by means of numerical evaluation of the integral for P_Δ and subsequent numerical solution of the equation for C_1.

At this point it is convenient to make our choice of Δ_0, whereas up to now we were concerned only with the numerical of smallness of Δ_0 rather than its specific value. A natural choice is to use the mean field definition itself, i. e.

$$\int\limits_{\mathbb{R}} d\xi P_\xi(\xi) \frac{1}{\sqrt{\xi^2 + \Delta_0^2}} := \frac{1}{g} \quad (115)$$

Substitution of the value of C_2 then leads to reduction of the first term, finally yielding the following self-consistency equation:

$$0 = \int\limits_{\mathbb{R}} d\Delta P_\Delta(\Delta) \cdot \frac{\Delta}{\Delta_0} \cdot \ln \frac{\Delta_0}{\Delta} \quad (116)$$

or, plugging in the expression (106) for P_Δ via R,

$$\begin{aligned} 0 &= \int\limits_{\mathbb{R}-i0} \frac{ds}{2\pi} \cdot \frac{\ln\{is\} + \gamma - 1}{(is)^2} \cdot \exp\{R(s)\} \\ &= \int\limits_{\mathbb{R}-i0} \frac{ds}{2\pi} \cdot \frac{\ln\{is\} + \gamma - 1}{(is)^2} \cdot \exp\left\{ gC_1 \cdot \left[is + 2P_\xi(0)\, v \cdot F\left(\frac{g}{v} s\right) \right] \right\} \end{aligned} \quad (117)$$

Because C_1 enters the definition of P_Δ, one can use the equations above to find the value of C_1 numerically.

7.2.5 Revision of results and main technical features

Let us revise the results and the main ingredients that led to a possibility of an explicit solution. One obtains the following expression for the probability function:

$$P_{\Delta}\left(\Delta\right)=\frac{1}{\Delta_0}\int\limits_{\mathbb{R}-i0}\frac{ds}{2\pi}\exp\left\{-is\frac{\Delta}{\Delta_0}+gC_1\cdot\left[is+2P_{\xi}\left(0\right)\cdot v\cdot F\left(\frac{g}{v}s\right)\right]\right\}, \tag{118}$$

where Δ_0 is the mean field zero-temperature order parameter defined as a solution to the following equation:

$$\int\limits_{\mathbb{R}}d\xi P_{\xi}\left(\xi\right)\frac{1}{\sqrt{\xi^2+\Delta_0^2}}:=\frac{1}{g}, \tag{119}$$

whereas the constant v and the function F are given by

$$v=K\Delta_0$$

$$F\left(u\in\mathbb{C}\right)=\int\limits_0^{\infty}dz\left(\exp\left\{iu\frac{1}{\sqrt{1+z^2}}\right\}-\left\{1+iu\frac{1}{\sqrt{1+z^2}}\right\}\right)$$

The integral (118) is convergent for any contour parallel to the real axis in the lower half-plane, which is ensured by asymptotic properties of F. The value of the constant C_1 should be determined as a solution to the following self-consistency equation:

$$0=\int\limits_{\mathbb{R}}d\Delta P_{\Delta}\left(\Delta\right)\cdot\frac{\Delta}{\Delta_0}\cdot\ln\frac{\Delta_0}{\Delta} \tag{120}$$

The solution presented rests on 3 main ideas:

1. The r.h.s. of the saddle-point equation (60) decays as ξ^{-1} at large ξ, so that its mean value w.r.t ξ is logarithmic in Δ_0. The resulting value of the mean field order parameter Δ_0 is exponentially small.

2. All higher powers of the r.h.s of saddle-point equation decay with ξ faster than ξ^{-1}, so that integrals of such quantities gain their value in a small region of $\xi\sim\Delta_0$. This allows the existence of the universal limit where the particular structure of the disorder distribution is rendered irrelevant.

3. The r.h.s. of the zero-temperature saddle-point equation (60) represents a homogeneous function of a single parameter $z=\xi/\Delta$. This provides an opportunity to decouple the connection between ξ and Δ in the equation and evaluate the probability distribution explicitly.

As a result of our consideration, we can observe that there exists a well defined field theory limit $\Delta_0\to 0$, $K\to\infty$ while $v=K\Delta_0=\text{const}$. Below we summarize the main features of this limiting regime:

1. In the regime of small Δ_0 and large K all physics is completely determined by 4 parameters:

   - Mean field value of the order parameter Δ_0

- Zero energy density of states of the disorder $\nu_0 = 2P_\xi(0)$
- Cooper coupling constant g
- Effective number of neighbors for a given site $v = K\Delta_0$

2. There's a natural rescaling of the physical quantities in the problem:

$$x = \frac{\xi}{\Delta_0}, y = \frac{\Delta_i}{\Delta_0}, \tag{121}$$

so that there exist a well defined limiting distribution function $P(y)$.

3. The form of the distribution function $P(y)$ is governed by the behavior of the generating function F of higher moments of the r.h.s of the saddle-point equation as well as the constant C_1 which is found self-consistently from the equation

$$\left\langle y \ln \frac{1}{y} \right\rangle_y := \int_0^\infty dy \cdot P(y) \cdot y \ln \frac{1}{y} = 0 \tag{122}$$

The physical sense of C_1 is to represent the actual mean value of the order parameter

$$\langle \Delta \rangle = \Delta_0 \cdot gC_1 \tag{123}$$

4. Averaging of various functions of Δ is made considerably easier due to analytical properties of the cumulant generating function

$$R(s) = \ln \left\langle \exp\left\{ -is\frac{\Delta}{\Delta_0} \right\} \right\rangle_\Delta, s \in \mathbb{C} \tag{124}$$

- It possesses universal limiting behavior in the region $|s| \lesssim s_0 = g^{-1}\Delta_0^{-1} \gg 1$, where it is know to lack any singularities.
- Outside the universal region the function $\exp\{R(s)\}$ is known to be an exponentially decaying function for $\mathrm{Im} s \to +\infty$.
- By means of Fourier transform, all integrals with weight P_Δ can be transformed into integrals over a suitable contour in the complex plane of s with the weight $\exp\{R(s)\}$. The latter is known to ensure exponential convergence of all integrals within the region $s \sim g^{-1}$. The only limitation on the integration contour is to have its ends be in the lower half-plane.
- Whenever the integrand of the aforementioned integrals over s decays at infinity, one is allowed to close the integration contour in the upper half-plane, thus allowing for explicit evaluation of all integrals via the residue technique.

7.2.6 Qualitative properties of the obtained distribution and comparison with the exact distribution

On Figure 9 we present a comparison of the developed model with the exact numerical solution of the saddle-point equation (60). One immediate observation is that both the model and the exact solution demonstrate a nontrivial body of the distribution. This confirms that below $v = 1$

statistics of the order parameter demonstrates nontrivial interplay between the tendency to the self-averaging as a result of large K and broad distribution of the r.h.s of the saddle-point equation due to the smallness of Δ.

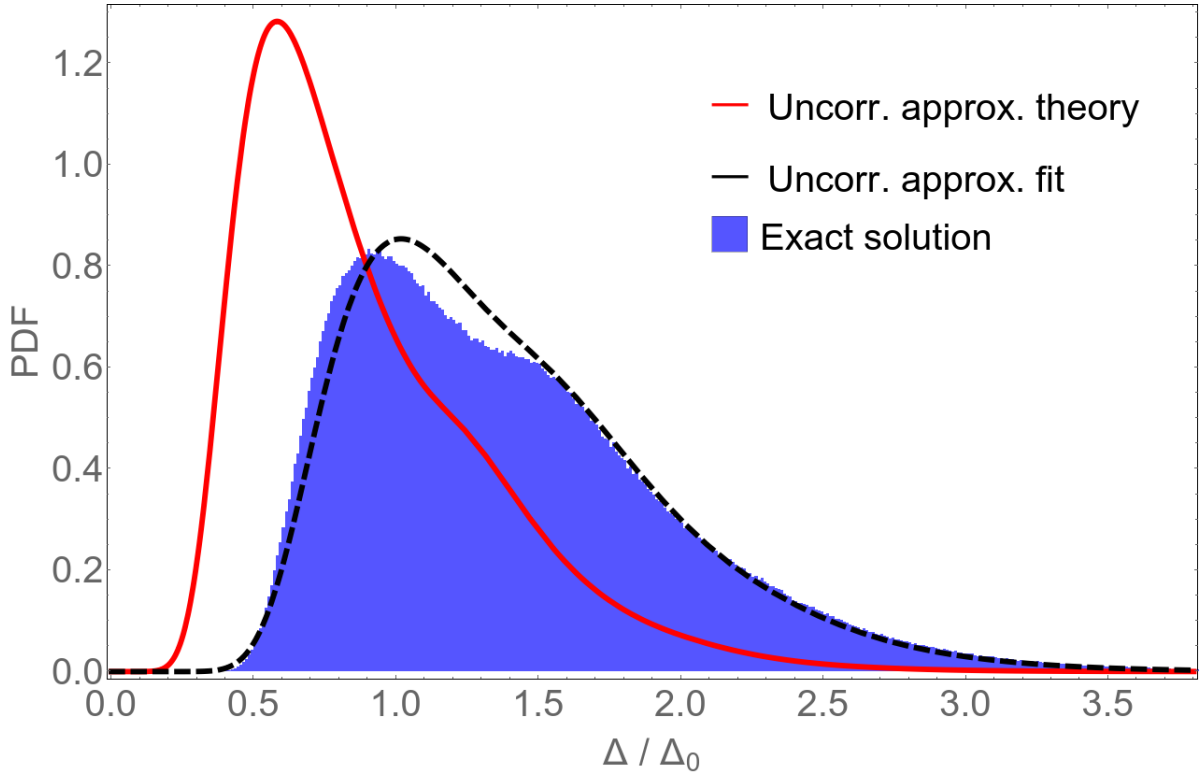


Figure 9: Plot of the distribution function of the dimensionless order parameter Δ/Δ_0. Parameters of the model: disorder has box-shaped distribution in the interval $[-1, 1]$, $g = 0.16$, $K = 50$, resulting in $\nu_0 = 1$, $v = 0.193$. The blue histogram represents statistics of the exact numerical solution of the self-consistency equation on a random regular graph with $N = 2^{22} \approx 4.2 \cdot 10^6$. The solid red line plots the prediction given by (118) with $C_1 = 5.620$ as found by solving (120). The dashed line corresponds to uncorrelated approximation with $C_1 = 8.727$ found directly from the mean value of numerical data (see the main text for an extended explanation).

Next obvious comment is that the model only succeeds to reproduce the qualitative form of the plot, while failing to deliver any quantitative convergence. However, on the same plot we also present the improved version of the model where the value of C_1 is replaced with the value that is obtained by using the connection between C_1 and the mean value (123). Indeed, instead of solving the equation (120) we can directly calculate the mean value from the numerical data and then use it to restore the value of C_1. While, in theory, those two procedure should evaluate to the same result, one observes a significant discrepancy between the two approaches. Using the "numerical" value of C_1 instead of the theoretical one then yields a much better quantitative agreement with the numerical data (dashed line on Figure 9), although still containing a noticeable discrepancy. As we will see later, both the discussed mismatch in the values of C_1 and the residual discrepancy after using the correct value of C_1 are owing to the cross-site correlations and are, therefore, beyond the scope of this model.

Let us also discuss the profile of the distribution at extreme (either large or small) values which is qualitatively reproduced by the amended approximation (i.e., with the corrected value

of C_1) and allegedly represents a robust feature of the distribution. At this point, we employ a simple saddle-point analysis for integral 118 to extract the large- and small-Δ tails of the probability density function. One can show that the leading contribution is always given by a saddle point at imaginary axis, where it can be seen that

$$F\left(u=ir\right),\left(i\cdot\mathrm{sgn}r\right)\cdot F'\left(u=ir\right),-F''\left(u=ir\right)>0,$$

so that exploring the saddle-point approximation of integral 118 amounts to solving a simple equation:

$$g\nu_0\cdot(-i)\,F'\left(u_0\right)=\frac{\Delta}{\Delta_0\cdot gC_1}-1\equiv\frac{\Delta}{\langle\Delta\rangle}-1, iu_0\in\mathbb{R} \tag{125}$$

And the estimation to the probability is then expressed via u_0 as

$$P_\Delta\left(k=\frac{\Delta}{\langle\Delta\rangle}\right)=\frac{1}{\Delta_0}\cdot\frac{1}{\sqrt{2\pi A\cdot\left[-\left(\frac{g}{v}\right)^2F''\left(u_0\right)\right]}}\exp\left\{A\cdot\left[F\left(u_0\right)-iu_0\frac{k-1}{\nu_0 g}\right]\right\} \tag{126}$$

$$A=gC_1\cdot\nu_0 v \tag{127}$$

Without going into further technical detail, we present the resulting asymptotic expressions for P:

$$P\left(0<k\lesssim 1\right)\approx\frac{1}{\Delta_0}\cdot\frac{\exp\left\{-A\cdot r_{\mathrm{small}}\right\}}{\sqrt{2\pi A\cdot\left(\frac{g}{v}\right)^2\cdot r_{\mathrm{small}}^{-1}}} \tag{128}$$

$$r_{\mathrm{small}}\approx\exp\left\{-\left(\frac{k-1}{\beta}+\ln 2+\gamma\right)\right\} \tag{129}$$

$$P_\Delta\left(k\gtrsim 1\right)\approx\frac{1}{\Delta_0}\cdot\frac{\exp\left\{-A\left(\left\{r_{\mathrm{large}}-1\right\}\left[\frac{k-1}{\beta}+\left(\gamma+\ln 2r_{\mathrm{large}}\right)\right]-r_{\mathrm{large}}\right)\right\}}{\sqrt{2\pi A\cdot\left(\frac{g}{v}\right)^2\cdot\left[\frac{k-1}{\beta}+\left(\gamma+\ln 2r_{\mathrm{large}}\right)\right]}} \tag{130}$$

$$r_{\mathrm{large}}=\ln\left[\sqrt{\frac{2}{\pi}}\left(\frac{k-1}{\beta}+\ln 2+\gamma\right)\right]+O\left(\ln\ln\left\{\frac{k-1}{\beta}\right\}\right) \tag{131}$$

The results of this approximation are shown on Figure 10.

The final property to be mentioned is the interpolation of the considered model to the Gaussian regime. While we can see that the model is only qualitatively correct in the regime of intermediate v, it demonstrates a proper quantitative matching with the large K regime. Indeed, the characteristic scale of integral 118 for the probability is in general case of the order unity due to the first terms. Therefore we can expand F in a power series:

$$F\left(\frac{g}{v}s\right)\approx\frac{1}{2}F''\left(0\right)\cdot\left[\frac{g}{v}s\right]^2=-\frac{1}{2}\cdot\frac{\pi}{2}\cdot\left[\frac{g}{v}s\right]^2,$$

which makes the integral easily computable:

$$\begin{aligned}&P_\Delta\left(\Delta\right)\\&\approx\frac{1}{\Delta_0}\int\limits_{\mathbb{R}-i0}\frac{ds}{2\pi}\exp\left\{-is\frac{\Delta}{\Delta_0}+gC_1\cdot\left[is-\nu_0 v\cdot\frac{1}{2}\frac{\pi}{2}\cdot\left[\frac{g}{v}s\right]^2\right]\right\}\\&=\frac{1}{\Delta_0 gC_1}\cdot\frac{1}{\sqrt{2\pi}\sqrt{\frac{\pi}{2}\frac{\nu_0 g^2}{v\cdot gC_1}}}\exp\left\{-\frac{1}{2}\left[\frac{\frac{\pi}{2}\nu_0 g^2}{v\cdot gC_1}\right]^{-1}\left(\frac{\Delta}{\Delta_0 gC_1}-1\right)^2\right\}\end{aligned} \tag{132}$$

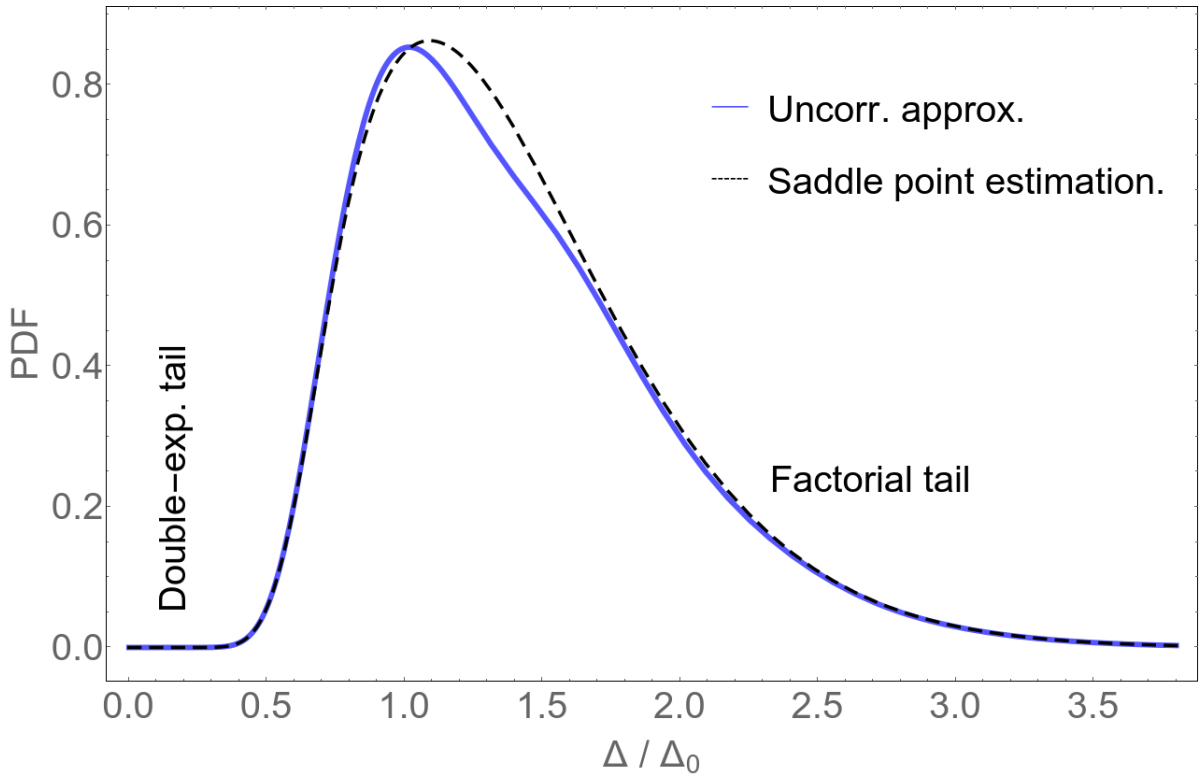


Figure 10: Comparison of the theoretical curve from 9 with experimental value of $C_1 = 8.727$ with the saddle point approximation for the probability integral (118). The blue line is the result of numerical integration. The dashed black line is the result of the saddle-point approximation (125)-(127). In the region of large $\Delta \gtrsim \langle\Delta\rangle$ the result is described by factorial dependence of the form $\exp\{A \cdot x \ln x\}$, see (128)-(129). In the region of small $\Delta < \langle\Delta\rangle$ the result decays at double-exponential rate $\exp\{A \exp\{x_0 - x\}\}$, see (130)-(131).

Plugging this approximate expression into the equation (120) for C_1 is also rather straightforward, as in the limit of large $v \propto K$ the characteristic width of the function P_Δ becomes small, so that the resulting expression simply reads

$$\left\langle y \ln \frac{1}{y} \right\rangle_y \approx gC_1 \cdot \ln \frac{1}{gC_1} = 0 \Leftrightarrow gC_1 = 1 \tag{133}$$

and we arrive at

$$P_\Delta(\Delta) = \frac{1}{\Delta_0 g C_1} \cdot \frac{1}{\sqrt{2\pi}\sqrt{\frac{\frac{\pi}{2}\nu_0 g^2}{v}}} \exp\left\{ -\frac{1}{2} \left[\frac{\frac{\pi}{2}\nu_0 g^2}{v} \right]^{-1} \left(\frac{\Delta}{\Delta_0} - 1 \right)^2 \right\}, \tag{134}$$

which is a Gaussian distribution with mean value Δ_0 and variance

$$\sigma^2 = \Delta_0^2 \cdot \frac{\pi}{2} \frac{\nu_0 g^2}{v} = \Delta_0^2 \cdot \frac{\pi}{2} \cdot 2P_\xi(0) \cdot \frac{g^2}{K\Delta_0}, \tag{135}$$

which is precisely the leading term (79) for the variance in the large K approximation. In this way we establish an explicit connection to the results of the previous section.

Let us summarize our key observations about the uncorrelated approximation that we developed:

- The approximation sheds some light on the origin and characteristic scales of the regime of intermediate value of K. In particular, it manages to qualitatively reproduce the nontrivial main body of the order parameter distribution. The observed distribution profile results from a balance between tendency towards self-averaging due to large value of K and strength of fluctuations of the r.h.s of the saddle-point equation stemming from the smallness of Δ_0.

- However, the developed approximation resembles the actual picture only when the mean value of the order parameter is provided. The self-consistency equation following from the theory itself describes the mean value inaccurately. Later we will see this to be the consequence of neglected inter-site correlations.

- Yet when the value of C_1 is provided, one can describe the tails of the distribution of the order parameter. From this prediction one deduces that the value of Δ significantly larger than the mean value are suppressed with a factorial rate (130), while small values of Δ are double-exponentially rare (128). Although the exact from of the tail demonstrates some discrepancy with these predictions, this approach provides a good intuition on the behavior of the true distribution.

- Finally, we can observe a clear transition of the described model to the large K regime, both in the form of the distribution itself and in the self-consistent equation on C_1.

8 Distribution of the order parameter

In this section we will present the theory describing the full statistics of the order parameter as found by solving the zero temperature saddle-point equation 60. We start with a set of leading examples and considerations that allow us to grasp the key intuition and the remaining ideas necessary to derive the full theory. The second section presents two integral equations on auxiliary functions r, $r_1^{(0)}$ and a set of formulae to express the probability distribution of the order parameter via these functions. As it can be shown, other physical quantities can also be expressed via r. An approximate scheme for an analytical solution of the aforementioned equations is then proposed. The results of these calculations are then compared to the numerical solution of the saddle-point equation with excellent agreement in place.

8.1 Description of mutual correlations

8.1.1 Qualitative importance of mutual influence

The main simplification of the model discussed previously in Subsection 7.2 was to completely neglect the mutual influence of both ξ and Δ fields on neighboring sites. Below we will present an instance of important effects omitted in this simplified model. In what follows, we will assume $v \lesssim g$, which corresponds to the nontrivial regime well beyond a simple Gaussian approximation discussed previously.

Suppose a given site i has its value of ξ_i abnormally small, within the range of several Δ_0. The naive approximation of uncorrelated sites predicts no effects in the onsite statistics of Δ whatsoever, as fields Δ and ξ are completely uncorrelated within that model. However, let us consider the effect of this small value of ξ for the neighboring sites.

Because appearance of $\xi \sim \Delta_0$ is an exponentially rare event (characteristic width of ξ distribution is of order unity), one can estimate a fraction of neighbors of site j that have this kind of smallness in the value of ξ:

$$n \sim 2P_\xi(0) \cdot \Delta_0,$$

which means that out of $K+1$ neighbors there are only

$$N_s \sim K \cdot \Delta_0 \cdot 2P_\xi(0) = v \cdot 2P_\xi(0) \tag{136}$$

sites that would have the value of ξ within the similar range. Given $v \sim 1$, it is evident that a given site typically neighbors only about a single other site with small enough values of ξ. However, this appears to be sufficient to produce a significant effect.

For each of the neighbors of such a site i with a decreased value of ξ the saddle-point equation reads:

$$j \in \partial i : \Delta_j = \frac{g}{K} \sum_{k \in \partial j \setminus i} \frac{\Delta_k}{\sqrt{\Delta_k^2 + \xi_k^2}} + \frac{g}{K} \cdot \frac{\Delta_i}{\sqrt{\Delta_i^2 + \xi_i^2}} \tag{137}$$

Let us estimate the magnitude of each term in the r.h.s of this equation. Assuming that all neighbors of j other than i have their value of ξ of the order unity, one estimates the first term to be of the order given by the mean field approximation:

$$\frac{g}{K} \sum_{k \in \partial j \setminus i} \frac{\Delta_k}{\sqrt{\Delta_k^2 + \xi_k^2}} \sim \Delta_0$$

The second term, however, has $\xi \sim \Delta_0$, so that for a typical value of the onsite order parameter Δ_i, the characteristic value of the second term reads:

$$\frac{g}{K} \cdot \frac{\Delta_i}{\sqrt{\Delta_i^2 + \xi_i^2}} \sim \frac{g}{K} \cdot 1 \sim \frac{g}{K\Delta_0} \cdot \Delta_0 \sim \frac{g}{v} \cdot \Delta_0 \tag{138}$$

This result tells us that whenever $v \lesssim g$, the considered effect of small ξ will lead to a noticeable enhancement in the value of Δ on neighboring sites, i. e.

$$\Delta_j \sim k\Delta_0, k \sim 1 + \frac{g}{v} \tag{139}$$

Provided that v described the typical fraction of the nearest neighbors of each site that have a suppressed value of ξ, one arrives at the fact that for $v \sim g$ there exists a fraction of $v \sim g$ of sites that do have such a neighbor and thus themselves experience an enhancement in the value of Δ.

Now let us get back to the initial site i with a decreased value of ξ. As we have observed, all neighbors of i will have their value of Δ increased by a noticeable fraction k. One can then examine the associated behavior of the saddle-point equation for site i:

$$\Delta_i = \frac{g}{K} \sum_{j \in \partial i} \frac{\Delta_j}{\sqrt{\Delta_j^2 + \xi_j^2}} \sim g \cdot \left\langle \frac{k\Delta_0}{\sqrt{(k\Delta_0)^2 + \xi^2}} \right\rangle_\xi \sim \Delta_0 \cdot \ln k \sim \Delta_0 \ln\left(1 + \frac{g}{v}\right), \tag{140}$$

where we estimated the typical amplitude of the r.h.s via its average value, which is applicable as long as neither of ξ_j is comparable to Δ_0. The result implies that provided the small onsite value

of ξ the value of Δ_i will be larger by a factor of $\ln\left(1+\frac{g}{v}\right)$ which can be a number significantly different from unity.

Formulated mathematically, this means the joint distribution $P(\Delta,\xi)$ of Δ on the given site will experience a noticeable deformation in the region $\xi \sim \Delta_0$. And indeed, statistics of the exact solution of the zero-temperature saddle-point equation (60) reveal that conditional distribution $P(\Delta|\xi) = P(\Delta,\xi)/P_\xi(\xi)$ possesses a nontrivial dependence on ξ in the region $\xi \sim \Delta$, where it demonstrates an enhanced value of Δ. This behavior is illustrated on Figure 11.

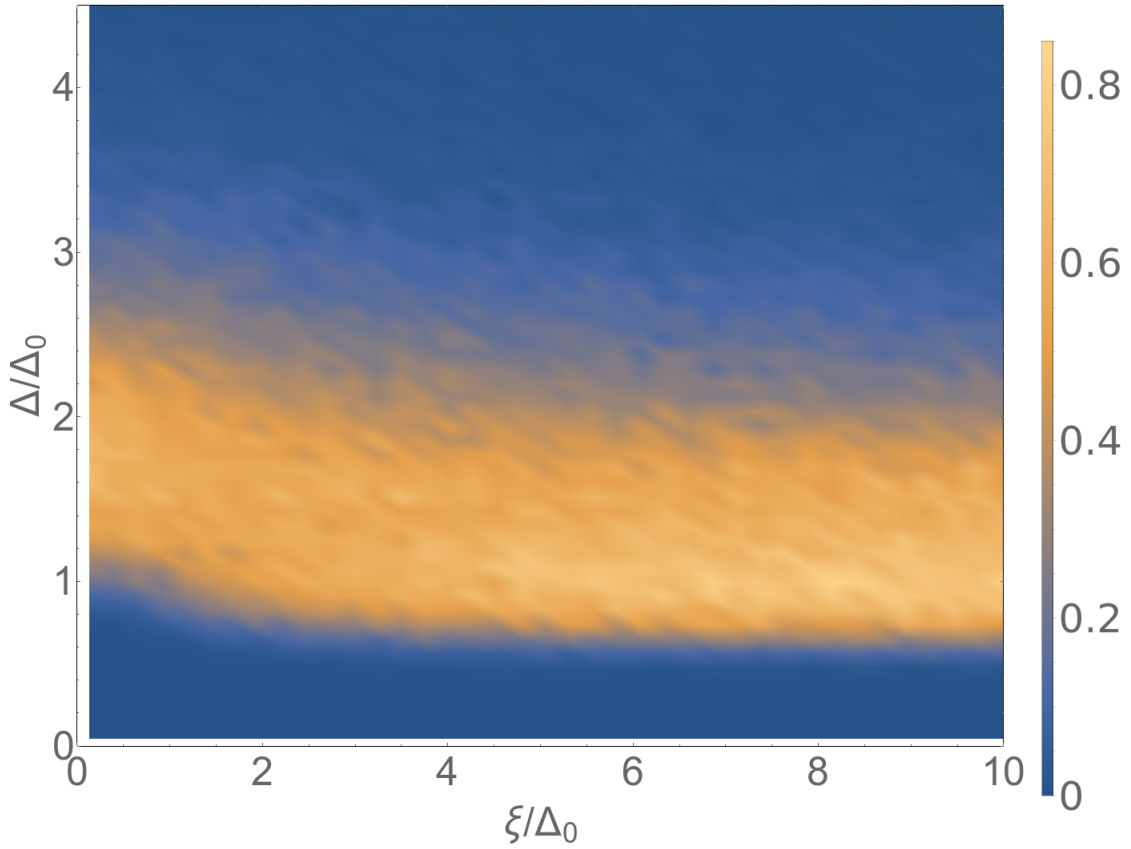


Figure 11: Density plot of the joint probability distribution $P(\Delta,\xi)$. The values are multiplied by $P_\xi^{-1}(\xi)\cdot\Delta_0^{-2}$, to reflect the data in a small region $\Delta \sim \xi \sim \Delta_0$. The color shows the magnitude of the resulting quantity, according to the legend on the right. Parameters of the simulation: $g = 0.16$, $K = 50$, total system size $N = 2^{22} \approx 4.2\cdot 10^6$. The perceived "noise" in the data is due to a finite sample size, as we are zooming in to a very small region of the ξ values, so that there are only $2\Delta_0 \cdot N \sim 3\cdot 10^4$ contributing to the areas showed on the plot.

For our consideration this conclusion is crucial for the following reason. As we have learned from the uncorrelated approximation, it is the region of $\xi, \Delta \sim \Delta_0$ that defines the quantitative behavior of the the whole distribution of Δ, as all integrals describing the probability distribution P_Δ gained their main contribution precisely in that region. And now we observe that a seemingly weak effect of a decreased value of ξ back-propagates via second order correlations and significantly deforms the joint distribution $P(\xi,\Delta)$ precisely in that region, so that its weakness due to central limit theorem is fully compensated by the smallness of the relevant region.

Of course, one can come up with a dozen of similar considerations, all of each will demonstrate that exponentially rare (with probability $\sim \Delta_0 \ll 1$) fluctuations play an essential role in the

formation of the whole distribution and lead to a much less trivial description even in the limit $\Delta_0 \to 0, K \to \infty, v = K\Delta_0 = \text{const}$, that we are going to capture below.

8.1.2 Motivating example: a chain of sites

With the help of a simple example that follows, we will present the main idea that allows to handle the full structure of mutual correlations, which is now evident to be essential for all features of the order parameter distribution in the region $v \lesssim g$.

Consider a chain of adjacent sites $i = \overline{1,N}$, with N being the total length of the chain. Each site hosts an independent fluctuating field x_i and a slave real scalar field $y_i \in \mathbb{R}$ defined by a set of coupled equations:

$$y_n = f_{n\to n+1}\left(x_{n+1}, y_{n+1}\right) + f_{n\to n-1}\left(x_{n-1}, y_{n-1}\right), n \in \overline{2, N-1} \tag{141}$$

When the site is at the edge of the chain, i. e. $i = 1$ or $i = N$, we will assume that the corresponding value of r.h.s is manually specified:

$$y_1 = f_{1\to 2}\left(x_2, y_2\right) + f_{0\to 1}, y_N = f_{N\to N+1} + f_{N\to N-1}\left(x_{N-1}, y_{N-1}\right), \tag{142}$$

where the numbers $f_{0\to 1}, f_{N\to N+1} \in \mathbb{R}$ are assumed to be known.

Let us also assume that there exists a unique stable solution to this system (e.g. it is given by a local minimum of a functional). We will denote this solution as

$$y_i = S_i^{(N)}\left(\{x_i\}_{i=1}^N | f_{0\to 1}, f_{N\to N+1}\right) \tag{143}$$

Note that all N functions $S_i^{(N)}, i \in \overline{1,N}$ depend non-trivially on the values of all independent fields x_i as well as the boundary values of f.

We will now exploit the locality of the equations. Let us cut the system along the connection $i_0 \leftrightarrow i_0 + 1$ for some $i_0 \in \overline{1, N-1}$ and formally treat the system as a set of 2 systems with consistent boundary values of f described by a separate equation:

$$\begin{cases} f_{i_0\to i_0+1} = f_{i_0\to i_0+1}\left(x_{i_0+1}, y_{i_0+1}\right) \\ f_{i_0+1\to i_0} = f_{i_0+1\to i}\left(x_{i_0}, y_{i_0}\right) \end{cases}, \tag{144}$$

whereas the values of y_i are given by

$$\begin{cases} y_i = S_i^{(i_0)}\left(\{x_i\}_{i=1}^{i_0} | f_{0\to 1}, f_{i_0\to i_0+1}\right), & i \in \overline{1, i_0} \\ y_i = S_i^{(N-i_0)}\left(\{x_i\}_{i=i_0}^{N} | f_{i_0\to i_0+1}, f_{N\to N+1}\right), & i \in \overline{i_0+1, N} \end{cases} \tag{145}$$

The key observation is that because the fields x_i are independent (i.e., one does not need to exchange any kind of information across the introduced cut to generate x_i), this system of equations is in all senses equivalent to the original system (141)-(142), i.e., they have the same values of y_i for each disorder realization as well as identical distribution of the original disorder fields x_i, so that all statistical properties of y_i are also identical.

Let us now explore the statistical properties of the solution to the new problem. In order to reflect the decoupling that occurs after we introduce the cut, let us consider the joint probability distribution for the fields on the sites $i_0, i_0 + 1$ on the edge of the cut:

$$\begin{aligned} &P_{i_0,i_0+1}^{(N)}\left(X_{i_0}, Y_{i_0}, X_{i_0+1}, Y_{i_0+1} | f_{0\to 1}, f_{N\to N+1}\right) \\ &= \left\langle \delta\left(Y_{i_0} - S_{i_0}^{(N)}\right) \delta\left(Y_{i_0+1} - S_{i_0+1}^{(N)}\right) \delta\left(X_{i_0} - x_{i_0}\right) \delta\left(X_{i_0+1} - x_{i_0+1}\right)\right\rangle_x, \end{aligned} \tag{146}$$

where the average is taken over all configurations of the independent fields x_i. The arguments of $S^{(N)}$ are omitted for brevity. Note that the joint probability also depends on externally specified values of the r.h.s $f_{0\to 1}, f_{N\to N+1}$ at the boundary.

Next step is to rewrite the δ-functions involving the full solution $S^{(N)}$ via the solutions $S^{(i_0)}$, $S^{(N-i_0)}$ in the subsystems:

$$\begin{aligned}
&\delta\left(Y_{i_0}-S_{i_0}^{(N)}\right)\delta\left(Y_{i_0+1}-S_{i_0+1}^{(N)}\right)\\
&=\delta\left(Y_{i_0}-S_{i_0}^{(i_0)}\right)\delta\left(Y_{i_0+1}-S_{i_0+1}^{(N-i_0)}\right)\cdot\left|\det\begin{pmatrix}1 & -\frac{\partial S_{i_0+1}^{(N-i_0)}}{\partial Y_{i_0}}\\ -\frac{\partial S_{i_0}^{(i_0)}}{\partial Y_{i_0+1}} & 1\end{pmatrix}\right|,
\end{aligned}\tag{147}$$

where we have used a 2D generalization of the 1D expression for the δ-function with another function as an argument (sometimes referred to as the Kac-Rice formula):

$$\begin{cases}
1\text{D}: & \delta\left(\phi\left(x\right)\right)=\delta\left(x-x_0\right)\cdot\left|\phi'\left(x_0\right)\right|\\
n\text{D}: & \delta\left(\phi\left(x\right)\right)=\delta\left(x-x_0\right)\cdot\left|\det\left\{\frac{\partial\phi}{\partial x}\left(x_0\right)\right\}\right|
\end{cases}\tag{148}$$

Here both ϕ and x are some n-dimensional vectors, and x_0 is asserted to be a solution to the equation $\phi\left(x\right)=0$. And again, we assume, that this solution is unique, as we otherwise would need to take into account an additive contribution from each such solution. In our case, we choose the function ϕ to reproduce the set of connecting equations (144)-(145):

$$\begin{aligned}
&Y_{i_0,i_0+1}=S_{i_0,i_0+1}^{(N)}\left(...\right)\\
&\Leftrightarrow\phi\left(Y_{i_0},Y_{i_0+1}\right)=\begin{pmatrix}Y_{i_0}-S_{i_0}^{(i_0)}\left(\left\{x_i\right\}_{i=1}^{i_0}|f_{0\to1},f_{i_0\to i_0+1}\left(Y_{i_0+1},X_{i_0+1}\right)\right)\\ Y_{i_0+1}-S_{i_0+1}^{(N-i_0)}\left(\left\{x_i\right\}_{i=i_0}^{N}|f_{i_0\to i_0+1}\left(Y_{i_0},X_{i_0}\right),f_{N\to N+1}\right)\end{pmatrix}=0
\end{aligned}\tag{149}$$

At this point we will assume that the solution to the system is also stable in a sense that it is given by a local minimum of some function of Y_i, so that the determinant in (147) is positive for all x_i and for $y_i=S^{(N)}$. In this case, we can drop the absolute value and continue the expression (147) as:

$$\begin{aligned}
&\delta\left(Y_{i_0}-S_{i_0}^{(N)}\right)\delta\left(Y_{i_0+1}-S_{i_0+1}^{(N)}\right)\\
&=\delta\left(Y_{i_0}-S_{i_0}^{(i_0)}\right)\delta\left(Y_{i_0+1}-S_{i_0+1}^{(N-i_0)}\right)\cdot\left[1-\frac{\partial S_{i_0+1}^{(N-i_0)}}{\partial Y_{i_0}}\frac{\partial S_{i_0}^{(i_0)}}{\partial Y_{i_0+1}}\right]
\end{aligned}\tag{150}$$

Finally, we note, that the second term can be rewritten as:

$$\begin{aligned}
...=&\,\delta\left(Y_{i_0}-S_{i_0}^{(i_0)}\right)\delta\left(Y_{i_0+1}-S_{i_0+1}^{(N-i_0)}\right)\\
&-\left[\frac{\partial}{\partial Y_{i_0+1}}\left\{\theta\left(Y_{i_0}-S_{i_0}^{(i_0)}\right)+C^{(i_0)}\right\}\right]\cdot\left[\frac{\partial}{\partial Y_{i_0}}\left\{\theta\left(Y_{i_0+1}-S_{i_0+1}^{(N-i_0)}\right)+C^{(N-i_0)}\right\}\right],
\end{aligned}\tag{151}$$

where θ is the Heaviside θ-function, and $C^{(i_0)}, C^{(N-i_0)}$ can be arbitrary functions of any combination of fields within the corresponding subsystem, while not contributing to the value of the final expression. This freedom will be important for us later to secure convergence of some integrals.

Note that the full expression now represents a product of quantities, each depending on the values of x field only within one of the two subsystems (i_0) or $(N-i_0)$. We can therefore average them independently, thus obtaining the following expression for the initial joint probability (pay attention to indices and arguments of the function in the r.h.s):

$$\begin{aligned}
&P^{(N)}_{i_0,i_0+1}\left(X_{i_0},Y_{i_0},X_{i_0+1},Y_{i_0+1}|f_{0\to1},f_{N\to N+1}\right)\\
&=P^{(i_0)}_{i_0}\left(X_{i_0},Y_{i_0}|f_{0\to1},f_{i_0\to i_0+1}\left(Y_{i_0+1},X_{i_0+1}\right)\right)\\
&\times P^{(N-i_0)}_{i_0+1}\left(X_{i_0+1},Y_{i_0+1}|f_{i_0\to i_0+1}\left(Y_{i_0},X_{i_0}\right),f_{N\to N+1}\right)\\
&-\left[\frac{\partial}{\partial Y_{i_0+1}}F^{(i_0)}_{i_0}\left(...\right)\right]\cdot\left[\frac{\partial}{\partial Y_{i_0}}F^{(N-i_0)}_{i_0+1}\left(...\right)\right],
\end{aligned}\tag{152}$$

where the functions $P^{(i_0)}$, $P^{(N-i_0)}$ represent statistics of x,y in the corresponding subsystem only, and the functions F are expressed via P's as

$$F^{(i_0)}_{i_0}\left(X_{i_0},Y_{i_0}|...\right)=\int\limits_{-\infty}^{Y_{i_0}}dy_{i_0}\cdot P^{(i_0)}_{i_0}\left(X_{i_0},y_{i_0}|...\right),\tag{153}$$

i.e., they represent the cumulative distribution function. The main achievement of the equation we have just derived is that it connects statistics of the slave field y_i in the full system with that in the 2 its subsystems connected by a single edge. For instance, one can now integrate out fields of either of the 2 sites to obtain an explicit expression for the distribution of x,y in the full system via that of the subsystem. Note, however, that the probabilities $P^{(i_0)},P^{(N-i_0)}$ are by no means conditional distributions, although they do share some common qualitative properties and can be connected with each other.

We now summarize the main ideas that deliver the result:

- The 2 subsystems are only connected via the edge that is eliminated. Once it is removed, the resulting parts are totally independent.
- The independent disorder field x is uncorrelated, allowing one to freely manipulate the structure of the underlying chain without distorting the ensemble of the disorder in each subsystem.
- The fields y are scalar, which allows one to make the final step (151)
- The solution to the equations (141)-(142) is assumed to be unique and locally stable.

Since these properties are sufficient to achieve the described decoupling, one can immediately come up with plenty of generalizations.

8.1.3 Generalization to the case of RRG

Let us now consider the immediate generalization of a simple system (141)-(142):

$$y_i=\sum_{j\in\partial i}f_{i\to j}\left(x_j,y_j\right),\tag{154}$$

where the summation in the r.h.s now is performed over all neighbors $j\in\partial i$ of a given site i on a given base graph (which is going to be a random regular graph in our case).

The developed above idea can be extended to exclude any small subgraph G of the full random regular graph, but for the sake of simplicity we will limit ourselves to a case of just one site i being eliminated, while more general cases will be presented elsehwere. Following the recipe of the previous section, one then considers the joint probability distribution of the the fields on the eliminated site i itself as well as its nearest neighborhood $j \in \partial i$:

$$P^{\text{full}}\left(X_i, Y_i, \{X_j, Y_j\}_{j\in\partial i}\right) := \left\langle \delta\left(Y_i - S_i^{\text{full}}\right)\delta\left(X_i - x_i\right)\cdot\prod_{j\in\partial i}\delta\left(Y_j - S_j^{\text{full}}\right)\delta\left(X_j - x_j\right)\right\rangle_x ,$$

where S_i^{full} is a function of all values of x_i that formally solves the full saddle-point equation (60), and the average is done over the full ensemble of x configurations.

In a familiar fashion, finding the solution S^{full} to the saddle-point equation can be seen as solving the 2 interconnected problems, see also Figure 12:

- Finding a set of solutions $S_j^{(j)}$ for each of the $K+1$ neighbors in the system where the values of x_i, y_i at site i are fixed.
- Finding a solution $S_i^{(i)}$ for the site i alone with the y_j for all nearest neighbors $j \in \partial i$ externally specified. This step is trivial for our case:

  $$S_i^{(i)} = \sum_{j\in\partial i} f\left(x_j, y_j\right),$$

  but we will use the notation $S^{(i)}$ for brevity.

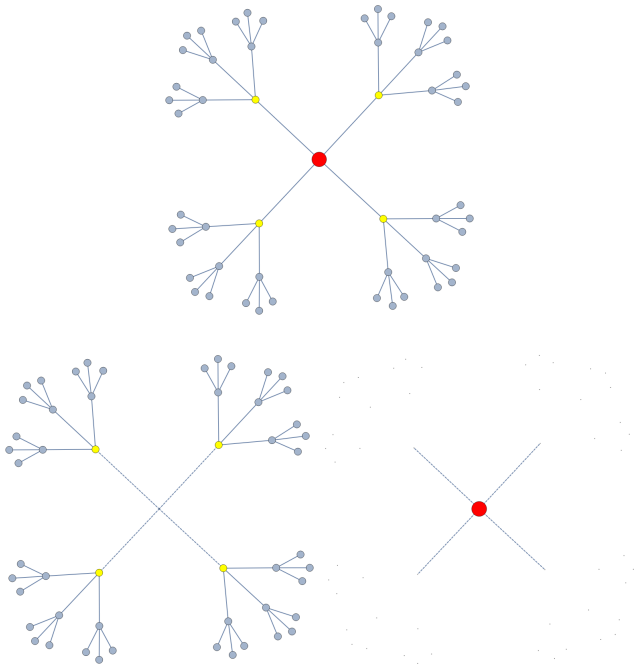


Figure 12: Illustration to the process of site decimation. The upper pane shows a neighborhood of a site to be eliminated (red), containing its closest neighbors (yellow). Solving a full system of equations (154) corresponds to first solving the equations for both subsystems on the lower figures, with the values of f corresponding to cut connections (dashed edges) replaced by externally specified values, and then synchronizing the values of f to satisfy the full system.

In order to move forward, we need the system obtained after elimination to factor in a union of independent subsystems, each of which interacts with the eliminated subsystem only via a scalar field. Luckily, this is almost the case in our problem as written—we only need to examine the independency assertion. Due to the locality of the equation (154), 2 given sites might only be correlated if there exists a path connecting these 2 sites. A random regular graph possesses a locally–tree-like structure, meaning that for any two close sites all but one paths connecting them are long, and this length grows with system size. Therefore, if one removes any site lying on that single short path, the two sites considered can only become correlated via a cycle with length of order of the system diameter, which is infinite in the thermodynamical limit. Because we know that all correlations in the order parameter decay exponentially with distance, and the decay rate is weakly sensitive to the particular value of the parameters, the locally loopless structure does lead to statistical independency of all $j \in \partial i$ once the site i itself is removed from the system. Identically, one can also say that the functions $S^{(j)}$ depend only on their private subsets of all independent fields x living on the corresponding branch of the local tree, while the dependence on values of x from different subsets is exponentially suppressed.

With this in mind we can now carry out the same type of decomposition of a product of δ-functions as done previously:

$$\delta\left(Y_i - S_i^{\text{full}}\left(\{x\}\right)\right) \cdot \prod_{j\in\partial i} \delta\left(Y_j - S_j^{\text{full}}\left(\{x\}\right)\right)$$

$$= \delta\left(Y_i - S_i^{(i)}\left(x_i, \{x_j, Y_j\}_{j\in\partial i}\right)\right) \cdot \prod_{j\in\partial i} \delta\left(Y_j - S_j^{(j)}\left(\{x_j\}_{j\neq i} \,|\, \{x_i, Y_i\}\right)\right) \cdot \left|\det\begin{pmatrix} 1 & -\frac{\partial S_i^{(i)}}{\partial Y_j} \\ -\frac{\partial S_j^{(j)}}{\partial Y_i} & 1 \end{pmatrix}\right|,$$

where we used yet another generalization of the Kac-Rice formula—this time of the product of δ-functions with components of a vector-function as an argument:

$$\prod_j \delta\left(\phi_j\left(\boldsymbol{x}\right)\right) = \delta\left(\boldsymbol{\phi}\left(\boldsymbol{x}\right)\right) = \delta\left(\boldsymbol{x} - \boldsymbol{x}_0\right) \cdot \left|\det\left\{\frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{x}}\right\}\right|, \tag{155}$$

where both x and ϕ are multidimensional vectors, and $\boldsymbol{x}_0$ is the unique solution to the vector-equation $\boldsymbol{\phi}\left(\boldsymbol{x}\right) = 0$. Similarly to the previous case, the role of ϕ is played by the those components of the full equation (154) that correspond to the neighboring sites.

Upon expanding the determinant of a block matrix, one arrives at:

$$... = \delta\left(Y_i - S_i\left(x_i, \{x_j, Y_j\}_{j\in\partial i}\right)\right) \cdot \prod_{j\in\partial i} \delta\left(Y_j - S_j\left(\{x_j\}_{j\neq i} \,|\, \{x_i, Y_i\}\right)\right) \cdot \left[1 - \sum_{j\in\partial i} \frac{\partial S_i}{\partial Y_j} \cdot \frac{\partial S_j}{\partial Y_i}\right],$$

where we have dropped the absolute value off the determinant because we still assert the solution to be unique and locally stable. Each factor in the resulting expression can now be averaged independently, each average performed only over the values of x in the corresponding branch of the local tree, which leaves us with the answer:

$$\begin{aligned} &P^{\text{full}}\left(X_i, Y_i, \{X_j, Y_j\}_{j\in\partial i}\right) \\ &= P_i\left(X_i, Y_i | \{X_j, Y_j\}_{j\in\partial i}\right) \times \prod_{j\in\partial i} P^{(j)}\left(X_j, Y_j | X_i, Y_i\right) \\ &- \sum_{j\in\partial i} \frac{\partial F_i\left(X_i, Y_i | \{X_j, Y_j\}_{j\in\partial i}\right)}{\partial Y_j} \cdot \frac{\partial F^{(j)}\left(X_j, Y_j | X_i, Y_i\right)}{\partial Y_i} \prod_{k\in\partial i\setminus\{j\}} P^{(k)}\left(X_k, Y_k | X_i, Y_i\right), \end{aligned} \tag{156}$$

$$F\left(X,Y|...\right)=\int\limits_{0}^{Y}dy\cdot P\left(X,y|...\right), \tag{157}$$

where the functions $P^{(j)},F^{(j)},j\in\partial i$ in the r.h.s represent the probability density functions and the cumulative distribution functions for the problem with site i dropped. Functions P_i, F_i represent the trivial distribution of the fields on site i in a system consisting of just this site i and all neighboring fields specified externally:

$$P_i\left(X_i,Y_i|\left\{X_j,Y_j\right\}_{j\in\partial i}\right)=P\left(X_i\right)\delta\left(Y_i-S^{(i)}\left(X_i|\left\{X_j,Y_j\right\}\right)\right) \tag{158}$$

$$F_i\left(X_i,Y_i|\left\{X_j,Y_j\right\}_{j\in\partial i}\right)=P\left(X_i\right)\theta\left(Y_i-S^{(i)}\left(X_i|\left\{X_j,Y_j\right\}\right)\right) \tag{159}$$

A final step is to integrate out the neighboring sites $j\in\partial i$ and obtain the answer for P^{full} in terms of solutions $P^{(j)}$ of the same problem on a truncated system:

$$P^{\text{full}}\left(X_i,Y_i\right)=\int\prod_{j\in\partial i}dx_jdy_jP^{\text{full}}\left(X_i,Y_i,\left\{x_j,y_j\right\}_{j\in\partial i}\right) \tag{160}$$

As it is already mentioned, we can approach a more general problem using the same sequence of steps. Importantly for our case, in a similar fashion we can derive a closed integral equation on the functions $P^{(j)}$ themselves by considering elimination of one more site neighboring to i, and the result reads:

$$\begin{aligned}&P\left(X_j,Y_j|X_i,Y_i\right)=P\left(X_j\right)\\&\times\int\limits_{\mathbb{R}-i0}\frac{ds}{2\pi}\exp\left\{isf\left(X_i,Y_i\right)\right\}\cdot\frac{\partial}{\partial Y_j}\left[\frac{\exp\left\{-isY_j\right\}}{-is}\cdot\left[\int dxdy\cdot P\left(x,y|X_j,Y_j\right)\exp\left\{isf\left(x,y\right)\right\}\right]^K\right],\end{aligned} \tag{161}$$

where we have substituted the integral representation of δ- and θ-functions in terms of their Fourier transforms, which allowed us to simplify the expression.

As a result, we now have a general framework to investigate statistics of quantities described by local equations with random independent noise.

8.2 Equations on the joint probabilities

In this section we will present a set of integral equations that fully describes the distribution of the order parameter in the limit $\Delta_0\to 0,K\to\infty,v=\Delta_0K=\text{const}$ and study the regime of "moderately small" $v\lesssim g$, that is already inaccessible by any kind of large K approximation but is nevertheless interesting for further practical considerations. The full derivation of the equations discussed below is now a purely technical task and is, therefore, left out, as at this point the reader should already be familiar with all key ideas that lead to the results below.

8.2.1 A few notations

In what follows, we will use the following set of already introduced notations:

- As we have been taught by the uncorrelated approximation, the most natural scale of the problem is the value of the mean-field order parameter found as the solution to the following self-consistency equation (62). Therefore, we will henceforth measure all quantities in units of Δ_0. In particular, we will use the already discussed dimensionless quantities

$$x_i := \frac{\xi_i}{\Delta_0}, y_i := \frac{\Delta_i}{\Delta_0} \tag{162}$$

- Following our previous experience, we also introduce dimensionless parameters of the theory—characteristic number of neighbors v and normal density of states:

$$v := K\Delta_0, \nu_0 := 2P_\xi(0) \tag{163}$$

  It is important that these are the only objects that know about the true form of the disorder distribution P_ξ, while every other aspect of the theory cares only about these two numbers.

- The r.h.s of the saddle-point equation is denoted as f and hence reads:

$$f(\xi_i, \Delta_i) = \frac{g}{K\Delta_0} \cdot \frac{\Delta_i}{\sqrt{\Delta_i^2 + \xi_i^2}} = \frac{g}{v} \cdot \frac{y_i}{\sqrt{x_i^2 + y_i^2}}$$

  So that the saddle-point equation now simply reads

$$y_i = \sum_j A_{ij} \cdot f(x_j, y_j) = \sum_{j \in \partial i} f(x_j, y_j), \tag{164}$$

  where A is just the adjacency matrix of the base graph.

- Just as in the case of the uncorrelated approximation, it is crucial for our results that f represents a homogeneous function:

$$f(x, y) = f(z = x/y), f(z) = \frac{g}{v} \cdot \frac{1}{\sqrt{1 + z^2}} \tag{165}$$

  Therefore we will use $f(z)$ instead of $f(x, y)$ whenever the discussed result relies heavily on the homogeneity of f.

- Following the notation of the previous section, by $F(a|b)$ we will denote functions that depend on the values of fields a in a system obtained from the original one by explicitly fixing the values of the fields b. Special care should be exercised as not to confuse this notation with the one used for conditional probability. Despite being qualitatively similar, our notation means a different thing.

- Finally, in order to increase the readability of the expressions below, we will write dummy variables (such as integration variables) in lower case, while external arguments for a given function are capitalized, for instance:

$$F(Y) = \int_0^Y dy \cdot P(y)$$

  This will prove useful when dealing with expressions involving multiple integrations and convolutions.

8.2.2 Exact form of the equations

Consider a system, where the values of the fields x_i, y_i on a given site i are fixed. This uniquely specifies the value of the r.h.s f of the saddle-point equation (164) for all nearest neighbors $j \in \partial i$ to be $f(x_i, y_i)$. As we have learned from the previous section, this also leads to statistical decoupling of all nearest neighbors $j \in \partial i$, and this decoupling becomes exact in the thermodynamical limit. The central object of our consideration is the cumulant generating function of the value $f(x_j, y_j)$ on a selected neighbor $j \in \partial i$ in the described system:

$$r(S|F_i) = \ln \langle \exp\{iSf(x_j, y_j)\} \rangle_{x|F_i},$$

where the averaging is done over all configurations of the disorder field x_i within the branch of the local–tree-like structure rooting from site j. The term $f(x_i, y_i)$ of the saddle-point equation (164) corresponding to the excluded site i is replaced by F_i.

One can then obtain the set of integral equations on the function $r(S|F)$ as well as the function r_1 which represents the mean value of $f(x_j, y_j)$, just as C_1 represented the mean value of the order parameter itself in the uncorrelated approximation:

$$r(S|F) = iS \cdot r_1(F) + \nu_0 \cdot \int_0^\infty dz \left(\exp\{iSf(z)\} - 1 - iSf(z)\right) \left[F + v \cdot \frac{d}{dz}\{z \cdot r_1(f(z))\}\right] \quad (166)$$

$$\begin{aligned} r_1(F) &= r_1(0) + \nu_0 \cdot v \cdot \int_0^\infty dz \cdot f(z) \cdot \frac{d}{dz}\{z[r_1(f(z)) - r_1(0)]\} \\ &+ \frac{F}{v} + \frac{g}{v} \cdot \nu_0 \cdot \int_0^\infty dy \cdot y \ln \frac{1}{y} \int_{\mathbb{R}-i0} \frac{d\omega}{2\pi} \exp\{i\omega F\} \cdot \exp\{v \cdot r(\omega|0)\} \cdot \exp\{-i\omega y\} \end{aligned} \quad (167)$$

The two equations above are a generalization of equations (105) and (117) from the uncorrelated approximation. One immediate observation is that the function r_1 now possesses a nontrivial dependence on the value of the fixed quantity F, which quantitatively reproduces the enhancement effect induced by neighboring sites, as discussed in the beginning of this section. Also, contrary to the uncorrelated approximation from the previous section, the equation on r cannot be solved explicitly, thus reflecting the nontrivial effect of mutual correlations, which, as we can see, goes well beyond a mere shift in the mean value of a counterpart of C_1.

Upon solving the two equations above, all physical quantities can be expressed as integrals containing $\exp\{v \cdot r\}$ as a weight function, as it becomes apparent form the previous section. For example, one can use (161) to access the joint probability distribution $P(X_i, Y_i)$ of the onsite values of the fields:

$$\begin{aligned} &P(X, Y) = /Z = X/Y/ \\ &= P(X) \cdot \int_{\mathbb{R}-i0} \frac{ds}{2\pi} \cdot \exp\{-isY\} \exp\{v \cdot r(s|f(X, Y))\} \left[1 + \frac{v}{Y} \cdot Z \frac{\partial}{\partial Z}\left\{\frac{r(s|f(Z))}{is}\right\}\right] \end{aligned} \quad (168)$$

8.2.3 Small v approximation

In the general case, equations (166)-(167) cannot be further simplified and neither can they be solved analytically, to the best of author's knowledge, so that the general case requires to use

numerical techniques to solve this coupled system of integrals equations. Yet for the case of moderately small $v \lesssim g$ the whole system can be significantly simplified. We will now consider the formal expansion in powers of v and retain only two leading orders, i.e., solve the equations for r, r_1 up to precision of $o\left(v^0\right)$. The resulting approximate system of equations appears to be more tractable by analytical means and presents a much easier problem for numerical solution.

The simplified version of the equation, after some algebra, amounts to the following:

$$r\left(S|F\right) \approx iS \cdot \left\{ \frac{F + f^{(0)}}{v} \right\} + \nu_0 \left[F + f^{(0)} \right] \Phi_0\left(S\right) + \nu_0 \Phi_1\left(S\right) + iS \cdot \left[r_1^{(0)}\left(F\right) - r_1^{(0)}\left(0\right) \right], \quad (169)$$

where $\Phi_{0,1}$ are familiar objects, similar to the F function (104) in the uncorrelated approximation:

$$\Phi_0\left(S\right) = \int\limits_0^\infty dz \left(\exp\left\{ iSf\left(z\right) \right\} - 1 - iSf\left(z\right) \right) \quad (170)$$

$$\Phi_1\left(S\right) = \int\limits_0^\infty dz \left(\exp\left\{ iSf\left(z\right) \right\} - 1 - iSf\left(z\right) \right) \frac{d}{dz}\left\{ zf\left(z\right) \right\} \quad (171)$$

The new function $r_1^{(0)}\left(F\right)$ has the following form:

$$\begin{aligned} & r_1^{(0)}\left(F\right) \\ & = \frac{g\nu_0}{v} \cdot \int\limits_0^\infty dy \cdot y \ln\frac{1}{y} \int\limits_{\mathbb{R}-i0} \frac{ds}{2\pi} \exp\left\{ isF \right\} \cdot \exp\left\{ v \cdot r\left(s|0\right) \right\} \cdot \exp\left\{ -isy \right\} \\ & = \frac{g\nu_0}{v} \cdot \int\limits_{\mathbb{R}-i0} \frac{ds}{2\pi} \exp\left\{ isF \right\} \cdot \exp\left\{ v \cdot r\left(s|0\right) \right\} \cdot \frac{\gamma - 1 + \ln\left\{ is \right\}}{\left(is\right)^2} \end{aligned} \quad (172)$$

Finally, the value of $f^{(0)}$ is defined via the solution to the following self-consistency condition:

$$0 = \nu_0 \int\limits_0^\infty dz \cdot f\left(z\right) \cdot \frac{d}{dz}\left\{ zf\left(z\right) \right\} + r_1^{(0)}\left(0\right) \quad (173)$$

It plays a role similar to that of C_1 in the uncorrelated approximation. While expressions (169)-(172) still are coupled to each other, both of them can be now explicitly solved, so that one simply has to deal with multiple numerical integrations to compute the value of both $r_1^{(0)}$ and r.

Moving further, small values of $v \lesssim 1$ allow us to come up with an approximate analytical expression for $r_1^{(0)}$. Namely, the integral over s in (172) is logarithmic and we can therefore drop even $O\left(v^0\right)$ terms from the exponent, thus obtaining

$$r_1^{(0)}\left(F\right) \approx \frac{g\nu_0}{v} \cdot \left(F + f^{(0)} \right) \ln\frac{1}{F + f^{(0)}} \quad (174)$$

Because the argument of r_1 only sweeps a finite range $F \in \left[0, \frac{g}{v} \sim 1\right]$, this expression appears to be a good quantitative approximation for the true function r_1 found by explicit integration of the equation (172). This, in turn, also allows us to find the value of $f^{(0)}$:

$$f^{(0)} = \frac{c}{W\left(c\right)}, \quad c = \frac{v}{g} \int\limits_0^\infty dz \cdot f\left(z\right) \cdot \frac{d}{dz}\left\{ zf\left(z\right) \right\} = \frac{g}{v} \cdot \frac{\pi}{4}, \quad (175)$$

where $W(c)$ is the Lambert W-function solving the equation $ye^y = x$, and the value of c is obtained for the particular form of f in our problem.

As a result, we have explicitly solved the exact set of equations (166)-(167) for small v, which allows us to analytically observe the complete set of features in statistics of the saddle-point equation.

8.2.4 Comparison with the numerical solution

It is time to conduct a comparison of the developed theory with the exact numerical solution. To keep things simple, we have chosen to examine the predictions for the joint probability function of the fields x, y on a given site. As we have seen previously, this object contains sufficient information to estimate the quality of reproduction of both the onsite distribution of the order parameter on its own and the effects of correlations that show up at small values of ξ.

To this end, one has to solve the system (166)-(167) numerically and then use the answer for r to obtain the form of P by means of (168), requiring yet another numerical integration. Although it is within the authors possibilities to conduct this full program, we opt for using the approximate expressions (169)-(172), as they succeed in reproducing the exact answer just as well, while being enormously simpler for numerical analysis. The results are presented on Figures 13-14 below.

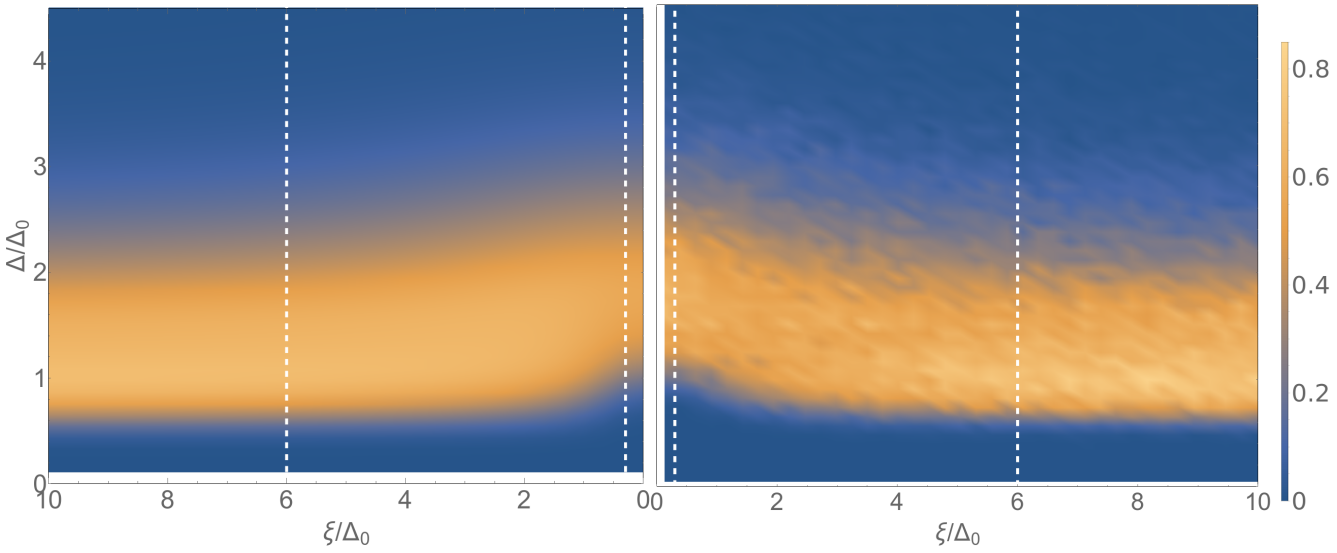


Figure 13: Density plots of the joint probability distribution $P(\Delta, \xi)$, as extracted from the numerical data (left) and the theoretical model (right). The values are multiplied by $P_\xi^{-1}(\xi) \cdot \Delta_0^{-2}$ to reflect the data in a small region $\Delta \sim \xi \sim \Delta_0$. Color represents the magnitude of the resulting quantity, according the legend on the right. Parameters of the simulation: $g = 0.16$, $K = 50$, total system size $N = 2^{22} \approx 4.2 \cdot 10^6$, which corresponds to $v = 0.19$. The perceived "noise" in the numerical data is due to a finite sample size, as we are zooming in to a very small region of the ξ values, so that there are only $2\Delta_0 \cdot N \sim 3 \cdot 10^4$ contributing to the areas showed on the plot. White lines show cross-sections of the plots that are presented on Figure 14.

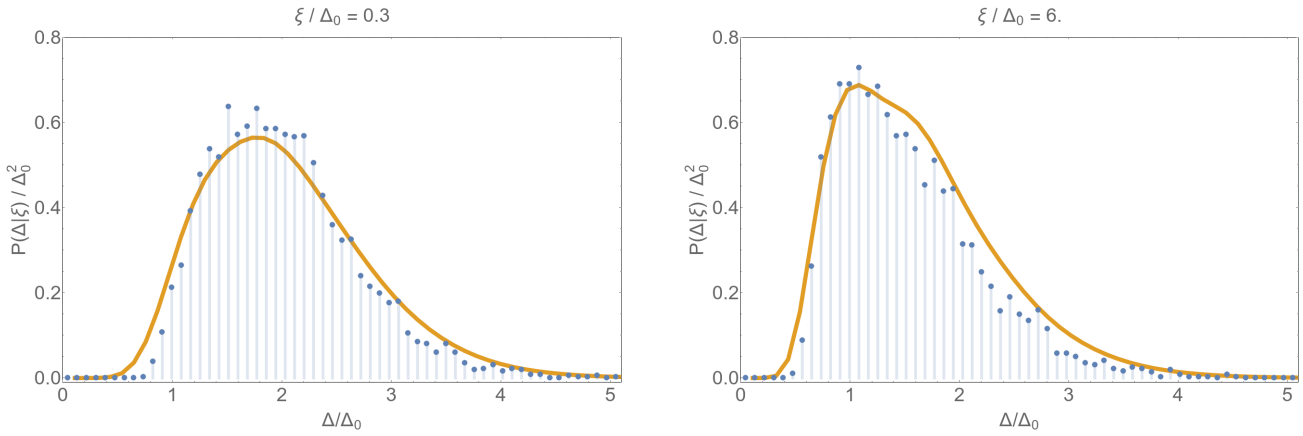


Figure 14: Plot of the probability distribution $P(\Delta, \xi)$. The values are multiplied by $P_{\xi}^{-1}(\xi) \cdot \Delta_0^{-2}$. All theoretical and numerical parameters are the same as on Figure 13. Left plot shows the distribution for $\xi = \xi_0 = 0.3\Delta_0$, and the right one stands for $\xi = \xi_0 = 6\Delta_0$. For numerical plots, a window of $\xi \in [\xi_0 - 0.15\Delta_0, \xi_0 + 0.15\Delta_0]$ was used to collect statistics. The perceived "noise" in the numerical data is due to a finite sample size, as we are zooming in to a very small region of the ξ values, so that there are only $0.3\Delta_0 \cdot N \sim 5 \cdot 10^3$ contributing to the data on the histogram.

In addition to what we have already observed during the investigation of the uncorrelated approximation, we highlight the following additional features:

- As we have expected, the value of Δ indeed bumps up for $\xi \sim \Delta$. This will be of crucial importance for further applications.
- This time no fitting is done whatsoever, yet we can see an excellent agreement between the theory and the experiment. As additional numerical experiments show, the discrepancy is explained by a finite value of v and can be removed either by considering smaller values of $g \sim v$, or by taking into account $O(v)$ corrections to the equations (169)-(172).

Part IV
Low energy fluctuations of the order parameter

9 Description of small fluctuations of the order parameter at low energies

9.1 Propagator of quadratic fluctuations

Away from the region of strong dynamical fluctuations, i.e., away from the superconducting transition itself, one expects to describe all properties of the fluctuations by Gaussian quadratic fluctuations. In terms of functional integrals this means expanding the order parameter action (17) up to the second power of deviation $\Delta - \langle \Delta \rangle$ from the saddle-point configuration. Note that by $\langle \bullet \rangle$ we denote the quantum mechanical averaging of a given field $\Delta(\tau)$ over all configurations according to functional integral (19) in a particular realization of static disorder fields ξ_j, as

opposed to statistical averaging over all configuration of ξ fields, as in Part III. The resulting quadratic action reads:

$$\delta_\alpha^i(\tau) := \Delta_i^\alpha(\tau) - \langle \Delta_i^\alpha \rangle, \tag{176}$$

$$S[\delta] = S[\langle \Delta \rangle] + \frac{1}{2}\int d\tau d\tau' \cdot \delta_\alpha^i(\tau) \cdot L_{ij}^{\alpha\beta}(\tau,\tau') \cdot \delta_\alpha^i(\tau') + O\left(\delta^3\right), \tag{177}$$

where the first term is the saddle-point action (57) that delivers the free energy, the linear term vanishes by virtue of the saddle-point equation, and the quadratic term is expressed via the following operator:

$$L_{ij}^{\alpha\beta}(\tau,\tau') = \left(\frac{1}{2}J\right)_{ij}^{-1}\delta^{\alpha\beta}\cdot\delta(\tau-\tau') - 4\left\langle\left\langle T_\tau\left\{s_i^\alpha(\tau)\, s_j^\beta(\tau')\right\}\right\rangle\right\rangle_{\langle\Delta\rangle} \tag{178}$$

where the second term represents the response function of non-interacting spins calculated at the saddle point configuration. One can express it either as the dynamic spin correlator with non-interacting Hamiltonian

$$H_{\text{spin}} = \sum_i 2\xi_i s_i^z - 2\langle\Delta_i^\alpha\rangle s_j^\alpha \tag{179}$$

or as a Matsubara semionic response loop with spin operators as vertices and a semionic Green function, as presented in [15]. Either way, one ends up with the following expression for the Fourier image of this correlator in the zero-temperature limit:

$$\Pi_{ij}^{\alpha\beta}(\omega) = -\delta_{ij}\frac{1}{4}\frac{\sqrt{\xi_i^2+|\Delta_i|^2}}{\xi_i^2+|\Delta_i|^2-\frac{\omega^2}{4}}\left(\frac{\Delta_i^\alpha\Delta_i^\beta}{\xi_i^2+|\Delta_i|^2} - \delta^{\alpha\beta} + \epsilon^{\alpha\beta}\frac{\omega\xi_i}{2\left(\xi_i^2+|\Delta_i|^2\right)}\right), \tag{180}$$

where $\epsilon^{\alpha\beta} = i\sigma^y$ is a fully antisymmetric matrix, and we have removed averaging brackets $\langle\bullet\rangle$ in notation of the saddle-point configuration of the order parameter. In this expression we have swiftly switched from imaginary bosonic Matsubara frequencies to real ones describing the zero-temperature limit.

Within the approximation of Gaussian fluctuations, the inverse of the introduced operator L gives the propagator of fluctuations δ of the order parameter:

$$G_{ij}^{\alpha\beta}(\tau,\tau') = \left\langle \delta_i^\alpha(\tau)\,\delta_j^\beta(\tau')\right\rangle \approx \left(L^{-1}\right)_{ij}^{\alpha\beta}(\tau,\tau'), \tag{181}$$

where L^{-1} implies full operator inverse, including component indices α,β, spatial indices i,j and time τ,τ', i.e.,

$$\int d\tau''\cdot\sum_j L_{ij}^{\alpha\beta}(\tau,\tau'')\cdot G_{jk}^{\beta\gamma}(\tau'',\tau') = \delta^{\alpha\gamma}\delta_{ik}\delta(\tau-\tau') \tag{182}$$

There is one important aspect to be mentioned at this point. By virtue of the saddle-point equation (60), the operator L possesses a unique zero eigenmode corresponding to the phase rotation of the order parameter, which corresponds to a rotation with $i\sigma^y$ matrix in vector notation:

$$\langle\Delta_i^\alpha\rangle(\phi) := \left[\exp\{i\phi\sigma^y\}\right]^{\alpha\beta}\left\langle\Delta_i^\beta\right\rangle, \tag{183}$$

so that the following expression is an immediate consequence of the saddle-point equation 60:

$$\int d\tau' \sum_j L_{ij}^{\alpha\beta}(\tau,\tau) \cdot \frac{d\left\langle \Delta_j^{\beta} \right\rangle(\phi)}{d\phi} = \sum_j L_{ij}^{\alpha\beta}(\omega=0) \cdot [i\sigma^y]^{\beta\gamma} \cdot \left\langle \Delta_j^{\gamma} \right\rangle = 0, \tag{184}$$

where the second expression is written in terms of a Fourier image at zero frequency. As it now becomes evident, the presented eigenmode is a direct consequence of $U(1)$ gauge symmetry of the problem. One generally expects this relation to be respected by a proper averaging technique. The same type of symmetry mode can be shown to be present at a finite temperature as well.

9.2 Transverse modes at low energies: formulation of the problem

As already discussed in Subsection 4.3.2, we expect that the propagator of the order parameter fluctuations possesses low energy eigenmodes that create a significant contribution to the current response function at subgap frequencies. While the loop term (34) of the response function can deliver additional contributions to the current response, at the very least it is true that if the correlator (181) does indeed possess eigenmodes at low energies, they will inevitably cause a finite current response. Additionally, as was demonstrated previously in Subsection 4.3.2, because of the structure of the current vertex it also makes sense to expect those modes to be coming from the transverse component of the fluctuations, i.e., the ones that are directed perpendicular to the static saddle-point order parameter $\langle \Delta_i^{\alpha} \rangle$ and thus are themselves fluctuations of the phase of the order parameter in complex representation.

However, the linear operator (178) also contains information about the dynamics of the longitudinal, or Higgs', modes of the fluctuations corresponding to alterations in the absolute value of the order parameter. Finally, it also includes mixing terms that lead hybridization of phase and amplitude modes. Generally this means that at a finite frequency the sought-for low-energy modes represent some superposition of amplitude and phase oscillations. However, in [15] it is argued that the amplitude mode is gapful with the corresponding energy scale of order $2\Delta_0$. Therefore, low-energy excitations, if any, will have only a small amplitude component. And even more so, when taking into account that the hybridizing term in (178) is proportional to frequency and thus is expected to bring in additional smallness and low energies. Having this in mind, we can basically ignore the existence of amplitude modes, and only work with that part of operator L that represents dynamics of the phase modes. Its Fourier image in zero temperature real frequency domain is given by:

$$L(\omega) = \left(\frac{1}{2}J\right)^{-1}_{ij} - \delta_{ij}\Pi_i, \quad \Pi_i = \frac{\sqrt{\xi_i^2+\Delta_i^2}}{\xi_i^2+\Delta_i^2-\frac{\omega^2}{4}}, \tag{185}$$

where the component indices are now absent because the operator is projected onto the subspace of phase fluctuations, and Δ_i denotes the saddle-point configuration of the order parameter, chosen to be aligned along the $(1,0)$ vector uniformly across the whole system, as was also done in the previous Part III. In what follows, we will drop the quantum mechanical average $\langle \bullet \rangle$ in notations of the saddle-point value of the order parameter, reserving it now for averages over disorder, as it should not cause any confusion.

One might note that the operator (185) seems to be ill-defined when the frequency reaches the value $2 \cdot \sqrt{\xi_i^2+\Delta_i^2}$. However, this event would imply that that either both ξ and Δ at a given site are abnormally low, or the frequency itself takes a value of at least several Δ_0 in magnitude.

While the former is an incredibly rare event—as we have seen in Part III, this is even less likely to happen than having a small value of Δ_i alone, which is, itself, a double exponentially rare event—the latter can by no means be referred to as a "low frequency". In other words, we expect such a singularity to never occur in our problem.

We are finally in position to formulate the immediate task of this part: we are studying the existence and properties of low energy transverse fluctuations of the order parameter. Mathematically this is formulated as searching for all pairs of frequency ω_λ and normalized vectors$\psi^{(\lambda)}$ satisfying the following equation:

$$\sum_j L_{ij}\left(\omega_\lambda\right)\cdot\psi_j^{(\lambda)}=0,\ \ \sum_i\left[\psi_i^{(\lambda)}\right]^2=1, \tag{186}$$

with L given by (185). We already know one such solution guaranteed to us by $U\left(1\right)$ symmetry (184):

$$\omega_0=0,\quad \psi_i^{(0)}=\frac{1}{\mathcal{N}}\Delta_i,\ \ \mathcal{N}^2:=\sum_i\Delta_i^2, \tag{187}$$

where $\mathcal{N}$ is a normalizing coefficient. Note, however, that this does not imply a finite thermodynamical density of states at low frequencies, as this mode is found to be a solitary one, so that its contribution to any physical quantities vanishes in the thermodynamical limit.

As already indicated, we are interested in the values of ω_λ well below the typical superconducting scales. The concrete choice of the upper limit for frequency beyond which it should be regarded as a high frequency is hard to define though. Indeed, from Part III we can infer that there is no clear quantitative scale for the magnitude of the order parameter, as the body of the order parameter distribution occupies an interval of several Δ_0. Therefore, below we will present data with frequencies up to several Δ_0 in magnitude, implying that it still provides valuable insights into the spectrum structure. Also, keep in mind that for all practical concerns, we are interested in characterizing the region of parameters in which the eigenmodes with arbitrarily low energy exist.

9.3 Important characteristics of the spectrum

9.3.1 Density of solutions

One important characteristic of the spectrum to keep track of its the density of solutions of the discussed eigenproblem (186), formally defined as

$$\rho_1\left(\omega\right):=\lim_{N\to\infty}\left\langle\frac{1}{N}\sum_\lambda\delta\left(\omega-\omega_\lambda\right)\right\rangle_\xi, \tag{188}$$

where N is the system size, λ enumerates all solutions to the eigenvalue problem (186) in a particular realization, and the average $\langle\bullet\rangle$ is taken over all realizations of the disorder field ξ. This quantity obviously describes the total density of eigenmodes on the frequency axis.

9.3.2 Inverse Participation Ratio

Among other characteristics of solutions to (186) we are also interested in localization properties, as they are responsible for whether the corresponding low-energy mode actually causes dissipation.

Indeed, while both localized and extended modes will show up in the current response function, one expects only the extended mode to contribute to the real part of conductivity, that would then induce dissipation. Of course, there are mechanisms that can cause dissipation to occur in localized modes as well, but those are well beyond the scope of the present model and deserve a separate investigation. Therefore, the classification of the spectrum into localized and extended parts provides valuable information. Mathematically, localization properties can be accessed by the so called Inverse Participation Ratio (IPR), defined as

$$I(\omega, N) := \left\langle \frac{1}{N} \sum_i \cdot \sum_\lambda |\langle \psi_\lambda | i \rangle|^4 \cdot \delta(\omega - \omega_\lambda) \right\rangle, \tag{189}$$

where ψ_λ is the normalized eigenvector corresponding to the frequency model ω_λ, and summation over i is performed across all sites of the system. Note that this quantity now contains a dependence on N, and the character of this dependence is what describes the localization properties. To illustrate that, let us consider two limiting cases:

- All eigenfunctions ψ_λ at a given frequency ω are extended. In this case, the eigenfunction is spread approximately uniformly across the whole system, so that a typical onsite amplitude of a normalized vector reads:

  $$1 = \sum_i |\langle \psi | i \rangle|^2 \sim N \cdot |\langle \psi | i \rangle|^2 \Rightarrow \langle \psi | i \rangle \sim \frac{1}{\sqrt{N}}$$

  For the IPR this results in

  $$I(\omega, N) \sim \left\langle \frac{1}{N} \sum_i \cdot \sum_\lambda \frac{1}{N^2} \cdot \delta(\omega - \omega_\lambda) \right\rangle \sim \frac{1}{N} \cdot \rho(\omega), \tag{190}$$

  i.e., one expects the IPR to be inversely proportional to the system size.

- In the opposite regime, when all eigenfunctions are confined to a finite number of sites defined by the localization volume, similar considerations result in

  $$I(\omega, N) \sim \frac{1}{N_{\text{loc}}(\omega)} \cdot \rho(\omega), \tag{191}$$

  where N_{loc} is the typical localization volume at a given energy.

The general tendency is that the typical value of IPR is given by the inverse support size of the eigenvector. Of course, there is a wide variety of intermediate cases, as well as there are possibilities to conduct a much more detailed research of eigenfunctions moments. However, our numerical capabilities are not quite enough to gather the necessary statistics. Hence we resort to a qualitative analysis of N-dependence of the IPR at a given energy. In the results presented below, we will only discriminate between power law decay of the IPR with system-size N, corresponding to delocalized states at the given energy, and N-independent behavior indicative of localization.

9.3.3 Imaginary part of correlator trace

Finally, we will be interested in the imaginary part of the diagonal matrix elements of transverse fluctuations correlator, given by the inverse of L:

$$\rho_2(\omega) := \lim_{N\to\infty}\left\langle \frac{1}{\pi}\cdot \mathrm{Im}\left\{\left[L^{-1}(\omega+i0)\right]_{ii}\right\}\right\rangle_\xi = \lim_{N\to\infty}\left\langle \frac{1}{\pi}\cdot \mathrm{Im}\left\{\frac{1}{N}\mathrm{Tr}\left[L^{-1}(\omega+i0)\right]\right\}\right\rangle_\xi, \quad (192)$$

where L is given by (185), the second expression assumes restoration of the translational invariance upon averaging over disorder, and the whole equation includes retarded regularization $+i0$ to define the inverse of L. As we have seen in Subsection 4.3.2, this expression is a candidate for creating a finite real part of conductivity at a given frequency. Additionally, this is the only quantity among the three considered for which we have developed quantitative predictions.

At this point it is important to understand that, contrary to what may seem initially, the quantities ρ_1 and ρ_2 are not equal to each other. Initially, one can spot an obvious dimensionality mismatch: ρ_1 has the dimension of inverse energy, while ρ_2 has a multiplier of dimension Δ_0^2 on top of that. Mathematically, this is because the problem (186) is not a standard eigenproblem of the form

$$H|\psi_\lambda\rangle = E|\psi_\lambda\rangle, \quad (193)$$

where the eigenvalue enters the equation in a scalar (in terms of matrix structure) and linear fashion, and one can then express the Green function via a standard expansion in orthogonal projectors:

$$(\omega+i0-H)^{-1} = \sum_\lambda \frac{|\psi_\lambda\rangle\langle\psi_\lambda|}{\omega+i0-E_\lambda}, \quad (194)$$

with λ enumerating all solutions 193. In our problem (186), the frequency dependence of L is much more complicated: not only is it nonlinear, but it also contains a nontrivial matrix structure of Π. The correct expression for the inverse of L in terms of spectral decomposition can be obtained via mapping to a standard eigenproblem:

$$\begin{aligned}
&\left((J/2)^{-1}-\Pi(\omega)\right)|\psi_\lambda\rangle = 0 \Leftrightarrow \Pi(\omega)\left(\Pi^{-1}(\omega)-\frac{J}{2}\right)(J/2)^{-1}|\psi_\lambda\rangle = 0\\
&\Leftrightarrow \left(\varepsilon-\frac{J}{2}\right)|v_\lambda\rangle = \lambda\cdot\frac{1}{\varepsilon}|v_\lambda\rangle, \quad |v_\lambda\rangle = a_\lambda\cdot(J/2)^{-1}|\psi_\lambda\rangle, \quad \lambda=\frac{\omega^2}{4}, \quad a_\lambda = \frac{1}{\sqrt{\langle\psi_\lambda|(J/2)^{-2}|\psi_\lambda\rangle}}\\
&\Leftrightarrow \sqrt{\varepsilon}\left(\varepsilon-\frac{J}{2}\right)\sqrt{\varepsilon}|\lambda\rangle = \lambda\cdot|\lambda\rangle, \quad |\lambda\rangle = b_\lambda\cdot\frac{1}{\sqrt{\varepsilon}}|v_\lambda\rangle, \quad b_\lambda = \frac{1}{\sqrt{\langle v_\lambda|\frac{1}{\varepsilon}|v_\lambda\rangle}},
\end{aligned} \quad (195)$$

where ε and $f(\varepsilon)$ stand for the following diagonal matrices:

$$\varepsilon = \mathrm{diag}\left\{\sqrt{\xi_i^2+\Delta_i^2}\right\}, \quad f(\varepsilon) = \mathrm{diag}\left\{f\left(\sqrt{\xi_i^2+\Delta_i^2}\right)\right\} \quad (196)$$

and, when passing on to the second line, we have taken advantage of the fact that the matrix Π can be expressed as:

$$\Pi(\omega) = \frac{\varepsilon}{\varepsilon^2-\omega^2/4} \quad (197)$$

The resulting problem (195) is now a standard eigenproblem for the operator $\varepsilon^2 - \sqrt{\varepsilon}\frac{J}{2}\sqrt{\varepsilon}$, so that the spectral decomposition for its inverse can be used:

$$\left[z - \sqrt{\varepsilon}\left(\varepsilon - \frac{J}{2}\right)\sqrt{\varepsilon}\right]^{-1} = \sum_{\lambda} \frac{|\lambda\rangle\langle\lambda|}{z-\lambda},\ z \in \mathbb{C} \tag{198}$$

One can then rewrite the inverse of the target operator (185) as follows:

$$\begin{aligned} L^{-1}(\omega) &:= \left((J/2)^{-1} - \Pi(\omega)\right)^{-1} = \frac{J}{2}\left[\Pi^{-1}(\omega) - \frac{J}{2}\right]^{-1}\Pi^{-1}(\omega) \\ &= \frac{J}{2}\left[\frac{\varepsilon^2 - z_\omega}{\varepsilon} - \frac{J}{2}\right]^{-1}\Pi^{-1}(\omega) = \frac{J}{2}\left[\frac{1}{\sqrt{\varepsilon}}\left(\varepsilon^2 - z_\omega\right)\frac{1}{\sqrt{\varepsilon}} - \frac{J}{2}\right]^{-1}\Pi^{-1}(\omega) \\ &= -\frac{J}{2}\sqrt{\varepsilon}\left[z_\omega - \sqrt{\varepsilon}\left(\varepsilon - \frac{J}{2}\right)\sqrt{\varepsilon}\right]^{-1}\sqrt{\varepsilon}\Pi^{-1}(\omega) \\ &= -\frac{J}{2}\sqrt{\varepsilon}\left(\sum_{\lambda}\frac{|\lambda\rangle\langle\lambda|}{z_\omega - \lambda}\right)\sqrt{\varepsilon}\Pi^{-1}(\omega), \end{aligned} \tag{199}$$

where we have denoted $z_\omega = \omega^2/4$. Finally, we use the definition of $|\lambda\rangle$ to make the expression explicitly symmetric:

$$\begin{aligned} &\frac{J}{2}\sqrt{\varepsilon}\,|\lambda\rangle = \frac{J}{2}\sqrt{\varepsilon}\cdot b_\lambda \cdot \frac{1}{\sqrt{\varepsilon}}\,|v_\lambda\rangle = \Big/\left(\varepsilon - \frac{J}{2}\right)|v_\lambda\rangle = \lambda\cdot\frac{1}{\varepsilon}\,|v_\lambda\rangle\Big/ \\ &= b_\lambda\cdot\left(\varepsilon - \lambda\cdot\frac{1}{\varepsilon}\right)|v_\lambda\rangle = b_\lambda\cdot\left(\Pi^{-1}(\omega) + (z_\omega - \lambda)\frac{1}{\varepsilon}\right)|v_\lambda\rangle \end{aligned} \tag{200}$$

So that we finally obtain to the desired spectral decomposition

$$L^{-1}(\omega) = -\Pi^{-1}(\omega)\sqrt{\varepsilon}\left(\sum_{\lambda}\frac{|\lambda\rangle\langle\lambda|}{z_\omega - \lambda}\right)\sqrt{\varepsilon}\Pi^{-1}(\omega) - \Pi^{-1}(\omega) \tag{201}$$

where we have used that $|\lambda\rangle$ form a complete basis, which allowed us to simplify the second term. In particular, one obtains for the imaginary part of the trace:

$$\begin{aligned} &\rho_2(\omega) \\ &= \frac{1}{\pi N}\left\langle \mathrm{Im}\left\{\sum_{\lambda}\frac{\left\langle\lambda\left|\frac{\left(\varepsilon^2 - (\omega+i0)^2/4\right)^2}{\varepsilon}\right|\lambda\right\rangle}{\lambda - (\omega+i0)^2/4}\right\} - \mathrm{Im}\mathrm{Tr}\left\{\Pi^{-1}(\omega+i0)\right\}\right\rangle_{\xi} \\ &= \frac{1}{N}\left\langle\sum_{\lambda}\delta(\omega-\omega_\lambda)\cdot\frac{\langle\lambda|\varepsilon^3|\lambda\rangle + 2\frac{\omega_\lambda^2}{4}\langle\lambda|\varepsilon|\lambda\rangle + \left(\frac{\omega_\lambda^2}{4}\right)^2\left\langle\lambda\left|\frac{1}{\varepsilon}\right|\lambda\right\rangle}{\omega_\lambda/2}\right\rangle_{\xi} \end{aligned} \tag{202}$$

In particular, we can see, that the weight of each eigenvector takes a nontrivial frequency-dependent form that leads to quantitative differences with ρ_1 even upon switching to dimensionless quantities. However, due to the presence of the common spectral weight $\delta(\omega - \omega_\lambda)$ for each value of frequency, the two densities ρ_1 and ρ_2 are either simultaneously equal to zero because of absence of states at that frequency or both take some positive value, as it is also discussed in [19, sec. 2.3.5].

9.3.4 Note about the qualitative character of physics conclusions

As a final note, let us draw attention to the fact that it is neither the current response function (32) that is being averaged over disorder and analyzed in the present research, nor is it any other physical object that can be directly accessed via measurement. As discussed in Part II, this is both because of an overly simplistic initial model in the first place and because any potential candidate for a physical quantity in the present model requires large technical effort to come up with an adequate averaging scheme, be it a numerical or an analytical one. All of the spectral quantities described above characterize some auxiliary objects that present a simple initial target for analysis. As a result, these quantities bear only qualitative information about the typical orders of relative scales and key elements of the phase diagram of the real materials. Nevertheless, it still is a meaningful task to seek qualitative convergence for any theoretical approach, as the latter might contain potential generalizations to actual systems. Having this convergence would imply that the underlying technical apparatus accurately reflects any delicate interplay between different physical mechanisms in the model and thus, if generalizable, claims to be a hopeful approach to real systems.

10 Numerical study

In building the qualitative picture and verifying our theoretical predictions we will rely on exact numerical diagonalization of the operator (185) in each particular disorder realization. Such a choice is a consequence of important correlations present in the order parameter distribution, which renders most of commonly applied indirect numerical schemes inapplicable. The framework, as well as a showcase of the results is presented below.

10.1 Implementation notes on the numerical algorithm

To this end, one has to conduct the following program for each particular realization of the disorder and then analyze statistics of quantities of interest across multiple disorder realizations:

1. Numerically solve the saddle-point equation (60), as discussed in Subsection 6.4
2. Calculate the operator (185) and solve the associated eigenvalue problem (186).
3. Calculate the required characteristics of eigenvalues and eigenvectors, such as ρ_1, ρ_2, I.

However, it is worth making several important implementation notes:

1. Because it is numerically expensive to calculate and store the inverse of a large adjacency matrix contained in J as well as to diagonalize the resulting dense matrix (185), we use the discussed mapping of the problem to a standard eigenvalue problem (195). Luckily for us, it is formulated in terms of the adjacency matrix itself, which is a sparse matrix. There is a number of good techniques to deal with numerical diagonalization of large sparse matrices, we opted for the FEAST algorithm [24].
2. At this point we also note that there is little to no profit in restoring the full spectrum of the target operator (185), as discussed in the previous section. We then take advantage of the fact that we are only interested in the presence or absence of eigenvalues in a small region of frequencies with order of several Δ_0 in magnitude. Out of the total large number

of eigenvalues, this constitutes only a small fraction of order of numerical value of Δ_0, which significantly reduces the required computational resources and time.

3. Our implementation of the program above allowed us to study systems with sizes of up to $2^{16} \approx 65 \cdot 10^3$ sites, thus being well beyond the $2^{22} \sim 4.2 \cdot 10^6$ that we have managed to reach for the saddle-point equation itself. The main limitation that bounds the accessible sizes is that the execution time of the algorithm grows quickly with the system size:

$$t_{\text{exec}} \sim n^2 N \sim (N\Delta_0)^2 N \sim N^3,$$

where n is the number of sought eigenvalues. Thus, studying systems with $N \sim 3 \cdot 10^3$ requires some 8 hours of computational time on available hardware. There is still a room for improvement, however, and the corresponding work is in progress.

4. However, the fact that the eigenvalues of interest are situated in a small region of several Δ_0 in size has its adverse sides: as Δ_0 is decreased, one needs to take larger system sizes to secure the same quality of statistical averaging. Together with the discussed limitation on the system sizes, this puts a lower limit on the value of Δ_0 that we can access numerically. For similar reasons, the exponential growth of the number of neighbors at distance d puts an upper limit on the numerically accessible values of K, as discussed in Subsection 6.4.

5. Finally, one should take care of practical difficulties emerging from δ-functions in the definitions of all spectral properties of interest. While these quantities are well defined for theoretical computation, estimating them numerically presents two complications. First of all, for any finite ensemble a sum of δ-functions remains a function with zero support, hence we have to introduce a finite regularization to smear the δ-peak enough to cover the level spacing between adjacent eigenvalues. This is done by replacing the δ-function with a Lorentzian of finite width ϵ:

$$\delta(x) \mapsto \frac{1}{\pi} \frac{\epsilon}{\epsilon^2 + x^2}, \tag{203}$$

which can be regarded as a finite level broadening from the physics point of view. Below we will specify the value of ϵ for all data presented. Practically, it should be large enough to have many eigenvalues contained within an interval of width ϵ. It then follows from typical magnitudes of the frequencies that ϵ should be significantly larger than $\Delta_0/(N \cdot M)$, where N is the system size and M is total number of realizations for the given system size. The smaller the ϵ can be while still containing a sufficiently large number of eigenvalues, the more accurately can the true profile of the quantity in question be reproduced. Additionally, a finite value of ϵ implies that the resulting estimate of the target quantity at a given frequency is only valid when the collected data is representative for the spectrum in an interval of several ϵ on either side of the target frequency. In particular, to present reliable data for frequency ω, we need to have also collected frequencies at least as large as $\omega \pm 10\epsilon$.

10.2 Demonstration of representative results

In this section we present and discuss a sample of our numerical results. Figure 15 shows the density of frequencies $\rho_1(\omega)$, as given by (188), and on Figure 16 one can find the Inverse Participation Ratio (189) for the same dataset.

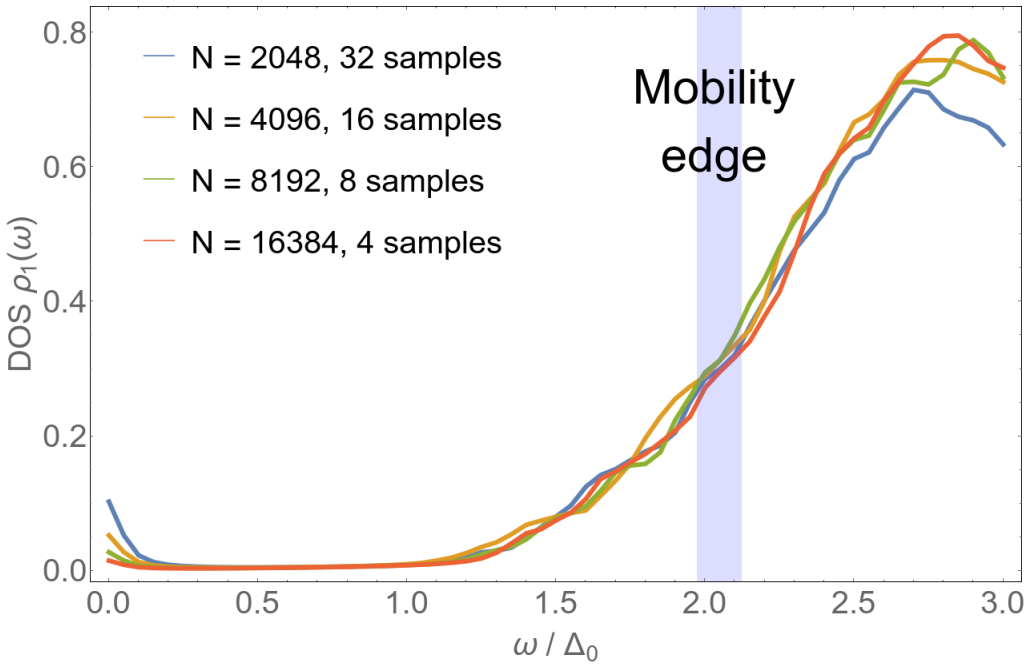


Figure 15: Frequency density of eigenmodes $\rho_1(\omega)$ of the operator (185) according to the results of a numerical diagonalization. The frequency is measured in native units of mean fields order parameter Δ_0. Parameters of the model: $g = 0.24$, $K = 8$, which corresponds to $\Delta_0 = 0.031$ and $v = K\Delta_0 = 0.25$. Each curve is labeled with the corresponding systems size and number of realizations. To calculate the density of states from a finite numerical dataset, Lorentzians with width $\epsilon = 0.05\Delta_0$ were used to approximate the δ-function, as described in (203). This corresponds to $\sim 10^2$ eigenvalues effectively contributing to each value on the plot. The blue strip indicates the approximate position of the localization transition, as judged by the Inverse Participation Ratio curves on Figure 16.

Based on Figure 15, the spectrum can be qualitatively characterized by 4 main regions:

- Firstly, we observe a well pronounced body of the spectrum, qualitatively reminiscent of a smeared BCS-like coherence peak. Note that the region climbs down to frequencies of $\omega \approx \Delta_0$, being well beyond the conventional threshold of $2\Delta_0$.

- Secondly, there is a gap region with no states detected. Numerical experiments with different parameters suggest that the eigenmodes can penetrate to arbitrary low energies, so that the gap eventually closes.

- Finally, there is a small peak exactly at zero frequency, originating from the discussed solitary $U(1)$ eigenmode (187). One can clearly see how the corresponding contribution diminishes with system size, indicating that there is indeed no finite density of states in the thermodynamical limit, but just a single eigenvalue pinned at zero frequency and protected by gauge invariance.

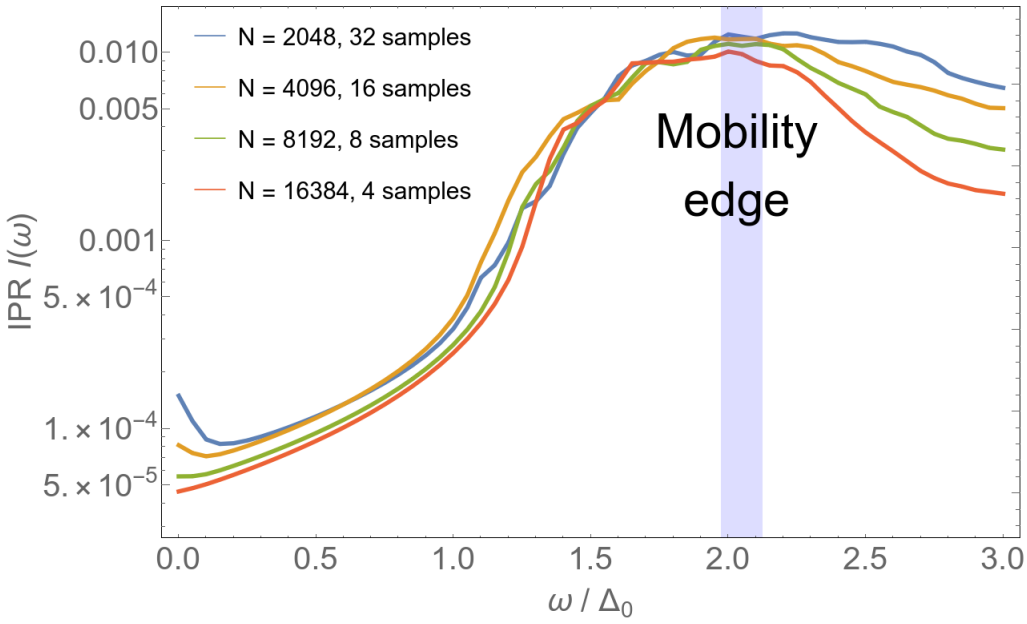


Figure 16: Inverse participation ration $I(\omega)$, as obtained from a numerical diagonalization. Frequency is measured in units of Δ_0. Note that the ordinate axis has a logarithmic scale, which reveals the power-law decay of the IPR with system size at large frequencies, as evident from the uniform spacing of the curves in that region. All parameters of the model, as well as parameters of the numerical procedure are identical to those on Figure 15. The position of the blue strip marks the approximate position of the localization transition w.r.t frequency as deduced from the qualitative change in system-size dependence of $I(\omega)$. As one can also see from from Figure 15, for frequencies lower than $\omega \approx \Delta_0$ the states are completely absent. As a result, the observed IPR below $\omega \approx \Delta_0$ is nonphysical and emerges from a finite value of regularization $\epsilon \approx 0.05\Delta_0$, with an exception of $\omega \approx 0$, where IPR is determined by the solitary $U(1)$ eigenvalue (184), whose contribution, however, vanishes in the thermodynamical limit, as also evident from the decay of the zero frequency density of states on Figure 15.

Information about localization properties on Figure 16 as well as additional numerical experiments performed provide some additional valuable insights:

- As it is evident from Figure 16, one can clearly distinguish both localized and extended modes in the spectrum of the problem. As it generally happens in localization theory, the density of states on Figure 15 shows no qualitative signs of the occurring transition, whose position is indicated on both Figures.

- Quite importantly, the position of the mobility edge is heavily dependent on the particular values of the parameters, as indicated by numerical experiments with different sets of parameters. The fact that this mobility edge falls within a close neighborhood of $2\Delta_0$ is a special feature of the particular set of parameters chosen.

- One should also take into account that the mere presence of localized states creates an additional technical difficulty in determining the overall spectral edge, naively situated at $\omega \approx \Delta_0$ on Figure 15. The main issue is caused by the fact that some mechanisms capable of delivering the existence of localized states are known to demonstrate exponentially slow saturation with system size. As an example, let us consider the general framework of the

so-called optimal fluctuation mechanism, describing the occurrence of spatially limited regions with certain properties that lead to eigenfunction localization within a given region. As one might understand, the probability of such an event decreases exponentially as the size of the region increases. Though, it is always mathematically possible that a particular configuration of a thermodynamically large configuration has an arbitrary large subregion with some special configuration driving the frequency of the resulting eigenstate to zero. As a result, the formal spectrum edge may be much lower or even be zero, but yet with a minuscule value of density of states, indistinguishable from zero from all practical points of view. This is precisely the case detected in [19] for a simplified version of the problem, which is a yet another indicator that such an event might take place in the current formulation of the problem.

- Finally, it is not hard to observe that the power law decay of average IPR with system size at large frequencies is described by $N^{-\alpha}$ with α clearly below 1, contrary to naive expectation suggested by simple estimation (190). A simple fitting procedure suggests that the decay exponent is approximately $\alpha \approx 2/3$. This fact is typically indicative of nontrivial fractal structure of the underlying eigenvectors, yet the presented analysis is too crude to come up with definite claims.

Finally, we note the characteristics of the order parameter distribution for the current set of parameters. The average value of the order parameter is given by $\langle\Delta\rangle \approx 1.66 \cdot \Delta_0$, while the double-exponential front of the distribution is numerically indistinguishable from zero approximately at $\Delta_{\min} \sim 0.3$, meaning that the probability to find Δ below $\Delta_{\min}$ is much smaller than the inverse of the total number of sites in all disorder realizations combined. Therefore it should be noted that the finite width of Δ distribution makes it harder to accurately determine the reference superconducting scales and thus prevents us from making any reliable qualitative interpretations of the data regarding the mechanism responsible for the formation of the spectrum. For instance, one might be tempted to explain the appearance of localized modes by existence of the corresponding optimal fluctuation of local order parameter field, but the data presented does not allow one to confidently make such a claim.

For this reason, we encourage the reader to exercise caution in giving the data any deep physical interpretation. As already mentioned in the beginning of the section, work to expand the practical numerical capabilities of the approach presented is in progress, so that much deeper statistical information would become accessible to both qualitative and quantitative analysis.

11 Simplified model: uniform order parameter

As a first step that provides the relevant scales of all quantities and paves the way towards the full analytical analysis, in this section we will consider a simplified version of the problem. Our approximation will be to completely ignore the exact value of the order parameter in (185) and replace it with a constant value $\Delta_i = \text{const} = \Delta_0$ across the whole system, with Δ_0 being the uniform approximation to the saddle-point equation, as given by mean field approximation (62) The resulting simplified expression for the inverse correlator of transverse modes simply reads:

$$L_{ij} \approx (J/2)^{-1}_{ij} - \Pi^0_i \cdot \delta_{ij}, \quad \Pi^{(0)}_i = \frac{\sqrt{\xi_i^2 + \Delta_0^2}}{\xi_i^2 + \Delta_0^2 - \omega^2/4} \tag{204}$$

In this way, the only source of disorder in the eigenproblem for (204) is that the original disorder fields ξ explicitly enters the expression for the spin response function $\Pi^{(0)}$. The resulting problem becomes a much simpler task both because the distribution of the disorder is thus fully known and because the resulting diagonal disorder Π_i in the target operator (204) is uncorrelated.

While this seems reasonable to do so in the limit of large $v = K\Delta_0$, where the order parameter assumes Gaussian distribution with small variance and quickly decaying correlations, as shown in Subsection 7.1, the results of our analysis in Part III also suggest that this might be a substantially incorrect approximation for smaller values of v. The question at this point is about the typical ranges of v, at which one can eventually observe the appearance of the subgap excitations.

Nevertheless, this simplified model is both of historical and methodological interest, hence we will devote some effort to analyze this model.

11.1 Review of previous studies

The simple model (204) with constant order parameter has already been circulating in the community for a while. For instance, calculations performed in [15] assume constant distribution of the order parameter and only deal with explicit disorder originating from ξ in the definition of the operator (204). A separate attention should be brought to the paper [1] that served as the primary motivation for the present research. In this paper the authors started with the simplified expression (204) for the inverse correlator and reformulated the emerging eigenproblem in terms of a standard eigenvalue problem for a local operator, similarly to how it is done in Subsection 9.3.3. Assuming that the densities of states for both problems demonstrate identical qualitative features (such as the localization edge, the edge of the spectrum), the authors then investigated the resulting local problem by means of Anderson upper limit technique [25], which initially looked perfectly applicable for the analysis of the resulting hopping problem on a random regular graph with diagonal uncorrelated disorder. As an outcome of such an analysis, the authors presented the phase diagram on Figure 17. Most importantly, it implies that one could indeed expect excitations with arbitrarily low energy at sufficiently low values of K, either localized or extended depending on the parameters.

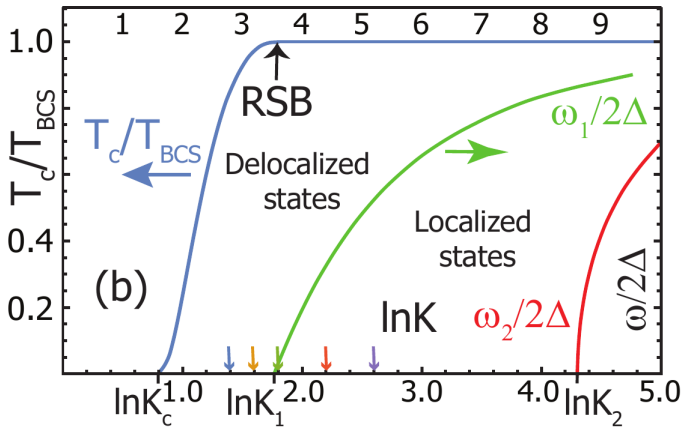


Figure 17: Phase diagram of the transverse modes in constant order parameter approximation, as presented in [1]. For a fixed value of g, the diagram depicts 4 regions of the spectrum in the plane of parameter K and frequency ω: region with no eigenmodes (area to the right of the red curve), region of localized eigenmodes (bounded by green and red curves), region of extended eigenmodes (to the left of the green line). The values of the characteristic values of K are found to be $K_1 = g \exp\{1/2g\} \sim g\sqrt{1/\Delta_0} \gg 1$ and $K_2 = g/4 \cdot \exp\{1/g\} \sim g/\Delta_0 \gg K_1$.

As a consequence of the absence of correlations in the diagonal matrix elements of simplified operator (204), the problem (or its equivalent reformulations) also admits extensive numerical analysis by means of population dynamics, as extensively discussed in the previous iteration of this research [19]. The method essentially solves the integral equation on the full distribution of the diagonal matrix elements of the Green function. The results of this numerical approach have revealed that the Anderson upper limit condition used in [1] is a rather crude approximation that fails to accurately describe the spectral properties of the simplified operator (204). The main reason for such a result is that the system always finds itself at the point of localization transition up to leading order in $1/K$, which implies that all physics including the localization transition itself is potentially sensitive to any effects given by subleading order in formal $1/K$ expansion.

We should note that due to the lack of correlations in diagonal disorder in (204) the problem admits exact solution by means of supersymmetric approach, as presented in [26]. The ability of the aforementioned numerical method to reproduce the full statistics of the diagonal matrix elements of the Green function can be traced back to the same ideas that lead to the existence of the exact solution presented in [26].

Finally, we note that in [19] it was pointed out that the described simplistic model destroys the $U(1)$ gauge symmetry of the original operator (185), because constant order parameter does not satisfy the exact saddle-point equation (60) and thus renders the zero-frequency mode (187) absent.

Nevertheless, below we propose a yet another scheme of analysis for the problem discussed that provides the information on the relevant scale of the problem and serves as a basis for a more accurate technique, that takes into account the full distribution of the order parameter.

11.2 Self-consistent non-Born approximation (SCnBA)

The technique that is about to be presented in this section can be understood as a logical extension of simple Self-consistent Born approximation (SCBA) for the case of strong impurities

(i.e., the ones that are beyond Born approximation for the scattering amplitude) that are still assumed to be sparse.

Such an approach rests on the following qualitative consideration, already familiar to the reader from Part III:

- Most of the time one has $\xi \sim 1$ with the resulting amplitude of the disorder being of order

$$\Pi\left(\xi \sim 1\right) \sim \frac{|\xi|}{|\xi|^2} \sim \frac{1}{|\xi|} \sim 1,$$

  which is a small disorder given the fact that each site has a large number of neighbors $K \gg 1$.

- However, with an exponentially small probability of order $\Delta_0 \ll 1$ the value of ξ falls within a narrow region of several Δ_0, causing the resulting disorder field to be large:

$$\Pi\left(\xi \sim \Delta_0\right) \sim \frac{1}{\Delta_0} \gg 1 \tag{205}$$

In other words, a situation can roughly be described as a set of strong impurities scattered at an exponentially small rate across the system, which seems to be a good basis for the aforementioned SCBA-like treatment. Of course, the situation is complicated by the slow decay of $1/\xi$ dependence, so that the background weak disorder from the majority of the impurities can add up to only logarithmically large effects, but those are not expected to change the qualitative situation.

The framework of self-consistent approximation originally deals with the notion of the self-consistent self-energy defined as

$$\langle G \rangle^{-1} = G^0 - \Sigma, \tag{206}$$

with $\langle G \rangle$ representing the target average of the Green function, i.e., $\langle L^{-1} \rangle$ in our case (we will actively use this notation further on); G^0 being some good initial approximation typically inferred from the properties in the absence of disorder, so that the bare adjacency matrix J of a graph is a good starting approximation to think of. The self-energy operator Σ is then evaluated self-consistently in terms of the average Green function $\langle G \rangle$. There is a number of known approaches to the derivation of such an approximation for Σ, among which the reader might probably be familiar with the Thomas-Fermi approximation or the impurity diagrammatic technique. The latter is discussed in all relevant detail in [3].

As it follows from the formulation of the method, this approach originally cannot capture the change of spectral structure occurring due to the presence of disorder, as it is ultimately based on perturbation theory, which by no means can be expected to capture the localization transition. In particular, the SCBA or its extensions are not intended for describing the localized part of the spectrum if the initial approximation G^0 contains only extended wave functions.

In what follows, we will actively exploit our favorite limit $K \to \infty$, $\Delta_0 \to 0$ with $v = K\Delta_0$ being fixed.

11.2.1 Analog of the full scattering amplitude for the RMT-like ensembles

For the operator (204) of our interest, one can easily construct the celebrated SCBA: the self-energy is given by the mean value and the variance of the disorder corresponding to the the

first 2 diagrams on Figure 18 in the impurity technique:

$$\Sigma_{ij} = \delta_{ij} \left[\langle \Pi \rangle_\xi + \left\langle \left\langle \Pi^2 \right\rangle \right\rangle_\xi \langle G \rangle_{ii} \right], \tag{207}$$

where $\langle\langle \Pi^2 \rangle\rangle$ stands for variance of Π. This expression results from neglecting two types of diagrams in the full perturbation series for the self-energy:

- Diagrams with interleaving scatterings on different impurities, also referred to as the vertex correction. The simplest contribution of that kind is given by the lower left diagram on Figure 18 and reads

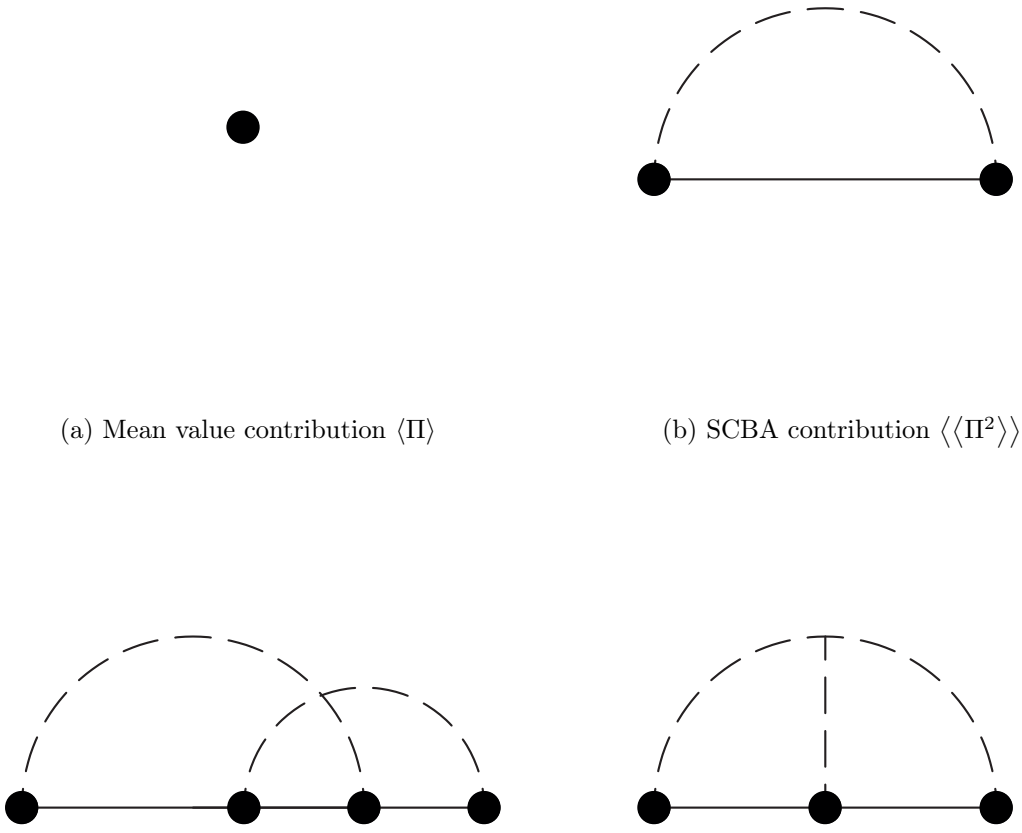
$$\delta\Sigma_{ik}^{(\text{vert.})} = \langle \Pi_i \langle G \rangle_{ik} \Pi_k \langle G \rangle_{ki} \Pi_i \langle G \rangle_{ik} \Pi_k \rangle = \left[\left\langle \Pi^2 \right\rangle \right]^2 \langle G \rangle_{ik}^3 \tag{208}$$

  The name originates from visual analogy of the impurity technique with the usual perturbation analysis of pair-wise interaction, in which case the corresponding diagram does indeed take the form of a vertex correction. We will later show that in our problem these corrections are as small as K^{-C} relative to the leading contribution, where C is the number of crossings of impurity lines. The situation is similar to what one observes in theory of dirty metals, where each crossing of impurity lines delivers another power of $(k_F l)^{-1} \ll 1$ [3].

- More importantly, the diagrams containing more than 2 scatterings on the same impurity are also neglected. The simplest such diagram is present on the bottom right of Figure 18 and reads:

$$\delta\Sigma_{ij}^{(\text{scat.})} = \delta_{ij} \cdot \langle \Pi_i \langle G \rangle_{ii} \Pi_i \langle G \rangle_{ii} \Pi_i \rangle = \delta_{ij} \left\langle \Pi^3 \right\rangle \cdot \langle G \rangle^2 \tag{209}$$

  As it can be shown for our problem, this and higher order diagrams of the same structure are all of the same order as the original contribution (207). This means that one has to collect and sum all such contributions. In the theory of dirty metals, this can be done by calculating the full scattering amplitude on a single impurity and then replacing the Born approximation with this exact value [3]. However, in our case some additional technical effort has to be put in, as it is explained below.



(a) Mean value contribution $\langle \Pi \rangle$

(b) SCBA contribution $\left\langle \left\langle \Pi^2 \right\rangle \right\rangle \cdot \langle G \rangle_{ii}$

(c) Vertex correction $\langle \Pi_i \langle G \rangle_{ik} \Pi_k \langle G \rangle_{ki} \Pi_i \langle G \rangle_{ik} \Pi_k \rangle$

(d) Higher order process $\langle \Pi_i \langle G \rangle_{ii} \Pi_i \langle G \rangle_{ii} \Pi_i \rangle$

Figure 18: Illustrations for the types of diagrams and corresponding contributions to the self-energy within the impurity technique.

To make a generalization similar to replacing Born scattering amplitude with the full one in the conventional theory of metals, one has to exercise some caution. In case of metals one deals with the case of randomly distributed identical impurities characterized by some short-range potential $V(r)$ identical across all impurities—we will refer to this case as the "metal-like ensemble". In our problem the diagonal disorder represents the ensemble of regularly positioned impurities with random magnitude—which we will henceforth call "RMT-like ensemble". In terms of a continuous problem the latter ensemble would correspond to something akin to a random weak potential drawn from some functional distribution (e.g., Gaussian), rather than a sum of copies of a fixed function translated to random positions. While within the SCBA these two ensembles are quite unsurprisingly identical in all physical properties, the corrections from higher order processes have different structure in these 2 ensembles.

Now, in a metal-like ensemble higher order diagrams correspond to higher numbers of scattering processes on the same impurity and thus add up to the full scattering amplitude of a particle, with Green function given by its average value on a solitary potential $V(r)$. In RMT-like ensemble the situation is different, as higher order diagrams correspond to higher cumulants of the onsite disorder potential $\langle\langle \Pi^n \rangle\rangle$, so that the self-energy reads

$$\Sigma_{ij} = \delta_{ij} \cdot \sum_{n=0}^{\infty} \left\langle\left\langle \Pi^{n+1} \right\rangle\right\rangle \left[\langle G \rangle_{ii}\right]^n, \tag{210}$$

where $\langle\langle \bullet \rangle\rangle$ denotes the cumulant of order n of a random variable Π. As one can see, the whole series does not evaluate to a simple object. This fact can be intuitively understood by considering some higher order contributions to Σ and noting that upon collecting the contributions corresponding to the average Green function itself (so called rainbow diagrams), the remaining series of Σ should contain no two identical impurities in a row, as those are already counted in the average Green function. However, instead of classifying all types of diagrams contributing to the self-energy, we will demonstrate this claim later in Section 12 using a functional approach.

At this point, however, a specific character of our problem comes into play. Namely, let us consider the formal expansion of Σ in powers of small quantity Δ_0, just as we did with the cumulant generating function for the order parameter in Subsection 7.2. Then each diagram in the aforementioned series for Σ can also be expanded in this way. And exactly in the same way as it happened during our consideration of the order parameter distribution, all higher order cumulants will be identical to the moments in the leading order, because all averages over ξ gain their contribution in a small region $\xi \sim \Delta_0$, i.e.,

$$\left\langle\left\langle \Pi^n \right\rangle\right\rangle = \left\langle \Pi^n \right\rangle \left(1 + O\left(\Delta_0\right)\right), \; n \geq 2 \tag{211}$$

In other words, in our particular problem the self-consistent self-energy is still equal to the exact scattering amplitude in the leading order w.r.t small parameter Δ_0. As a result, we have

$$\Sigma_{ij} = \delta_{ij} \cdot \sum_{n=0}^{\infty} \left\langle \Pi \left(\langle G \rangle_{ii} \Pi\right)^n \right\rangle_{\xi} + O\left(\Delta_0\right) = \delta_{ij} \cdot \left\langle \frac{\Pi}{1 - \Pi \langle G \rangle_{ii}} \right\rangle_{\xi} + O\left(\Delta_0\right), \tag{212}$$

with the leading term containing contributions of the order $O\left(\ln 1/\Delta_0\right) = O\left(1/g\right)$ and $O\left(\Delta_0^0\right)$, with the former originating from the same type of logarithmic divergence at large ξ as in Part III and the latter computed by replacing the limits of integration over ξ with infinity, in full analogy with Part III:

$$\Sigma_{ij} = \left[\ln \frac{2}{\Delta_0} + \nu_0 \cdot t\left(\mathcal{G} = \frac{\langle G \rangle_{ii}}{\Delta_0}; \zeta = \frac{\omega}{2\Delta_0}\right)\right] \delta_{ij} + O\left(\Delta_0\right), \tag{213}$$

$$t\left(\mathcal{G}; \zeta\right) = \int_0^{\infty} dx \cdot \left[\frac{\sqrt{x^2+1}}{\left(x^2+1\right) - \zeta^2 - \mathcal{G}\sqrt{x^2+1}} - \frac{1}{\sqrt{x^2+1}}\right] = \frac{\left(\zeta^2 + \mathcal{G}c_+\right) f\left(c_+\right) - \left(\zeta^2 + \mathcal{G}c_-\right) f\left(c_-\right)}{c_+ - c_-}, \tag{214}$$

where $\nu_0 = 2P_{\xi}\left(\xi = 0\right)$ is a familiar density of states of the original disorder at the Fermi level and

$$f\left(c\right) = \frac{1}{\sqrt{1-c^2}}\left(\frac{\pi}{2} + \arcsin c\right), \quad c_{\pm} = \frac{1}{2}\left(\mathcal{G} \pm \sqrt{\mathcal{G}^2 + 4\zeta^2}\right) \tag{215}$$

with the branch selected in accordance with the rule $\sqrt{u+i0}=+\sqrt{u}+i0$. As a result we can also observe that both frequency ζ itself and the target Green function $\langle G\rangle$ enter the problem as natural dimensionless combinations, thus also indicating the proper scaling in the $\Delta_0\to 0$ limit. Indeed, in accordance to our expectations to obtain a proper low energy field theory, both the correlator G and the frequency ω enter the expression as natural dimensionless combinations:

$$[\omega]=[\Delta]\,,\quad [G(\omega)]=\frac{[\Delta]^2}{[\Delta]}=[\Delta]\,, \tag{216}$$

where $[\bullet]$ denotes the typical scale of the corresponding quantity, and in the second expression one dimension of $[\Delta]$ in the denominator occurs from the fact that upon switching to Fourier image the correlator loses one energy dimension due to time integration in the Fourier integral.

11.2.2 Self-consistency equation for the diagonal matrix element of the Green Function

To complete the theory, it only remains to calculate the value of the diagonal matrix element of the average Green function according to (206). This can is easily be done by calculating the diagonal matrix element of the average Green Function, or equivalently by calculating its trace due to expected translational invariance. We opted for the latter since it is straightforward to compute as the self-energy is given by a scalar matrix. Exploiting the McKay's expression for the density of states of a large RRG (38), one obtains:

$$\mathcal{G}=\frac{1}{\Delta_0}\cdot\frac{1}{N}\mathrm{Tr}\left\{\langle G\rangle\right\}=\frac{1}{\Delta_0}\int\frac{\rho(\lambda)\,d\lambda}{\frac{1}{g}\frac{K}{\lambda}-\Sigma}=\left/\rho(\lambda)=\frac{K+1}{2\pi}\frac{\sqrt{4K-\lambda^2}}{(K+1)^2-\lambda^2}\right/=\frac{g}{v}\cdot F(g\Sigma;K)\,, \tag{217}$$

where $v=K\Delta_0$ is a warmly welcome dimensionless parameter of the theory, $\mathcal{G}=G_0/\Delta_0$ is properly rescaled diagonal matrix elements of the Green function, and F is given by

$$\begin{aligned}F(s;K)&=K\int\limits_{-1}^{1}\frac{K+1}{2\pi}\frac{4K\sqrt{1-y^2}}{(K+1)^2-4Ky^2}\frac{1}{\frac{K}{2\sqrt{K}y}-s}dy\\&=\frac{2Ks}{\frac{K}{K+1}K-2s^2+\frac{K}{K+1}\sqrt{K^2-4Ks^2}}\end{aligned} \tag{218}$$

Finally, as long as we expect the universal low-energy limit to take place, one expects s to be of order $O(\Delta_0^0)$ (up to logarithmic accuracy), so that in the large K limit the expression above simplified to just

$$F(s;K)\approx s,\quad s\ll\sqrt{K} \tag{219}$$

Substituting the expression (213) for the self-energy, we finally arrive to a simple equation for the diagonal matrix element of the average Green Function:

$$\mathcal{G}=\frac{g}{v}\left[1+\nu_0 g\cdot t(\mathcal{G};\zeta)\right], \tag{220}$$

with t given by (214). Solving this simple algebraic equation for $\mathcal{G}$ for a given ζ yields a correct description for the delocalized part of the spectrum of the simplified operator (204). We also note that in full analogy with what we have observed in III, the theory is only controlled by 4 parameters: Δ_0, v, g, ν_0.

11.2.3 Results for the imaginary part of the correlator

Plotted on Figure 19 is the imaginary part of the diagonal matrix element of the average correlator $\langle G \rangle = \langle L^{-1} \rangle$, as found from the SCnBA equation (220).

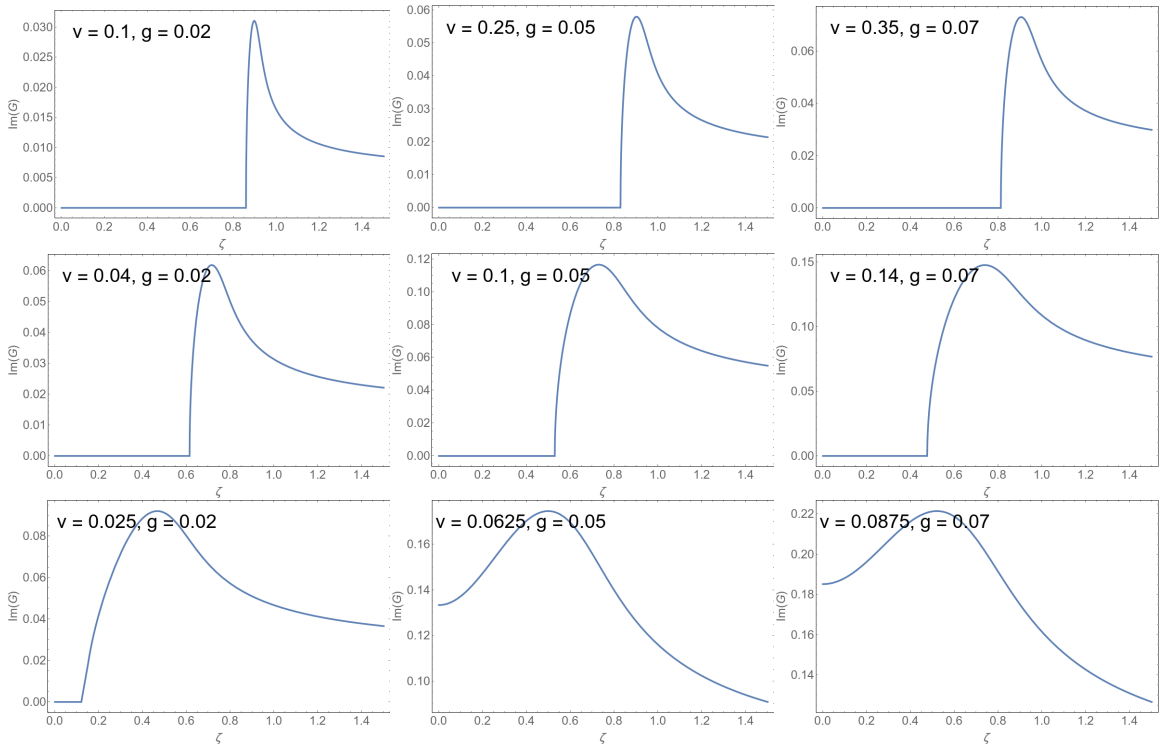


Figure 19: Plots of rescaled imaginary part of the average $\langle G \rangle / \Delta_0$ as a function of dimensionless frequency $\zeta = \omega / 2\Delta_0$ for various values of v and g, as found from the SCnBA equation (220). The plots are sorted by increasing g (horizontal axis) and decreasing v (vertical axis).

The demonstration reveals an important set of qualitative features of the theory:

- The general form of the dependence at moderately small v resembles a celebrated smeared BCS peak with a sharp spectrum edge, originally presented in [27] for the problem of electronic Green function in a superconducting system with magnetic impurities. Closer to our topic, similar profile was obtained in [28] for the electronic Green function with frozen disorder of the order parameter. The direct comparison is not applicable, however, because we are dealing with the correlator of fluctuations on top of the saddle-point configuration of the order parameter, while the Gor'kov's electronic Green function is purely a saddle-point object, corresponding to the mean field semionic propagator in our theory.

- At large v one can observe the sharpening of the BCS singularity and growth of spectral gap towards its mean-field value of $2\Delta_0$, as qualitatively expected from this limit. Because in the limit of large v constant Δ_0 approximation becomes ever more applicable, as follows from our results in Part III, the described treatment demonstrates how a transition to a text-book situation occurs. We also note here that one should not be confused with the apparent increase of peak width and decrease of spectral gap with the increase of g, because having a constant v while increasing g also implies decreasing K to keep the value of v intact.

- However, at small v one also expects the appearance of a considerable tail of localized states instead of a sharp spectrum edge, as our numerical results in Section 10 suggest. However, not only are those states inaccessible by the constant Δ_0 approximation, but it also is the fault of the approach that fails to reproduce localization properties, as announced earlier.
- Additionally, the model shows that the spectral gap closes at sufficiently small v, that are yet well above the superconductor-insulator transition itself. While one does not expect the simplistic approximation of constant Δ_0 to reproduce the gap closure quantitatively accurately, this is an indicator that the low energy transverse fluctuations can indeed be pushed to arbitrarily low energies, as it was initially suggested in [1].

11.2.4 Position of the spectral edge of extended states

Within the simple equation (220), one can analytically identify the position on the spectrum edge. As it follows from the definition of the density of states, absence of any states tautologically corresponds to existence of a purely real stable solution of equation (220). By means of graphical analysis one can establish that the r.h.s of this equation represents a concave real function for real arguments. Therefore, the moment when equation (220) loses real solutions corresponds to the moment when the straight line in the l.h.s of (220) touches the curve corresponding to the r.h.s. This can be formulated by the following tangency condition:

$$1 = \frac{\partial}{\partial \mathcal{G}} \left\{ \frac{g}{v} \left[1 + \nu_0 g \cdot t\left(\mathcal{G}; \zeta\right) \right] \right\} = \nu_0 g \cdot \frac{\partial t}{\partial \mathcal{G}} \tag{221}$$

Being solved simultaneously with (220) w.r.t values of both $\mathcal{G}$ and ζ, this equation provides the boundary value of frequency ζ at which the spectrum edge is situated. The visualization of the solution is presented in Figure 20.

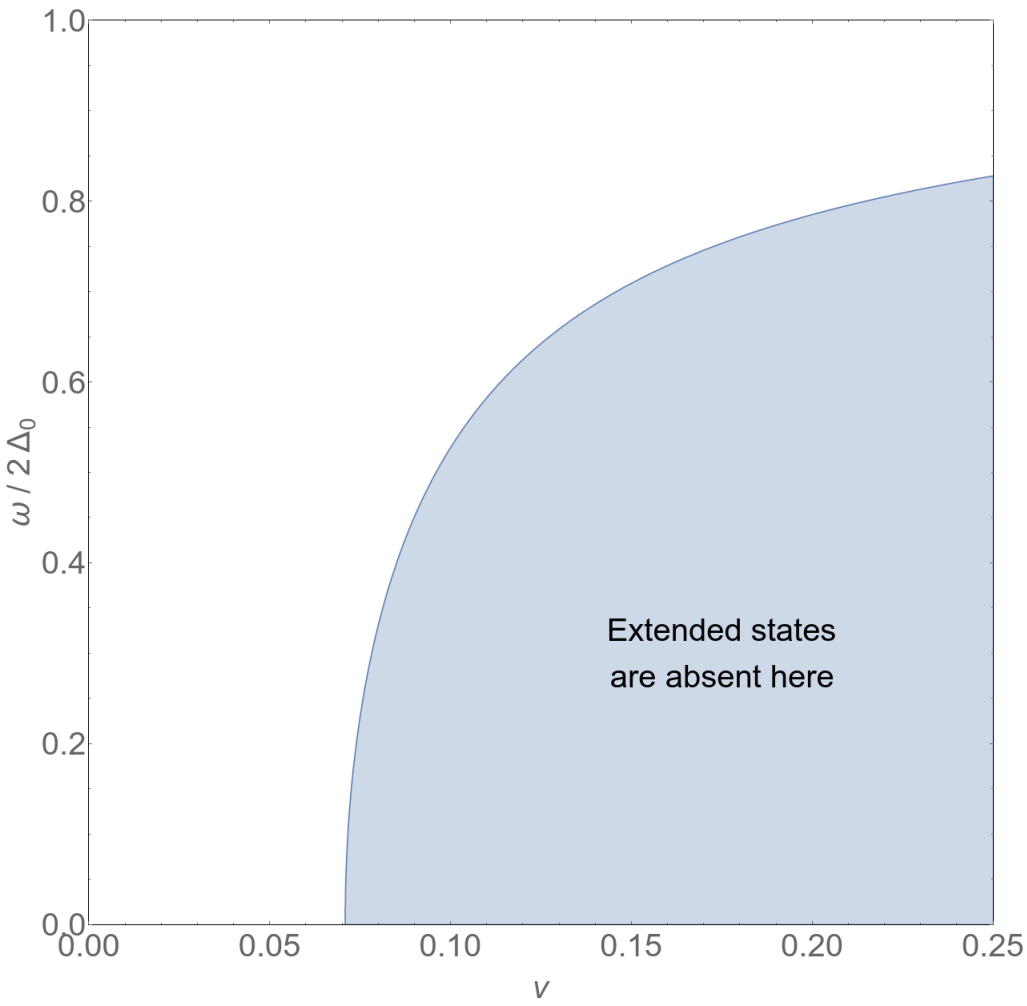


Figure 20: Phase diagram of the SCnBA Green function in $v - \zeta$ variables for $g = 0.05$. For a constant g, v literally translates to the value of K, so that a comparison with Figure 17 can be conducted.

One particular observation is that the resulting phase diagram in $\zeta - v$ plane can be qualitatively compared to Figure 17. We can see that for sufficiently small v the gap does indeed close off, as already showcased in the previous section. Yet we obtain at our disposal a quantitative instrument to characterize this event. At this point, we also note that the technical approach being discussed is similar to that of [28].

One final remark is that the equation allows a full analytical analysis in the limit of small g. In particular, one can analytically extract the value of v at which the gap closure occurs, as it corresponds to seeking a real solution for $\mathcal{G}$ to the system (220 - 221) with $\zeta = 0$. The result reads

$$v_{\text{cl.}} = g \cdot \left[1 + O\left(g^{2/3}\right)\right], \tag{222}$$

with subleading corrections also accessible analytically. While this result is analytically exact for the present theory with constant Δ_0, one expects it to be different in the full theory, as $v \sim g$ is already well beyond the regime of Gaussian fluctuations of the order parameter, as we have seen in Part III. However, the result (222) is still meaningful as it provides the relevant scales of parameters at which the spectral gap closes. In other words, one expects the exact answer to be given by the same type of parametric dependence, but potentially with a different numerical

prefactor. Finally, the form of the subleading correction $O\left(g^{2/3}\right)$ also indicates that at finite values of g one can expect significant numerical deviation from the leading asymptotic. For instance, one can see from Figure 20 that already for $g = 0.05$ this correction constitutes about 30% of the answer.

11.2.5 Estimation of contributions from vertex corrections

Finally, let us revisit the announced smallness of vertex corrections. There are two major effects stemming from the existence of such corrections:

- direct corrections to the diagonal matrix elements of the self-energy
- induction of off-diagonal matrix elements of the self-energy, absent within the SCnBA due to the absence of correlations in disorder.

We start with the first effect using the example of the leading vertex correction (208) to the SCBA contribution in (207). The estimation reads:

$$\frac{\delta\Sigma_{ii}^{(\text{vert.})}}{\Sigma_{ii}^{(\text{SCBA})}} = \frac{\left[\langle\Pi^2\rangle\right]^2 \langle G\rangle_{ii}^3}{\langle\langle\Pi^2\rangle\rangle_\xi \langle G\rangle_{ii}} \approx \langle\Pi^2\rangle \langle G\rangle_{ii}^2 \sim \frac{1}{\Delta_0} \cdot \Delta_0^2 \cdot \mathcal{G}^2 = \Delta_0 \cdot \mathcal{G}^2 \ll 1,$$

where in the third expression we used that $\langle\langle\Pi^2\rangle\rangle \approx \langle\Pi^2\rangle \sim \Delta_0^{-1}$, and the last estimate is valid since $\mathcal{G} = O(1)$, as we have seen previously—it can be shown that $\mathcal{G}$ never exceeds $O(1/g)$, which is still not enough to compensate an exponential smallness of Δ_0. As a result, this correction is by all means negligible. The estimation can be extended to all vertex corrections.

The second effect can be estimated in a similar way:

$$\frac{\delta\Sigma_{ik}^{(\text{vert.})}}{\Sigma_{ii}^{(\text{SCBA})}} \sim \Delta_0 \cdot \left(\frac{\langle G\rangle_{ik}}{\Delta_0}\right)^2 \tag{223}$$

However, this correction can also be shown to be unimportant, because the off-diagonal matrix elements $\langle G\rangle_{ik}$ obey the same type of scaling with Δ_0, so that $\langle G\rangle_{ik} \sim \Delta_0$. Moreover, out of the normalization of the translationally invariant operators on RRG discussed in Subsection 5.2.4, one can deduce that the off-diagonal matrix elements decay with distance $d = |i-k|$ at least as

$$\frac{\langle G\rangle_{ik}}{\Delta_0} \sim \frac{1}{K^{d/2}} \tag{224}$$

so that the estimate (223) reads

$$\frac{\delta\Sigma_{ik}^{(\text{vert.})}}{\Sigma_{ii}^{(\text{SCBA})}} \sim \Delta_0 \cdot K^{-|i-k|} \tag{225}$$

But as we will show later in Subsection 12.3, to produce a finite contribution to the properly rescaled matrix element of the average Green function $\langle G\rangle$ in the limit of large K, the matrix elements of Σ should scale as

$$\Sigma_{ik} = \frac{O\left(\Delta_0^0\right)}{K^{|i-k|}} \tag{226}$$

Already at this point we can see that the induced off-diagonal matrix element of Σ is negligible. It will be even more so, once we demonstrate in Subsection 12.3 that the actual off-diagonal

matrix elements of the average correlator decay even faster than (224). And again, the same type of estimate can be seen to hold for all other types of vertex corrections to the leading SCnBA expression (210).

As a result, we can see that vertex correction vanishes in the limit of large K and small Δ_0, thus making the remaining theory easily solvable. The situation qualitatively true in a more general case is about to be discussed below.

12 Theoretical approach to average correlator of transverse fluctuations at low energies

Guided by the simple consideration presented above, we now generalize this approach to take into account the exact structure of the order parameter, known to us via the results of Section 8. The main idea persistent in the entire theoretical approach of this work will be to exploit Δ_0 and K as the smallest and the biggest parameters of the theory, respectively. Other than this separation of scales and all its consequences, the approach below does not contain any independent approximations.

In this section we choose to present all major steps of the derivation directly within the text, as it contains a sufficient amount of independent ideas and technical peculiarities leading to the final result. Sections 12.1 - 12.2 contain a detailed derivation of generalization for the SCnBA expression (212) to the case of correlated disorder. Then we present the expression for arbitrary matrix elements of average correlator $\langle L^{-1}(\omega)\rangle$ in Subsection 12.3. Finally, in Subsection 12.4 we formulate the resulting theory for the diagonal matrix element of the average correlator, exact in the discussed limit of small Δ_0 and large K, while in Subsection 12.5 we present a comparison of the theoretical results with those of the numerical study.

12.1 General expression for the self-energy

12.1.1 Functional representation for the average correlator

We start off by rewriting the target correlator as a replica field functional integral:

$$\begin{aligned}&\left[L^{-1}(\omega+i\epsilon)\right]_{ij}\\&=\lim_{n\to 0}\left[\int D\psi\cdot\frac{(+i)\cdot\psi_i^\alpha\psi_j^\alpha}{n}\cdot\exp\left\{-\frac{i}{2}\psi^T\left\{\left(\frac{1}{2}J\right)^{-1}-\Pi(\omega+i\epsilon)\right\}\psi\right\}\right],\end{aligned}\tag{227}$$

where $\psi\in\mathbb{R}^n$ is a static (i.e., without time dynamics) n-component replica field being integrated over, with n eventually approaching zero. The exponent is implied to have a diagonal replica structure. The Gaussian integral then performs the required inversion of the target linear operator, while the replica limit eliminates the partition function of the Gaussian integral by setting it to 1. For brevity, we will henceforth imply the $n\to 0$ limit without stating that explicitly. The convergence of integral (227) is guaranteed by the causality of the diagonal part:

$$\mathrm{Im}\left\{\Pi(\omega+i\epsilon)\right\}=\mathrm{Im}\left\{\frac{\varepsilon_i}{\varepsilon_i^2-\omega^2/4+\mathrm{sgn}\omega\cdot i\epsilon}\right\}=+\epsilon\cdot\frac{\varepsilon_i\mathrm{sgn}\omega}{\left(\varepsilon_i^2-\omega^2/4\right)^2}>0\tag{228}$$

Next step is to perform the averaging over disorder.

$$\begin{aligned}
&\left\langle \left[L^{-1}\left(\omega+i\epsilon\right)\right]_{ij}\right\rangle \\
&=\left\langle \int D\psi\cdot\frac{(+i)\cdot\psi_i^\alpha\psi_j^\alpha}{n}\cdot\exp\left\{-\frac{i}{2}\psi^T\left\{\left(\frac{1}{2}J\right)^{-1}-\Pi\right\}\psi\right\}\right\rangle \\
&=\int D\psi\cdot\frac{(+i)\cdot\psi_i^\alpha\psi_j^\alpha}{n}\cdot\exp\left\{-\frac{i}{2}\psi^T\left[J/2\right]^{-1}\psi\right\}\left\langle\exp\left\{+\frac{i}{2}\psi^T\Pi\psi\right\}\right\rangle
\end{aligned}\tag{229}$$

Therefore, we need to compute the following average

$$\exp\left\{-iS_{\text{dis}}\left[d_j=\psi_j^\alpha\psi_j^\alpha\right]\right\}=\left\langle\exp\left\{+\frac{i}{2}\sum_j\Pi_j d_j\right\}\right\rangle,\tag{230}$$

where we have taken into account that Π is a diagonal matrix. The averaging implies taking into account the exact value of the order parameter field in each disorder realization and evaluating the value of Π according to its original definition (185). In particular, the averaging also takes into account the fact that the values of Π at sites close to each other get entangled via spatial correlations of the order parameter discussed in Section 8.

Next step is to write an explicit expression for the disorder action S_{dis}. By direct inspection one can conclude that it can be written as a power series in terms of cumulants over the exact distribution over disorder:

$$-iS_{\text{dis}}\left[d\right]=\sum_{s=1}^{\infty}\frac{1}{s!}\sum_{\{i_k\}}\left\langle\left\langle\prod_{k=1}^{s}\left[\frac{i}{2}\Pi_{i_k}d_{i_k}\right]\right\rangle\right\rangle,\tag{231}$$

where $\langle\langle\bullet\rangle\rangle$ denotes the corresponding cumulant, and each term in summation over s comprises a summation over all possible sets of s indices $\{i_k\}$, with all of them iterating over all sites of the system.

In this last expression (231), as well as in what follows, by cumulants of multivariate distribution of values of Π at all sites of the system we understand the following: given a set of random variables $\{X_i\}_{i=1}^n$, one defines the following generating function:

$$F\left(\{t_i\}_{i=1}^n\right)=\ln\left\langle\exp\left\{\sum_i t_iX_i\right\}\right\rangle,\tag{232}$$

with $\{t_i\}$ implying that the function F depends on n arguments $t_1,...,t_n$, and average $\langle\bullet\rangle$ is taken over the joint distribution of all X_i. One then defines all possible cumulants as corresponding coefficients of formal Taylor series of F:

$$F\left(\{t_i\}_{i=1}^n\right):=\sum_{s=1}^{\infty}\frac{1}{s!}\sum_{\{i_k\}}\left\langle\left\langle\prod_{k=1}^{s}X_{i_k}t_{i_k}\right\rangle\right\rangle,\tag{233}$$

i.e., function F is a formal generating function for the cumulants. This definition is a logical extension of the special case of a single variable. In this way, one can notice that expression (231) is almost the definition of the joint cumulants. However, one can prove the following recursive

combinatorial formula for the cumulants which connects them with the values of standard moments of the joint distribution:

$$\left\langle \prod_{k=1}^{s} X_k \right\rangle := \sum_{\pi} \prod_{A \in \pi} \left\langle \left\langle \prod_{k \in A} X_k \right\rangle \right\rangle, \tag{234}$$

where π iterates over all possible partitions of set $[1, k]$ into nonempty subsets $\{A_i\} \in \pi$, and cumulants can thus be recursively restored via this equation. For instance:

$$\langle X_1 \rangle = \langle\langle X_1 \rangle\rangle,$$

$$\langle X_1 X_2 \rangle = /\{\pi\} = \{\{1\}, \{2\}\}, \{\{1,2\}\}/ = \langle\langle X_1 \rangle\rangle \langle\langle X_2 \rangle\rangle + \langle\langle X_1 X_2 \rangle\rangle$$

So that

$$\langle\langle X_1 X_2 \rangle\rangle = \langle X_1 X_2 \rangle - \langle X_1 \rangle \langle X_2 \rangle$$

is just a covariance. Going further,

$$\begin{aligned} \langle X_1 X_2 X_3 \rangle &= /\{\pi\} = [1,2,3], [12,3], [13,2], [23,1], [123]/ \\ &= \langle\langle X_1 X_2 X_3 \rangle\rangle + \sum_i \langle\langle X_i \rangle\rangle \langle\langle X_j X_k \rangle\rangle + \prod_i \langle\langle X_i \rangle\rangle \end{aligned}$$

so that

$$\begin{aligned} \langle\langle X_1 X_2 X_3 \rangle\rangle &= \langle X_1 X_2 X_3 \rangle + 2 \langle X_1 \rangle \langle X_2 \rangle \langle X_3 \rangle \\ &\quad - \langle X_1 \rangle \langle X_2 X_3 \rangle - \langle X_2 \rangle \langle X_1 X_3 \rangle - \langle X_3 \rangle \langle X_1 X_2 \rangle \end{aligned}$$

and so on.

12.1.2 Diagrammatic technique at the functional saddle-point

This functional integral (227) along with the explicit expression for disorder action (231) spans a certain diagrammatic technique that corresponds to the perturbation series around the trivial saddle-point $\psi \equiv 0$ of functional integral (227). Below we are going to state the resulting diagrammatic rules with a few brief comments that clarify some technical details, while the full derivation is done in a standard fashion. The reader might consult [17, pt. 5] for a typical framework of deriving the diagrammatic rules from the functional integral formulation. The rules of the technique are the following:

- The propagator of the ψ-field is read off from the quadratic part of the full action of the functional integral:
$$G_0 = \left([J/2]^{-1} - \langle \Pi \rangle \right)^{-1} \cdot \delta^{\alpha\beta}, \tag{235}$$
where the last multiplier indicates diagonal structure in replica space.

- Higher order terms in the disorder action 231 correspond to interaction vertices of the associated order $s \geq 2$:
$$\Gamma^{(s)}_{\{i_k\}} = \left\langle \left\langle \prod_{k=1}^{s} \Pi_{i_k} \right\rangle \right\rangle, s \geq 2 \tag{236}$$

- Because of the replica structure, only diagrams with no loops of ψ propagators contribute to the final answer. Indeed, each such loop is proportional to a number of replicas n due to an independent contraction over replica indices along the loop, and thus is eliminated by the replica limit—at least as long as the corresponding integral is not divergent in the replica limit, which is the case in our theory.

- Each remaining diagram is then represented by lines of ψ-propagators and interaction vertices with s pairs of external ends, with each line corresponding to $(-i)\, G^{(0)}$ and each vertex corresponding to $i^s \Gamma^{(s)}$, and because all diagrams are loopless there are no combinatorial coefficients whatsoever. The latter is the case because:

  - In each order s the coefficient in the series for S_{dis} eliminates the one coming from permutation of d fields in the expression for the perturbation itself.
  - In cases when there are more than one identical vertices in the expression which then can be permuted, such a permutation is compensated by the coefficient coming from the expansion of the exponent of the disorder action S_{dis} in power series.
  - Finally, the coefficient $1/2$ in the definition of S_{dis} is compensated by permutation of ψ fields inside d.

- We can make use of one more simplification: because there are no loops, we can essentially drop the coefficients $\pm i$ in front of all expressions, as they will always exactly compensate each other. In other words, we can identify each line with $G^{(0)}$ and each vertex with $\Gamma^{(s)}$. To understand this, one can consider a simple counting argument: each d fields attached to interaction vertex Γ generates an i multiplier, but it also represents a pair of ψ fields to be paired. Each such pairing generates a $-i$ multiplier from the corresponding propagator. In every diagram the total number of ψ fields to be paired is $2 + 2 \cdot \#\,(d)$, with $\#\,(d)$ being a shorthand for "number of d fields". Hence the total multiplier of a diagram is

$$(+i)^{\#(d)} \cdot (-i)^{\#(\psi)/2} = (+i)^{\#(d)} \cdot (-i)^{1+\#(d)} = (-i)\,, \tag{237}$$

  which is precisely the multiplier in front of the expression for the propagator line, so that it can be simultaneously eliminated along with the aforementioned redefinition of lines and vertices.

12.1.3 Expression for the self-energy in absence of vertex corrections

Within the described diagrammatic technique, the self-energy for the average correlator arises naturally as a sum of all single-particle irreducible diagrams by virtue of the Dyson equation:

$$\left\langle L^{-1} \right\rangle^{-1} =: G = \left[G^{(0)} \right]^{-1} - \Sigma, \tag{238}$$

where we have introduced a shorthand for the target average correlator for brevity. Note that there is an averaging in between the two inversion operations, so that the r.h.s should not be evaluated to $\langle L \rangle$.

First several orders of the whole series for Σ read:

$$\begin{aligned}
&\Sigma_{ij} \\
&= \Gamma^{(2)}_{ij} G^{(0)}_{ij} \\
&+ \Gamma^{(3)}_{ik_1 j} G^{(0)}_{ik_1} G^{(0)}_{k_1 j} \\
&+ \Gamma^{(2)}_{ij} G^{(0)}_{ik_1} \Gamma^{(2)}_{k_1 k_2} G^{(0)}_{k_2 j} + \Gamma^{(2)}_{ik_1} G^{(0)}_{ik_2} \Gamma^{(2)}_{k_2 j} G^{(0)}_{k_1 j} + \Gamma^{(4)}_{ik_1 k_2 j} G^{(0)}_{ik_1} G^{(0)}_{k_1 k_2} G^{(0)}_{k_2 j} + ...
\end{aligned} \tag{239}$$

with summation over internal indices k_i of the expressions implied. Next step is to collect all rainbow diagrams (such as, for instance, the third term in the expression above) by replacing the propagator line with the exact average correlator, in the spirit of self-consistent approach discussed in the previous section:

$$\begin{aligned}
\Sigma_{ij} &= \Gamma^{(2)}_{ij} G_{ij} \\
&+ \Gamma^{(3)}_{ik_1 j} G_{ik_1} G_{k_1 j} \\
&+ \Gamma^{(2)}_{ik_1} G_{ik_2} \Gamma^{(2)}_{k_2 j} G_{k_1 j} + \Gamma^{(4)}_{ik_1 k_2 j} G_{ik_1} G_{k_1 k_2} G_{k_2 j} + ...
\end{aligned} \tag{240}$$

This procedure takes into account all diagrams with nested nonintersecting propagator lines, so that they are now excluded from the series.

Finally, one neglects all diagrams that correspond to correction to any interaction vertices, such as the third term in the latter series (240), as all of them can be shown to be at least as small as $\Delta_0/g \to 0$ relative to the leading contribution in the same order. We will omit the explicit demonstration, as it is done in absolutely the same way as presented in Subsection 11.2.5.

As a result, one is left with the following simple series:

$$\Sigma = \Gamma^{(2)}_{ij} G_{ij} + \sum_{s=1}^{\infty} \Gamma^{(2+s)}_{i,\{k_1\}_{l=1}^{s},j} \cdot G_{ik_1} \cdot \left(\prod_{l=1}^{s-1} G_{k_l k_{l+1}} \right) \cdot G_{k_s j}, \tag{241}$$

again, with summation convention over internal indexes k_i in place. At this point, it is convenient to explicitly include $\langle \Pi \rangle$ in the definition of Σ:

$$G^{-1} = [J/2]^{-1} - \Sigma \tag{242}$$

$$\Sigma = \langle \Pi_i \rangle \, \delta_{ij} + \Gamma^{(2)}_{ij} G_{ij} + \sum_{s=1}^{\infty} \Gamma^{(2+s)}_{i,\{k_1\}_{l=1}^{s},j} \cdot G_{ik_1} \cdot \left(\prod_{l=1}^{s-1} G_{k_l k_{l+1}} \right) \cdot G_{k_s j} \tag{243}$$

We can now see that the self-energy is indeed given by a series of cumulants, as promised in the previous section. Expression (210) is then restored by using the fact that when the disorder field Π is uncorrelated on different sites, the joint cumulants are all equal to zero except when all indices coincide, in which case they are reduced to their single-variable counterparts:

$$\left\langle\left\langle \prod_{k=1}^{s} \Pi_{i_k} \right\rangle\right\rangle = \begin{cases} \langle\langle \Pi^s \rangle\rangle \,, & i_1 = i_2 = ... = i_s \\ 0, & \text{otherwise} \end{cases} \tag{244}$$

as can be proved, for example, by direct inspection of the cumulant generating function (232). In this way one obtains that Σ has only diagonal matrix elements of the form given by (210).

12.2 Matrix elements of the self-energy in the limit of small Δ_0

12.2.1 General expression

As the reader might already expect, we will further evaluate all cumulants up to leading power in Δ_0, which will in some specific sense correspond to replacing them with central moments. However, for a generalized case with joint cumulants in a distribution with correlations, this step requires additional technical effort.

We can conveniently rewrite series (243) for the self-energy by using the cumulant generating function, that happens to coincide with S_{dis}:

$$S_{\text{dis}}[d] = i \cdot \langle \Pi \rangle \sum_i \left[\frac{i}{2} d_i \right] + i \cdot \sum_{s=2}^{\infty} \frac{1}{s!} \sum_{\{i_k\}} \Gamma^{(s)}_{\{i_k\}} \cdot \prod_{k=1}^{s} \left[\frac{i}{2} d_{i_k} \right] \tag{245}$$

so that we can obtain each term in series (243) by differentiation:

$$\langle \Pi \rangle = (-i) \cdot \partial_i S_{\text{dis}}(\{d_i\}), \quad \Gamma^{(s)}_{\{i_k\}} = (-i) \cdot \left[\prod_{k=1}^{s} \partial_{i_k} S(\{d_i\}) \right]_{d_i=0}, \tag{246}$$

where ∂_i is the following differentiation operator:

$$\partial_{i_k} := -2i \cdot \frac{\partial}{\partial d_{i_k}} \tag{247}$$

Series (243) then reads:

$$\Sigma_{ij} = (-i) \cdot \left(\left[\delta_{ij} \partial_i + \partial_i \partial_j \left(G_{ij} + \sum_{s=1}^{\infty} \prod_{l=1}^{s} \partial_{k_l} \cdot G_{ik_1} \cdot \prod_{l=1}^{s-1} G_{k_l k_{l+1}} \cdot G_{k_s j} \right) \right] S(\{d\}) \right)_{d=0}, \tag{248}$$

where the r.h.s contains a differential operator that acts on $S_{\text{dis.}}$. This operator can be conveniently rewritten in matrix notation:

$$D_{ij} := \delta_{ij} \cdot \partial_i \tag{249}$$

$$\Sigma_{ij} = (-i) \cdot \left(\left\{ D + D \left[G \sum_{s=0}^{\infty} (DG)^s \right] D \right\}_{ij} S_{\text{dis.}}(\{d\}) \right)_{d=0} \tag{250}$$

implying that the resulting matrix differential operator acts component-wise on S_{dis}, i.e., each component of the matrix Σ is obtained by acting on S_{dis} with the corresponding component of the matrix and then setting $d = 0$ in the resulting expression.

It is now convenient to switch to Fourier transform of the function S_{dis}, as differential operators act trivially on each component. We introduce direct and inverse transforms as

$$S_{\text{dis.}}(\{d_i\}) = \int D\phi \cdot U(\{\phi_i\}) \, e^{+i\phi_i d_i/2}, \quad \int D\omega = \prod_i \int \frac{d\phi_i}{2\pi}, \tag{251}$$

$$U(\{\phi_i\}) = \int Dd \cdot S(\{d_i\}) \, e^{-i\phi_i d_i/2}, \quad \int Dd = \prod_i \int dd_i, \tag{252}$$

where integration is performed over all arguments of $S_{\text{dis.}}$, one for each site of the system. Substituting the result in (250) gives

$$\Sigma = (-i) \cdot \int D\omega \cdot U\left(\{\phi_i\}\right) \cdot \left(\Phi + \Phi\left(G^{-1} - \Phi\right)^{-1} \Phi\right), \quad \Phi_{ij} = \phi_i \cdot \delta_{ij}, \tag{253}$$

where Φ is a diagonal matrix, whose components contain the integration variables, and we have explicitly performed summation of the geometric series.

Now we can exploit the smallness of Δ_0. To do that, we use the same basic idea as in Part III: the quantity $\langle \exp\{-it_i\Pi_i\} - 1\rangle$ is as small as $O(\Delta_0)$ because all integrals over ξ gain the most of their value in a small region $\xi \sim \Delta_0$. Therefore, it is logical to expand the disorder action as:

$$\begin{aligned}
&S_{\text{dis}}\left(\{d\}\right) \\
&= \ln\left\langle \exp\left\{\frac{i}{2}\sum_i d_i\Pi_i\right\}\right\rangle = \ln\left\langle \prod_i \left\{1 + \left[\exp\left\{\frac{i}{2}d_i\Pi_i\right\} - 1\right]\right\}\right\rangle \\
&= /\exp\{-it_i\Pi_i\} - 1 =: \Delta_0 \cdot Y_i/ \\
&= \ln\left\langle 1 + \sum_{A \subset \overline{1,N}} \Delta_0^{|A|} \prod_{i\in A} Y_i\right\rangle = \ln\left\{1 + \sum_{n=1}^{\infty} \Delta_0^n \cdot \frac{1}{n!} \sum_{\{i_k\}-\text{diff}} \left\langle \prod_{k=1}^{n} Y_{i_k}\right\rangle\right\},
\end{aligned} \tag{254}$$

where the sum in the last expressions is done over all sets of different indexes, i.e.,

$$\begin{aligned}
&n = 1 : \text{all } i \\
&n = 2 : i \neq j \\
&n = 3 : i \neq j, k \text{ and } k \neq j \\
&\quad ...
\end{aligned}$$

Now, expanding expression (254) in powers of Δ_0, we arrive at

$$\begin{aligned}
&S_{\text{dis}}\left(\{d\}\right) \\
&= \sum_i \langle \exp\{i\Pi_i d_i/2\} - 1\rangle \\
&+ \frac{1}{2}\sum_{i\neq j} \left[\langle \exp\{i\Pi_i d_i/2 + i\Pi_j d_j\}\rangle - \langle \exp\{i\Pi_i d_i/2\}\rangle \langle \exp\{i\Pi_j d_j/2\}\rangle\right] \\
&+ ...
\end{aligned} \tag{255}$$

with n-th term being of order $O(\Delta_0^n)$, essentially one power of Δ_0 per one independent integration over ξ. For the U function (252) this results in:

$$\begin{aligned}
&U\left(\{\phi\}\right) \\
&= 2\pi \sum_i \left[P_\Pi\left(\phi_i\right) - \delta\left(\phi_i\right)\right] \prod_{k\neq i} \delta\left(\phi_k\right) \\
&+ \frac{(2\pi)^2}{2} \sum_{i,j} \left[P_\Pi\left(\phi_i, \phi_j\right) - P_\Pi\left(\phi_i\right) P_\Pi\left(\phi_j\right)\right] \\
&+ ...
\end{aligned} \tag{256}$$

where each successive term is again of order $O\left(\Delta_0^n\right)$, and P_Π is the full probability distribution of the correspondent subset of Π values, i.e.,

$$\begin{aligned}
&P_\Pi\left(\phi\right) = \left\langle \delta\left(\phi - \Pi\left(\xi, \Delta\right)\right)\right\rangle \\
&P_\Pi\left(\phi_i, \phi_j\right) = \left\langle \delta\left(\phi_i - \Pi_i\left(\xi, \Delta\right)\right) \delta\left(\phi_j - \Pi_j\left(\xi, \Delta\right)\right)\right\rangle \\
&...
\end{aligned}$$

with average taken over the exact ensemble of disorder field ξ and emergent values of Δ defined by the saddle-point equation. Finally, expression (253) for Σ takes the following form:

$$\begin{aligned}
&\Sigma \\
&= \sum_m \cdot \int d\Pi \cdot P\left(\Pi\right) \cdot \left\{\Pi \cdot E_m + \Pi \cdot E_i \left(G^{-1} - \Pi \cdot E_i\right)^{-1} \Pi \cdot E_{mi}\right\} \\
&+ \frac{1}{2} \sum_{m \neq n} \int d\Pi_m d\Pi_n \cdot \left[P\left(\Pi_m, \Pi_n\right) - P\left(\Pi_m\right) P\left(\Pi_n\right)\right] \\
&\times \left\{\left(\Pi_m E_m + \Pi_n E_n\right)\right. \\
&\left. + \left(\Pi_m E_m + \Pi_n E_n\right) \left(G^{-1} - \left(\Pi_m E_m + \Pi_n E_n\right)\right)^{-1} \left(\Pi_m E_m + \Pi_n E_n\right)\right\} \\
&+ ...
\end{aligned} \tag{257}$$

where by E_m we denote the matrix basis element $\left(E_m\right)_{ij} = \delta_{im}\delta_{mj}$, the first term in (256) with $\delta\left(\phi\right)$ vanishes because the integrand in (253) is proportional to ϕ, and a shorthand $d\Pi \cdot P\left(\Pi\right)$ stands for integration w.r.t values of fields ξ and Δ according to the full joint probability distribution of the corresponding group of fields $\left\{\xi\right\}, \left\{\Delta\right\}$ with the value of Π set to be $\Pi\left(\xi, \Delta\right)$, i.e,

$$\begin{aligned}
&\int d\Pi \cdot P\left(\Pi\right) \cdot f\left(\Pi\right) = \int d\xi d\Delta \cdot P\left(\xi, \Delta\right) \cdot f\left(\Pi\left(\xi, \Delta\right)\right) \\
&\int d\Pi_m d\Pi_n \cdot P\left(\Pi_i, \Pi_j\right) \cdot f\left(\Pi_i, \Pi_j\right) \\
&= \int d\xi_i d\Delta_i d\xi_j d\Delta_j \cdot P\left(\xi_i, \Delta_i, \xi_j, \Delta_j\right) \cdot f\left(\Pi\left(\xi_i, \Delta_i\right), \Pi\left(\xi_j, \Delta_j\right)\right) \\
&...
\end{aligned}$$

Again, the hierarchy of orders is still in place: each successive term contains one more power of Δ_0 because of additional integration over ξ. The full form of each term can be restored by expanding the series (254) up to corresponding order in U_i and then plugging in the result into (253).

As a result, expression (257) for Σ can be understood as an expansion in the intensity of correlations in the distribution of disorder field Π, with leading order containing just the onsite distribution and thus having no information about the correlations, and all successive terms introducing corrections from correlations between groups of sites with increasing number. For instance, assuming Gaussian distribution for Π would leave the 2 leading terms in (257), as Gaussian distribution is completely characterized by its mean value and pair-wise correlator. The whole expression (257) allows one to control corrections to any such approximation.

12.2.2 Demonstration: small distance matrix elements Σ_0, Σ_1

To better understand the structure of the general series (257) for Σ, let us go through the first 2 terms. According to (257), we first need to calculate the following 2 objects:

$$\Gamma^{(m)}\left(\phi_m\right) = \phi_m E_m + \phi_m E_m \cdot \left(G^{-1} - \phi_m E_m\right)^{-1} \cdot \phi_m E_m, \tag{258}$$

$$\begin{aligned}
&\Gamma^{(m,n)}\left(\phi_m, \phi_n\right) \\
&= \left[\phi_m E_m + \phi_n E_n\right] \\
&+ \left[\phi_m E_m + \phi_n E_n\right] \cdot \left(G^{-1} - \left[\phi_m E_m + \phi_n E_n\right]\right)^{-1} \cdot \left[\phi_m E_m + \phi_n E_n\right],
\end{aligned} \tag{259}$$

where $\left(E_m\right)_{ij} = \delta_{im}\delta_{mj}$. As it appears, the explicit expression for Γ amounts to inversion of matrices of sizes 1 and 2, accordingly. Let us first rewrite the expression back as a geometric series:

$$\begin{aligned}
&\Gamma^{(m,n)}\left(\phi_m, \phi_n\right) \\
&= \left[\phi_m E_m + \phi_n E_n\right] \\
&+ \left[\phi_m E_m + \phi_n E_n\right] \cdot \left\{\sum_{k=0}^{\infty} \left(G\left[\phi_m E_m + \phi_n E_n\right]\right)^k\right\} G \cdot \left[\phi_m E_m + \phi_n E_n\right]
\end{aligned}$$

The first terms can be inspected manually:

$$\begin{aligned}
&\left(\left[\phi_m E_m + \phi_n E_n\right] \cdot \left(G\left[\phi_m E_m + \phi_n E_n\right]\right)^1 G \cdot \left[\phi_m E_m + \phi_n E_n\right]\right)_{ab} \\
&= \phi_m^3 \delta_{am} G_{mm}^2 \delta_{mb} + \phi_m^2 \phi_n \left[\delta_{an} G_{nm} G_{mm} \delta_{mb} + \delta_{am} G_{mn} G_{nm} \delta_{mb} + \delta_{am} G_{mm} G_{mn} \delta_{nb}\right] \\
&+ \phi_m \phi_n^2 \left[\delta_{an} G_{nn} G_{nm} \delta_{mb} + \delta_{an} G_{nm} G_{mn} \delta_{nb} + \delta_{am} G_{mn} G_{nn} \delta_{nb} +\right] + \phi_n^3 \delta_{an} G_{nn}^2 \delta_{nb}
\end{aligned}$$

As we can see, only 4 matrix elements of G are relevant to the problem: G_{nm}, G_{mn}, G_{nn} and G_{nn}. Similarly, Γ_{ab} itself has nonzero matrix elements only for $a, b \in \{m, n\}$. In this case we can project the problem just on these 2 sites, assemble the whole geometric series back together and obtain:

$$\Gamma^{(m,n)}\left(\phi_m, \phi_n\right) = \begin{pmatrix} \Gamma_{mm} & \Gamma_{mn} \\ \Gamma_{nm} & \Gamma_{nn} \end{pmatrix} = \hat{\phi} + \hat{\phi} \cdot \left(\hat{h} - \hat{\phi}\right)^{-1} \cdot \hat{\phi}, \tag{260}$$

$$\hat{\phi} = \begin{pmatrix} \phi_m & 0 \\ 0 & \phi_m \end{pmatrix}, \hat{h} = \begin{pmatrix} G_{mm} & G_{mn} \\ G_{nm} & G_{nn} \end{pmatrix}^{-1}, \tag{261}$$

where $\bullet^{-1}$ implies an inversion of a corresponding 2x2 matrix. The result is easily computed:

$$\begin{aligned}
&\Gamma^{(m,n)}\left(\phi_m, \phi_n\right) = \begin{pmatrix} \phi_m & 0 \\ 0 & \phi_m \end{pmatrix} \\
&+ \frac{1}{\left(h_{mm} - \phi_m\right)\left(h_{nn} - \phi_n\right) - h_{mn} h_{nm}} \begin{pmatrix} \phi_m^2\left(h_{nn} - \phi_n\right) & -\phi_m \phi_n h_{mn} \\ -\phi_m \phi_n h_{nm} & \phi_n^2\left(h_{mm} - \phi_m\right) \end{pmatrix}
\end{aligned} \tag{262}$$

In a similar fashion, one can see that $\Gamma^{(m)}$ evaluates to

$$\Gamma^{(m)}\left(\phi_m\right) = \phi_m + \frac{\phi_m^2 G_{mm}}{1 - \phi_m G_{mm}} \tag{263}$$

At this point it is instructive to restore expression (212) for the uncorrelated case. Starting from the general expression (257), it is obtained by noting the fact that the joint distribution of multiple Π fields factors into a product of onsite distributions:

$$P\left(\Pi_1, \Pi_2, ...\right) = P\left(\Pi_1\right) \cdot P\left(\Pi_2\right) \cdot ... \tag{264}$$

so that that the structure of expression (257) simply eliminates all terms except the first one, which, upon substituting the expression for $\Gamma^{(m)}$, is identical to the uncorrelated approximation (212). A more general statement is that when 2 or more sites are weakly correlated, the structure of expression (257) implies that a term involving the correspondent group of sites will contain this smallness, in addition to that of Δ_0^n coming from integration over ξ. This also provides a natural measure for the strength of correlations:

$$\Delta P = P\left(\{\Pi_i\}\right) - \prod_i P\left(\Pi_i\right) \tag{265}$$

Expression (257) can then be perceived as a formal expansion w.r.t this difference. Because we expect the disorder to be weakly correlated, as our analysis in Part III suggests, the higher order terms in (257) will contain the correspondent smallness.

Note also when the off-diagonal matrix elements of the $G_{nm} \sim G_{mn}$ are small, the answer reduces back to that of a pair of independent sites:

$$\Gamma^{(m,n)}\left(\phi_m, \phi_n\right) = \begin{pmatrix} \Gamma^{(m)}\left(\phi_m\right) & 0 \\ 0 & \Gamma^{(n)}\left(\phi_n\right) \end{pmatrix} + O\left(G_{mn}, G_{nm}\right) \tag{266}$$

And again, the structure of the corresponding term in (257) discards everything that factorizes into a product of functions of a single argument ϕ_m or ϕ_n, so that the first term in 266 is eliminated upon integration over ϕ variables. Therefore, if a given off-diagonal matrix element of G is small, then so is every expression in (257) involving this matrix element.

The generalization of (260) for higher order terms in (257) is now quite straightforward. Suppose we are interested in a term of (257) that contains only sites $i_1, ..i_n$. We then introduce the following projector on the subspace of sites $\{i_k\}$:

$$P = \sum_{k=1}^{n} |v_{i_k}\rangle\langle i_k| \in \mathrm{Mat}\left(n \times N\right), \tag{267}$$

where $\left(E_m\right)_{ij} = \delta_{im}\delta_{mj}$ and v_{i_k} denotes basis vectors in $\mathbb{R}^n$, i.e., in components:

$$\left(u_{i_k}\right)^j = \delta^{kj} \tag{268}$$

The corresponding matrix contribution to be averaged then reads

$$\Gamma\left(\phi_{i_1}, .., \phi_{i_n}\right) = P^T\left(\hat{\phi} + \hat{\phi}\left(\hat{h} - \hat{\phi}\right)^{-1}\hat{\phi}\right)P \in \mathrm{Mat}\left(N, N\right), \tag{269}$$

$$\hat{\phi} = P\Phi P^T, \hat{h} = \left\{PGP^T\right\}^{-1} \in \mathrm{Mat}\left(n, n\right), \tag{270}$$

where inversion in the definition of $\hat{h}$ deals with just an instance of $n \times n$ matrix.

12.2.3 Estimation of typical magnitude of matrix elements

Now, expression (269) allows us to explicitly estimate the magnitude of each matrix element of Σ. The latter is fully characterized by the distance $d = |i - j|$ between the 2 sites i and j. The matrix element Σ_d can be produced by one of the following ways:

- Considering the contribution of order 2 between the sites i and j. Such a contribution would contain smallness coming from that of the corresponding matrix element G_{ij}, as suggested by (266). In the upcoming Subsection 12.3 we will demonstrate that this smallness is of order K^{-d}. Moreover, sites at large distances are weakly correlated according to the results of Part III, so that quantity (265) is also expected to be small, thus adding up to the total smallness of the resulting contribution. In total, we obtain that such contributions are at least as small as
$$\Sigma_{ij}^{(2)} \sim \frac{1}{K^{|i-j|}} \tag{271}$$
- Another way is to consider a group of d sites arranging a chain of sites from i to j. Then there are no issues with correlation intensity or smallness of the required off-diagonal matrix elements of G. However, in this case the smallness is attributed to the fact that each additional site brings additional integration over the value of ξ, which is the primary source of smallness in the whole expression (257). As a result, this type of contributions is estimated as
$$\Sigma_{ij}^{(d+1)} \sim \Delta_0^{|i-j|} \tag{272}$$
and because RRG is a loopless structure there is always only one such contribution.
- Such an analysis can be extended to an arbitrary path (possibly, self-intersecting) connecting sites i and j. Each unique site in such a path adds a power of Δ_0, and if the total site count is $2 \leq a \leq d+1$, the distance d can be covered only by excluding hops at distances larger than 1, which will then introduce a total factor of K^{d-a} via the corresponding smallness of the Green function matrix elements. As a result, any such contribution is estimated as
$$\Sigma_{ij}^{(a)} \sim \frac{\Delta_0^a}{K^{d-a}}, \tag{273}$$
which incorporates both expressions for $a = 2, d+1$.

We thus arrive at the following scaling of the matrix elements of Σ:

$$\Sigma_d \sim \sum_{a=2}^{d+1} k_a \cdot \frac{\Delta_0^{a-1}}{K^{d+1-a}} \sim \frac{\sigma_d(v)}{K^d}, \tag{274}$$

where k_a is some numerical coefficients of order unity, found by means presented in the previous section, and $\sigma_d \sim O(\Delta_0^0)$ is some dimensionless function of the control parameter $v = K\Delta_0$

12.3 Explicit expression for arbitrary matrix elements of the correlator

Next step is to calculate the expression for the matrix elements of the averaged correlator, assuming a known form of the self-energy. To do that, we need to note that upon averaging over disorder one expects all objects to restore their translational invariance in a sense described in Subsection 5.2.4. In other words, one expects all matrix elements of averaged operators,

including G and Σ, to depend solely on the distance between the sites i and j. Having this in mind, we can exploit the full power of the analog of Fourier analysis for random regular graph developed in Subsection 5.2.4. It allows us to write down the explicit expression for any matrix element G_d of the average correlator by direct inversion of (242):

$$\begin{aligned}\frac{G_d}{\Delta_0} &= \frac{g}{v}\cdot\int\limits_{-\pi}^{\pi}\frac{d\theta}{2\pi}\cdot\frac{K}{\frac{K}{2\sqrt{K}\cos\theta}-g\Sigma\left(\theta\right)}\cdot\frac{-2iK\sin\theta}{Ke^{-i\theta}-e^{i\theta}}\cdot\left[\frac{e^{i\theta}}{\sqrt{K}}\right]^{d}\\ &= \frac{g}{v}\cdot\int\limits_{|z|=1}\frac{dz}{2\pi i}\cdot\frac{1}{z}\frac{K}{\frac{\sqrt{K}}{z+z^{-1}}-g\Sigma\left(z=e^{i\theta}\right)}\cdot\frac{K\left(1-z^{2}\right)}{K-z^{2}}\cdot\left[\frac{z}{\sqrt{K}}\right]^{d},\end{aligned}\tag{275}$$

where in the second line we have switched to an integral over a unit circle in the complex plane, and $\Sigma\left(\theta\right)$ is the spectral decomposition of the self-energy matrix, as dictated by (52 - 54):

$$\Sigma\left(\theta\right)=\frac{\Pi\left(\theta\right)\sigma\left(\theta\right)+\Pi\left(-\theta\right)\sigma\left(-\theta\right)}{\Pi\left(-\theta\right)+\Pi\left(\theta\right)},\Pi\left(\theta\right)=\frac{-2iK\sin\theta}{Ke^{-i\theta}-e^{i\theta}},$$

$$\sigma\left(\theta\right):=\mathrm{tr}\left\{\left[\Sigma_{ij}-\frac{1}{N}\Sigma\left(K\right)I_{ij}\right]\cdot\left[\theta\right]^{\left|i-j\right|}\right\},\left[\theta\right]=\frac{e^{i\theta}}{\sqrt{K}},$$

$$\Sigma\left(K\right)=\mathrm{tr}\left\{\Sigma I\right\}=\frac{1}{N^{2}}\sum_{ij}\Sigma_{ij}$$

We now switch to a proper scaling form (274) for the matrix elements of the self-energy:

$$s_d=g\Sigma_d\cdot K^d\tag{276}$$

So that the spectral decomposition is rewritten as:

$$\Sigma\left(K\right)=\frac{1}{g}\tag{277}$$

$$\sigma\left(\theta\right):=\mathrm{tr}\left\{\left[X_{ij}-\frac{1}{N}X\left(K\right)I_{ij}\right]\cdot\left[\theta\right]^{\left|i-j\right|}\right\}=\frac{1}{g}\sum_{d=0}^{\infty}s_d\cdot\left(\frac{e^{i\theta}}{\sqrt{K}}\right)^{d}\tag{278}$$

$$\begin{aligned}g\Sigma\left(\theta\right) &= g\frac{\Pi\left(\theta\right)\sigma\left(\theta\right)+\Pi\left(-\theta\right)\sigma\left(-\theta\right)}{\Pi\left(-\theta\right)+\Pi\left(\theta\right)}=g\frac{\Pi\left(\theta\right)\sigma\left(\theta\right)+\Pi\left(-\theta\right)\sigma\left(-\theta\right)}{\Pi\left(-\theta\right)+\Pi\left(\theta\right)}\\ &= \sum_{d=0}^{\infty}s_d\cdot\frac{1}{K^{d/2}}\frac{z^{d+1}-z^{-(d+1)}}{z-z^{-1}},\end{aligned}\tag{279}$$

where we have explicitly evaluated the $K\to\infty$ limit in everything but the oscillating term containing z. Substituting this back into expression (275) for the correlator renders:

$$\frac{G_d}{\Delta_0}=\frac{g}{v}\cdot\int\limits_{|z|=1}\frac{dz}{2\pi i}\cdot\frac{K}{z}\cdot\left[\frac{\sqrt{K}}{z+z^{-1}}-\sum_{d=0}^{\infty}s_d\cdot\frac{1}{K^{d/2}}\frac{z^{d+1}-z^{-(d+1)}}{z-z^{-1}}\right]^{-1}\cdot\frac{K\left(1-z^{2}\right)}{K-z^{2}}\cdot\left[\frac{z}{\sqrt{K}}\right]^{d}\tag{280}$$

Next step is to notice that all roots of the denominator are of order $K^{\pm 1/2}$, because presence of Σ brings only $O\left(K^{-1/2}\right)$ correction to the leading expression. Moreover, we can see that there is a universal leading asymptotic for all of these roots:

$$\begin{cases} z_0 = O\left(\sqrt{K}\right): & \frac{1}{z}\cdot\left(1-O\left(\frac{1}{K}\right)\right) - \frac{1}{\sqrt{K}}\cdot\sum_{d=0}^{\infty} s_d \cdot \frac{z^d}{K^{d/2}}\left(1+O\left(\frac{1}{K}\right)\right) = 0 \\ z_0 = O\left(\frac{1}{\sqrt{K}}\right): & z\left(1-O\left(\frac{1}{K}\right)\right) - \frac{1}{\sqrt{K}}\cdot\sum_{d=0}^{\infty} s_d \cdot \frac{z^{-d}}{K^{d/2}}\left(1+O\left(\frac{1}{K}\right)\right) = 0 \end{cases} \tag{281}$$

or, unifying the 2 cases, all roots are given by

$$z_0 \approx K^{\pm 1/2}\cdot u_0\left(1+O\left(\frac{1}{K}\right)\right) \tag{282}$$

where u_0 iterates over all roots of the following equation:

$$\frac{1}{u} - \sum_{d=0}^{\infty} s_d \cdot u^d = 0 \tag{283}$$

Let us introduce the generating function of s_d coefficients:

$$\phi(u) = \sum_{d=0}^{\infty} s_d \cdot u^d, \tag{284}$$

which is assumed to be convergent in some neighborhood of $u=0$. We can now evaluate integral (280) using the residue technique. And because all relevant values of z are in the $O\left(K^{-1/2}\right)$ vicinity of zero, we can expand the integrand up to the leading order in powers of K:

$$\begin{aligned} I(z) &= \frac{K}{z}\cdot\left[\frac{\sqrt{K}}{z+z^{-1}} - \sum_{d=0}^{\infty} s_d \cdot \frac{1}{K^{d/2}}\frac{z^{d+1}-z^{-(d+1)}}{z-z^{-1}}\right]^{-1}\cdot\frac{K\left(1-z^2\right)}{K-z^2}\cdot\left[\frac{z}{\sqrt{K}}\right]^d \\ &= \left/ z = \frac{u}{\sqrt{K}} \right/ \approx \frac{\sqrt{K}}{K^{d-1}}\cdot\frac{u^{d-1}}{u-\phi\left(\frac{1}{u}\right)}\cdot\left[1-\frac{u^2}{K}\left(\frac{\phi(0)-2\phi\left(\frac{1}{u}\right)}{u-\phi\left(\frac{1}{u}\right)}\right)+O\left(\frac{1}{K^2}\right)\right], \end{aligned} \tag{285}$$

so that expression (280) can be now rewritten as

$$\frac{G_d}{\Delta_0} = \frac{g}{v}\cdot\frac{1}{K^{d-1}}\cdot\int\limits_{u=\sqrt{K}}\frac{du}{2\pi i}\cdot\frac{u^{d-1}}{u-\phi\left(\frac{1}{u}\right)}\cdot\left[1-\frac{u^2}{K}\left(\frac{\phi(0)-2\phi\left(\frac{1}{u}\right)}{u-\phi\left(\frac{1}{u}\right)}\right)\right] \tag{286}$$

Now the integral possesses only one singularity $u=\infty$ outside the integration contour, which then delivers the value of the integral:

$$\begin{aligned} &\frac{G_d}{\Delta_0} \\ &= -\frac{g}{v}\cdot\frac{1}{K^{d-1}}\cdot\operatorname*{res}_{u=\infty}\cdot\left\{\frac{u^{d-1}}{u-\phi\left(\frac{1}{u}\right)}\cdot\left[1-\frac{u^2}{K}\left(\frac{\phi(0)-2\phi\left(\frac{1}{u}\right)}{u-\phi\left(\frac{1}{u}\right)}\right)\right]\right\} \\ &= +\frac{g}{v}\cdot\frac{1}{K^{d-1}}\cdot\left[(1-\delta_{d,0})\cdot\frac{1}{(d-1)!}\left[\frac{d^{d-1}}{du^{d-1}}\left\{\frac{1}{1-u\phi(u)}\right\}\right]_{d=0} + \delta_{d,0}\frac{\phi(0)}{K}\right], \end{aligned} \tag{287}$$

so that we have explicit formulas for all matrix elements of the correlator

$$\frac{G_0}{\Delta_0} = \frac{g}{v}\phi\left(0\right) \tag{288}$$

$$\frac{G_{d\geq 1}}{\Delta_0} = +\frac{g}{v}\cdot\frac{1}{K^{d-1}}\cdot\frac{1}{(d-1)!}\cdot\left[\frac{d^{d-1}}{du^{d-1}}\left\{\frac{1}{1-u\phi\left(u\right)}\right\}\right]_{u=0} \tag{289}$$

As we have claimed multiple times up to now, the matrix elements of the Green functions indeed have the following scaling form in the limit $\Delta_0 \to 0$, $K \to \infty$, $v = K\Delta_0 = \text{const}$:

$$G_{d\geq 1} \sim \frac{\Delta_0}{K^{d-1}}, G_0 \sim \Delta_0 \tag{290}$$

and thus express exponential decay with distance d.

Additionally, expressions (288 - 289) also prove the two claims made in Subsection 11.2.5 about the role of off-diagonal matrix elements of the self-energy:

- We can see that any contribution to Σ that falls off faster than the scaling form (274) does not contribute to the value of any matrix element G upon proper rescaling
- Expression (288) also demonstrates that the diagonal matrix element of the Green function G_0 is only sensitive to the diagonal matrix element of the self-energy $\phi\left(0\right) = s_0 = g\Sigma_0$. Moreover, because we know that each successive term in (257) has smallness of order Δ_0^n, one Δ_0 for each independent site involved, we can evaluate the value of s_0 using only the first term in (257), as all other terms die out in the scaling limit.
- More generally, matrix element G_d at distance d knows about matrix elements Σ_d as far as distance $d-2$.

12.4 Self-consistency equations on diagonal matrix elements

Let us consolidate everything we have gathered. We have derived a coupled set of equations on the diagonal matrix elements of the target correlator G and its self-energy Σ in the limit

$$\Delta_0 \to 0, \quad K \to \infty, \quad v = K\Delta_0 = \text{const} \tag{291}$$

In this limit, the aforementioned operators possess translational invariance in a sense discussed in Subsection 5.2.4 with the following scaling of the matrix elements:

$$G_0 \sim \Delta_0, \;\; G_{d\geq 1} \sim \frac{\Delta_0}{K^{d-1}}, \;\; \Sigma_d \sim \frac{g^{-1}}{K^d}, \tag{292}$$

so that it is appropriate to introduce the following dimensionless quantities:

$$\mathcal{G}_0 := \frac{G_0}{\Delta_0}, \;\; \mathcal{G}_{d\geq 1} = \frac{G_{d-1}}{\Delta_0}K^{d-1}, \;\; s_d = g\Sigma_d \cdot K^d \tag{293}$$

In terms of these quantities, the equations take the following simple form:

$$\mathcal{G}_0 = \frac{g}{v}\cdot s_0, \tag{294}$$

$$s_0 = \int dxdy \cdot \frac{P(x,y)}{\Delta_0} \cdot \left[\pi(x,y) + \frac{\pi^2(x,y) \cdot \mathcal{G}_0}{1 - \pi(x,y) \cdot \mathcal{G}_0} \right], \tag{295}$$

$$\pi(x,y) = \frac{\sqrt{x^2+y^2}}{x^2+y^2-\zeta^2}, \tag{296}$$

where P is the joint probability distribution of rescaled onsite values of $x = \xi/\Delta_0$ and $y = \Delta/\Delta_0$. One then has to solve this system (257) for s_0 and $\mathcal{G}_0$. As it is demonstrated in subsections 12.2 - Subsection 12.3, there are no perturbative corrections to this set of equations, i.e., they become exact in the limit (291).

By use of the previously derived expression (168) for P one then completes the analytical description of the diagonal matrix element of G. In particular, after some algebra that heavily exploits the analytical properties of P discussed in Part II, one can rewrite equations (294 - 296) via the universal r function:

$$\frac{\mathcal{G}_0}{g/v} = 1 + g \cdot [\Delta\Pi(0) + \Delta\Sigma(\mathcal{G}_0;\zeta)] \tag{297}$$

where $\Delta\Pi$ is a number given by

$$\Delta\Pi(0) = \nu_0 \left[\gamma + \int\limits_{\mathbb{R}-i0} \frac{d\omega}{2\pi} \cdot \frac{\ln i\omega}{i\omega} \cdot \int\limits_0^\infty \frac{dz}{(1+z)^{3/2}} \cdot \exp\{v \cdot r(\omega | f(z))\} \right] \tag{298}$$

and $\Delta\Sigma$ is given by

$$\begin{aligned} \Delta\Sigma(\mathcal{G}_0;\zeta) = \sum_{a=\pm} (ic_a s_a) \cdot \nu_0 \cdot \int\limits_0^\infty \frac{dz}{1+z^2} \cdot \int\limits_0^\infty d\psi \cdot \exp\{v \cdot r(-s_a\psi | f(z))\} \\ \times \left[\exp\left\{ \frac{is_a\varepsilon_a\psi}{\sqrt{1+z^2}} \right\} + \frac{\exp\left\{ \frac{is_a\varepsilon_a\psi}{\sqrt{1+z^2}} \right\} - 1}{\frac{i\varepsilon_a\psi}{\sqrt{1+z^2}}} \left(-z\frac{\partial f}{\partial z}(z) \right) \left[v \cdot \frac{\partial r}{\partial f}(-s_a\psi | f(z)) \right] \right], \end{aligned} \tag{299}$$

$$\varepsilon_\pm = \frac{\mathcal{G}_0}{2} \pm \sqrt{\left(\frac{\mathcal{G}_0}{2}\right)^2 + \zeta^2}, \quad s_\pm = \mathrm{sgnIm}\varepsilon_\pm, \quad c_\pm = \varepsilon_\pm \mp \frac{\zeta^2/2}{\sqrt{\left(\frac{\mathcal{G}_0}{2}\right)^2 + \zeta^2}} \tag{300}$$

Similarly to the simple case of constant Δ approximation, one can observe that equations (294 - 296) possess a purely real solution for $\mathcal{G}_0$ for some values of the parameters, which algebraically can be traced down to the fact that s_0 is a concave function of $\mathcal{G}_0$. In perfect analogy with Subsection 11.2.4, one can then formulate the additional tangential condition that would then describe the position of the spectral edge.

Finally, as it is discussed in subsections 12.2 - Subsection 12.3, it is in principle possible to come up with a similar type of equations for all off-diagonal matrix elements of G, with the same universality taking place in the limit (291). However, the equations rapidly become rather complicated. However, for each finite distance d, the matrix element G_d is described by a closed set of equations involving all $G_{d'}$ with $d' < d$.

12.5 Comparison with the numerical solution

We now calculate the the diagonal matrix element of the average correlator using a particular set of model parameters and compare it with the results of the numerical exact diagonalization. The theoretical calculation is done according to expressions (297 - 300), where r is found via small v approximation, as described in Subsection 8.2.3. Comparison of the theoretical value for $\rho_2 = \mathrm{Im}\left\{\mathrm{Tr}G/N\right\}$ with that obtained from numerical diagonalization is presented on Figure 21.

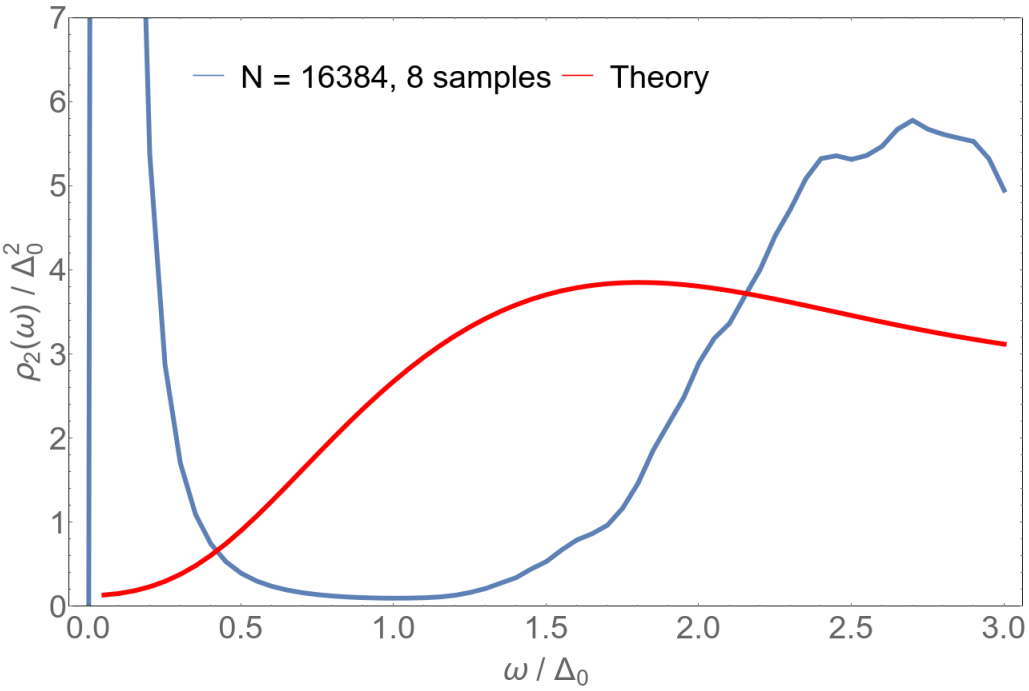


Figure 21: Frequency dependence $\rho_2(\omega)/\Delta_0^2$ of the (properly normalized) imaginary part of the average correlator G, as computed by the exact numerical diagonalization and the theoretical SCnBA method with the value of r restored from the small v approximation. The parameters of the model: $g = 0.21$, $K = 30$, box shaped distribution of disorder $P_\xi(\xi) = \theta(1 - |\xi|)/2$. This corresponds to the following values of the control parameters for the theory: $\nu_0 = 1$, $\Delta_0 = 0.017$ $v = K\Delta_0 = 0.51$. Experimental results were obtained from averaging over 8 samples with size $N = 16384$ with level-broadening $\epsilon = 0.05\Delta_0$. As discussed in the main text, the peak at zero frequency in experimental data is a consequence of finite system sizes and thus is absent in the theoretical curve that represents the infinite N limit.

First of all, let us comment the apparent peak at low frequencies, which is present only in the numerical data. This peak is attributed to a solitary eigenvalue of $U(1)$-symmetry mode (187), whose contribution is supposed to vanish in the thermodynamical limit and thus is absent in the theoretical curve. In the numerical data, however, this peak shows up as a consequence of the fact that the spectral weight (202), expressing the contribution of each eigenmode to the value of ρ_2, contains an explicit $1/\omega$ singularity at low frequencies. However, this singularity is absent in the thermodynamical limit $N \to \infty$ for any finite level broadening ϵ. In the numerical data it survives because large N is compensated by a numerical smallness of ϵ. To illustrate that, the numerical data for the density of solutions ρ_1 for various system sizes is presented on Figure 22. In ρ_1 such a singularity in the spectral weight at small frequencies is absent, and one can thus

see a decrease in $\rho_1(0)$, because it is provided by a single eigenvalue, whose contribution is not extensive with system size.

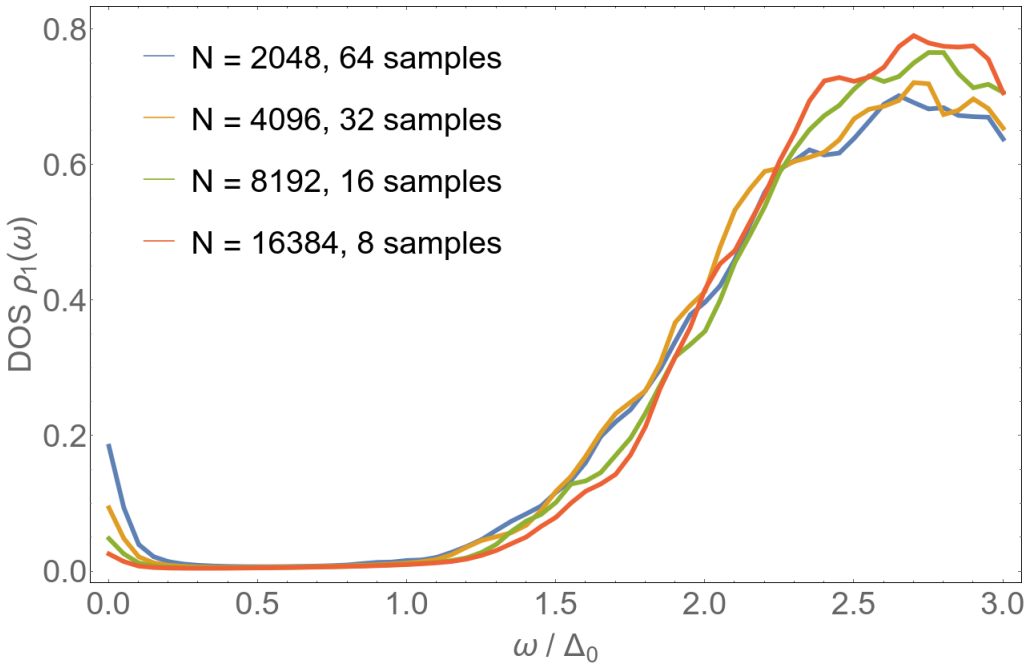


Figure 22: Reference numeric density of solution for the numeric dataset presented on Figure 21. All parameters of the numerical experiment are presented on Figure 21.

Next, we address the fact that the convergence between the theory and the numerical data is not quite as outstanding as was observed in Section 8 for the saddle-point equation itself—we ignore the behavior of the numerical data at zero frequency, as explained in the comments to Figure 21. The explanation lies in a simple fact that the value $v = 0.51$ is not small enough compared to unity to expect accurate results from the small v approximation that we used to compute the r function, thus rendering the quantitative performance of the resulting theory quite poor. A straightforward fix to this problem would be to use the exact set of equations on r, which is also presented in Section 8, but the corresponding numerical routine is not yet implemented. Another way would obviously be to use parameters yielding a smaller value of v, but due to the technical limitations of the exact diagonalization outlined in Section 10, this is not possible at the moment.

With that said, our theoretical approach gives a much better qualitative approximation of the actual distribution, than the constant Δ approximation discussed previously. In particular, the developed model accurately reproduces the smooth form of the imaginary part of the correlator ρ_2 observed numerically, which is also consistent with the behavior of the density of solutions ρ_1.

Part V
Conclusion

13 Achieved results

Below we summarize the immediate results of our theoretical research of the saddle-point solution and small fluctuations of the order parameter within the Anderson Pseudospin model on a random regular graph, presented in Part II and analyzed in Parts III and IV:

1. The complete theoretical investigation has been performed in the limit of small values of the typical superconducting scale Δ_0 and large values of K, in which case the theory demonstrates a well-defined universal behavior

   (a) Exponentially small value of Δ_0 arises naturally due to the assumed smallness of Cooper coupling constant g. The large value of K is then required to preserve the very existence of the superconducting phase, as demonstrated in previous studies. In our research we chose to characterize the superconducting scale by its mean field value.

   (b) All relative physical quantities are then showed to acquire a native characteristic scale determined by the dimensionality w.r.t energy units. Upon proper rescaling, the superconducting energy scale is then removed from the problem, while all other energy scales may well be considered to be infinite, with the only exception of the equation determining the value of Δ_0 itself.

   (c) The emerging theory for the rescaled quantities is then completely described by just 3 control parameters:

      i. Density of states at the Fermi-energy: $\nu_0 = 2P_\xi(0)$
      ii. The value of the Cooper constant g
      iii. Effective number of neighbors within the superconducting scale of energy: $v = K\Delta_0$

2. For both the saddle-point problem and the problem of low energy transverse fluctuations we have developed a numerical approach that allowed us to conduct an explicit verification of all the theoretical results

   (a) For the saddle point equation, we have managed to come up with a good numerical scheme for a solution that allowed us to reach system sizes of up to $2.4 \cdot 10^6$ sites. This enabled us not only to probe onsite distribution of the order parameter itself, but also to reliably probe the local correlations, that have been shown to be essential in the problem. Yet, there still is a room for improvement to dramatically increase the range of accessible types of statistical data.

   (b) For the eigenproblem of the low-energy transverse fluctuations of the order parameter we have also developed a numerical tool capable of restoring the profile of density states and discriminate between localized and extended parts of the spectrum, with both types of the states being confidently observed in the present problem. However, our current approach is rather limited in what concerns the accessible system sizes,

thus preventing us from providing a reliable quantitative description for the most interesting ranges of control parameters. In this regard, a substantial piece of technical work remains to be done, as the developed approach can be significantly improved.

3. In Part III we presented a detailed and practically complete theory for statistics of the order parameter

   (a) It was showed that there is a single function r of 2 variables that completely describes all statistical quantities in the theory, such as the full probability distribution of the order parameter, as well as any joint probability distribution of fields on any local group of sites, such as joint probability distribution on the 2 neighboring sites of the system.

   (b) The only assumption used in the derivation is the absence of short loops in the system, which is asymptotically true for large random regular graphs. Within this accepted assumption, all equations are exact in the limit of small Δ_0 and large K

   (c) The aforementioned function r is shown to satisfy a nonlinear integral equation. In the general case, this equation requires a complex procedure of obtaining a numerical solution that is yet to be developed.

   (d) However, in the regime of large v the system is trivially described by small fluctuations and thus assumes a proper analytical solution. In the opposite limit of small v we were able to simplify the integral equation on r and solve it analytically, thus obtaining the full description of the statistics of the order parameter for $v \lesssim g$.

   (e) The results of our calculations are in excellent agreement with the numerical solution, as judged by the full structure of the joint probability distribution of ξ and Δ fields on a given site, thus fully verifying our approach.

4. Finally, in Part IV we have presented an analytical approach to calculate the average correlator of low-energy transverse fluctuations of the order parameter

   (a) Our treatment is shown to be exact in a sense that all perturbation corrections vanish in the studied limit of small Δ_0 and large K. The only source of corrections remaining to be studied is the non-perturbative corrections from other functional saddle points.

   (b) The outcome of our approach is a single analytical equation on the value of the diagonal matrix elements of the average correlator.

   (c) The solution to this equation demonstrates good qualitative agreement with the numerical calculation. However, there exists a discrepancy due to the used approximation of small v that is poorly satisfied for all parameters accessible by the developed numerical method.

   (d) Once the apparent source of the aforementioned discrepancy is eliminated, the developed approach is expected to provide a valid quantitative description for the main body of the spectrum. However, this is yet to be verified. The question of whether the developed approach reproduces the detected localized states also needs to be challenged.

14 Physics conclusions

In this section we discuss some general conclusions following from our results so far, as well as a realistic interpretation of the observed physics:

1. One particularly important qualitative conclusion from our analysis is that the exact structure of the order parameter field is crucial to describe any physics on top of it. Ignoring the distribution of the order parameter might even lead to qualitatively incorrect results when one deals with low energy physics, as the order parameter is found to demonstrate a nontrivial distribution in the relevant region of parameters $v \approx g$.

2. The fact that our model is described only by 3 control parameters suggests that the proposed analytical approach can be extended to a more complicated model of disordered superconductors to include possible effects from nontrivial structure of localized eigenstates.

3. The demonstrated universal behavior of the low-energy physics also qualitatively justifies our claims about the irrelevance of some particular design choices for the initial model, as the obtained results are based on rather general principles that seemingly remain intact in a rather general setting.

4. Regarding our understanding of the low energy transverse fluctuations of the order parameter, we do not yet possess a confident quantitative description of the full phenomenology. To date, most of the general physics conclusions drawn from our analysis are of qualitative nature, both because of an overly simplified initial model and incomplete analysis of the model. The cumulative outcome of our research of transverse fluctuations is summarized in a qualitative phase diagram on Figure 23.

15 Further development

In this section we discuss potential directions of further theoretical and numerical research of the problem. We start by noting a list of purely technical tasks to be done in order to complete the analysis of the proposed theoretical model:

1. Among the top priorities is a technical implementation of the full numerical solution to the integral equation describing the universal r function. While we have already developed such a routine for relatively simple limit of small v, we are compelled to tackle the problem for arbitrary v in order to be able to conduct a fair comparison with numerical experiments.

2. Secondly, the developed approach begs for additional analytical investigation in order to infer the key qualitative features of the solution, such as the described position of the spectral edge in the full SCnBA approach considered.

3. Finally, we are interested in collecting a more reliable numerical dataset to improve both the statistical quality and the available range of parameters.

Secondly, we draw some long-term plans for future theoretical analysis:

1. One obvious aspiration is to study other saddle points of the functional integral for the average correlator. One can intuitively expect that these saddle points are important for describing the lowest energy scales.

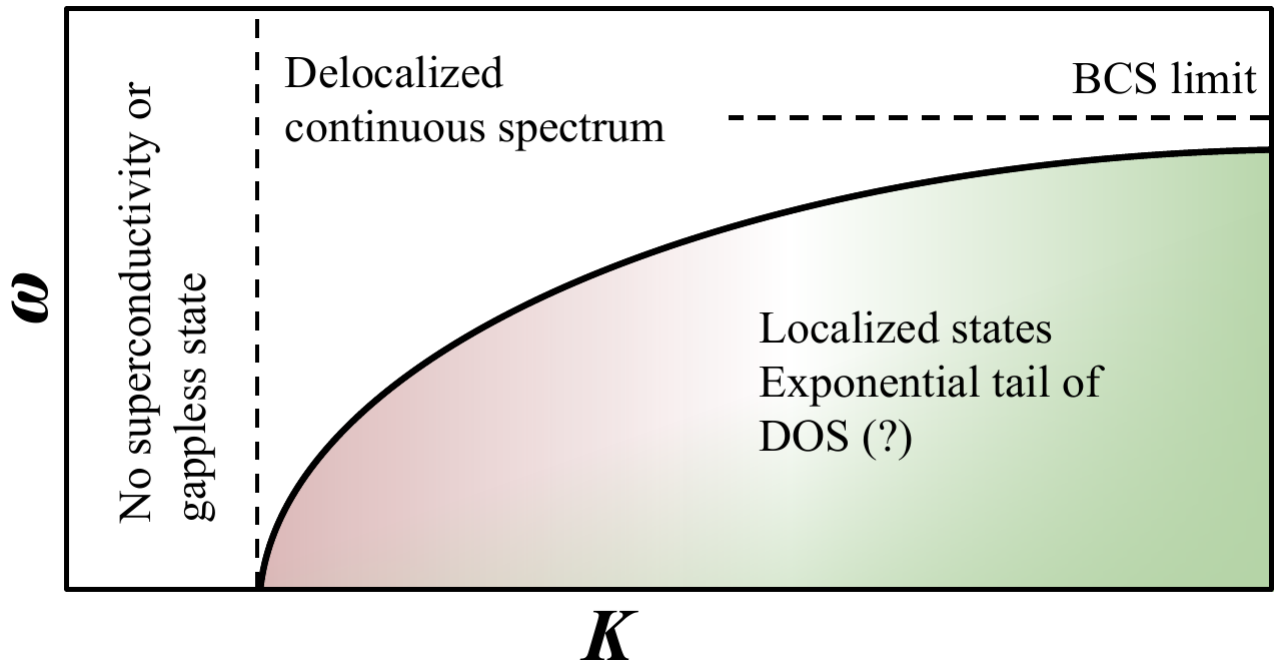


Figure 23: Sketch of an apparent phase diagram, as perceived from both analytical and numerical studies. Green color reflects the presence of localized low-energy eigenstates, while red color marks the region allegedly containing extended low-energy excitation, albeit with insufficient data available for that region to draw certain conclusions.

2. A more general observation is that we can notice all signs of existence of the underlying universal field theory that can correctly capture everything derived so far as well as yet-to-be-obtained results of the analysis of other saddle points of the model.

3. Finally, we are interested in exploiting the universality of the studied limit to possibly describe a more physical version of the current model, that would eventually allow one to apply our results to real world experiments.

References

[1] M. V. Feigel'man and L. B. Ioffe. Microwave Properties of Superconductors Close to the Superconductor-Insulator Transition. Physical Review Letters, 120(3), jan 2018.

[2] Pierre-Gilles De Gennes. Superconductivity of metals and alloys. CRC Press, 2018.

[3] Леонид Левитов and Андрей Шитов. Функции Грина. Задачи и решения. Litres, 2018.

[4] Bing Cheng, Liang Wu, N. J. Laurita, Harkirat Singh, Madhavi Chand, Pratap Raychaudhuri, and N. P. Armitage. Anomalous gap-edge dissipation in disordered superconductors on the brink of localization. Physical Review B, 93(18), may 2016.

[5] Daniel Sherman, Uwe S. Pracht, Boris Gorshunov, Shachaf Poran, John Jesudasan, Madhavi Chand, Pratap Raychaudhuri, Mason Swanson, Nandini Trivedi, Assa Auerbach, Marc Scheffler, Aviad Frydman, and Martin Dressel. The Higgs mode in disordered superconductors close to a quantum phase transition. Nature Physics, 11(2):188–192, jan 2015.

[6] Uwe S. Pracht, Nimrod Bachar, Lara Benfatto, Guy Deutscher, Eli Farber, Martin Dressel, and Marc Scheffler. Enhanced Cooper pairing versus suppressed phase coherence shaping the superconducting dome in coupled aluminum nanograins. Physical Review B, 93(10), mar 2016.

[7] M.V. Feigel'man, L.B. Ioffe, V.E. Kravtsov, and E. Cuevas. Fractal superconductivity near localization threshold. Annals of Physics, 325(7):1390–1478, jul 2010.

[8] Thomas Dubouchet, Benjamin Sacépé, Johanna Seidemann, Dan Shahar, Marc Sanquer, and Claude Chapelier. Collective energy gap of preformed cooper pairs in disordered superconductors. Nature Physics, 15(3):233–236, dec 2018.

[9] D. Shahar and Z. Ovadyahu. Superconductivity near the mobility edge. Physical Review B, 46(17):10917–10922, nov 1992.

[10] Benjamin Sacépé, Thomas Dubouchet, Claude Chapelier, Marc Sanquer, Maoz Ovadia, Dan Shahar, Mikhail Feigel'man, and Lev Ioffe. Localization of preformed cooper pairs in disordered superconductors. Nature Physics, 7(3):239–244, jan 2011.

[11] P. W. Anderson, K. A. Muttalib, and T. V. Ramakrishnan. Theory of the "universal" degradation ofTcin high-temperature superconductors. Physical Review B, 28(1):117–120, jul 1983.

[12] M. V. Feigel'man, L. B. Ioffe, and M. Mézard. Superconductor-insulator transition and energy localization. Physical Review B, 82(18), nov 2010.

[13] Michael Ma and Patrick A. Lee. Localized superconductors. Physical Review B, 32(9):5658–5667, nov 1985.

[14] Igor Poboiko and Mikhail Feigel'man. Paraconductivity of pseudogapped superconductors. Physical Review B, 97(1):014506, 2018.

[15] A. V. Shtyk and M. V. Feigel'man. Collective modes and ultrasonic attenuation in a pseudogapped superconductor. Physical Review B, 96(6), aug 2017.

[16] Анатолий Вадимович Свидзинский. Пространственно-неоднородные задачи теории сверхпроводимости. Наука. Гл. ред. физ.-мат. лит., 1982.

[17] Alexander Altland and Ben Simons. Condensed Matter Field Theory. Cambridge University Pr., 2010.

[18] V. N. Popov and S. A. Fedotov. The functional integration method and diagram technique for spin systems. Sov. Phys. JETP, 67(3):535–541, 1988.

[19] Антон Хвалюк и Михаил Фейгельман. Низкоэнергетические возбуждения в сверхпроводниках вблизи перехода "сверхпроводник-изолятор". Дипломная работа бакалавра, Московский физико-технический институт (ГУ), 2018.

[20] K. S. Tikhonov and A. D. Mirlin. Fractality of wave functions on a Cayley tree: Difference between tree and locally treelike graph without boundary. Physical Review B, 94(18), nov 2016.

[21] Brendan D McKay. The expected eigenvalue distribution of a large regular graph. Linear Algebra and its Applications, 40:203–216, 1981.

[22] Fan Chung and S.-T. Yau. Coverings, heat kernels and spanning trees. The Electronic Journal of Combinatorics, 6(1), dec 1998.

[23] Brendan D. McKay and Nicholas C. Wormald. Uniform generation of random regular graphs of moderate degree. Journal of Algorithms, 11(1):52–67, mar 1990.

[24] Eric Polizzi. Density-matrix-based algorithm for solving eigenvalue problems. Physical Review B, 79(11), mar 2009.

[25] R. Abou-Chacra, D. J. Thouless, and P. W. Anderson. A selfconsistent theory of localization. Journal of Physics C: Solid State Physics, 6(10):1734–1752, may 1973.

[26] Alexander D. Mirlin and Yan V. Fyodorov. Localization transition in the anderson model on the bethe lattice: Spontaneous symmetry breaking and correlation functions. Nuclear Physics B, 366(3):507–532, dec 1991.

[27] A A Abrikosov. MAGNETIC IMPURITIES IN NONMAGNETIC METALS. Soviet Physics Uspekhi, 12(2):168–181, feb 1969.

[28] AI Larkin and Yu N Ovchinnikov. Density of states in inhomogeneous superconductors. Sov. Phys. JETP, 34:1144–1150, 1972.