Minimal vertex covers on finite-connectivity random graphs – a hard-sphere lattice-gas picture

Martin Weigt and Alexander K. Hartmann
Institute for Theoretical Physics,
University of Göttingen, Bunsenstr. 9, 37073 Göttingen, Germany
E-mail: hartmann/weigt@theorie.physik.uni-goettingen.de

December 6, 2000

Abstract

The minimal vertex-cover (or maximal independent-set) problem is studied on random graphs of finite connectivity. Analytical results are obtained by a mapping to a lattice gas of hard spheres of (chemical) radius one, and they are found to be in excellent agreement with numerical simulations. We give a detailed description of the replica-symmetric phase, including the size and the entropy of the minimal vertex covers, and the structure of the unfrozen component which is found to percolate at connectivity $c \simeq 1.43$. The replica-symmetric solution breaks down at $c = e \simeq 2.72$. We give a simple one-step replica symmetry broken solution, and discuss the problems in interpretation and generalization of this solution.

1 Introduction

The last few years have seen an increasing interest of theoretical computer scientists, mathematicians and, more recently, of statistical physicists in random combinatorial optimization and decision problems, see *e.g.* [1, 2]. Traditional complexity theory [3] characterizes combinatorial problems with respect to the worst-case dependence of solution times for algorithms on the problem size, or, more precisely, on the memory size needed to encode a problem. Some of the most challenging problems are collected in the class of *NP-complete problems*: In such problems a potential solution can be verified (or falsified) very effectively in polynomial time, whereas the search for a solution among the exponential number of candidates becomes very slow due to entropic reasons. The completeness property refers to the fact that once an effective, *i.e.* polynomial algorithm is found for any NP-complete problem, it can be modified to effectively solve every other such problem. The question whether or not such algorithms can be constructed is however still open, and belongs to the important open questions of modern mathematics. Famous members of the class of NP-complete problems are *e.g.* Boolean satisfiability, number partitioning, vertex cover, or the traveling-salesman problem.

This worst-case classification does however give no information on typical solution times. For almost ten years now, randomized optimization and decision problems have been studied therefore, for an overview see the special issues [1, 2]. It was realized, that exponentially longest solution times typically appear when the problems are situated at phase boundaries and therefore critically constraint [4].

Due to the analogy between such combinatorial optimization problems and statistical-mechanics models with discrete degrees of freedom at low temperature, many methods developed in physics can be applied directly to theoretical computer science. This was done *e.g.* for Boolean satisfiability [5, 6, 7, 8], for number partitioning [9], the traveling-salesman problem [10], for Euclidean matching [11] and recently also for vertex cover [12]. Also the relations between phase transitions and the appearance of hardest instances was recently analyzed for specific algorithms using statistical-mechanics methods [13, 14].

In this paper, we give a detailed description of the statistical mechanics approach to minimal vertex covers on finite connectivity random graphs. For this reason, the model will be mapped to a random lattice gas of hard spheres of radius one.

The plan of the paper is the following. In the next section we define the model and give an overview over some rigorously known results. In section 3 the model is shown to be equivalent to a hard-sphere lattice gas. Section 4 explains the numerical methods used to check the analytical results. The latter are based on the replica approach presented in sections 5-7, starting with a general calculation of the replicated free energy. In section 6, the most important results are presented: the size, entropy and structure of minimal vertex covers are described in a replica-symmetric approach, whereas the simplest one-step replica symmetry broken solution is explained in section 7. The paper is closed by a concluding section. Several technical details are delegated to three appendices.

2 The model

In this section, we will introduce the terminology and some rigorously known results about vertex cover and related problems.

2.1 Vertex cover and related problems

Let us start with the definition of vertex covers. We consider a graph $G = (V, E)$ with N vertices $i \in \{1, 2, ..., N\}$ and undirected edges $\{i, j\} \in E \subset V \times V$ connecting pairs of vertices. Please note that $\{i, j\}$ and $\{j, i\}$ both denote the same edge.

Definition 1: *A vertex cover V_{vc} is a subset $V_{vc} \subset V$ of vertices such that for all edges $\{i, j\} \in E$ at least one of the endpoints is in V_{vc}, i.e. $i \in V_{vc}$ or $j \in V_{vc}$.*

Later on also subsets V_{vc} are considered, which are not covers. Anyway, we call all vertices in V_{vc} *covered*, all others *uncovered*. Also edges from $E \cap (V_{vc} \times V \cup V \times V_{vc})$ are called covered. This means that if V_{vc} is a vertex cover, all edges are covered.

We will study the minimal vertex-cover problem, which consists in finding a vertex cover V_{vc} of minimal cardinality, and calculate the minimal fraction $x_c(G) = |V_{vc}|/N$ needed to cover the whole graph.

This problem is equivalent to other optimization problems:

- An *independent set* is a subset of vertices which are pairwise disconnected in the graph G. Due to the above-mentioned properties, any set $V \setminus V_{vc}$ thus forms an independent set, and maximal independent sets are complementary to minimal vertex covers.
- A *clique* is a fully connected subset of vertices, and thus an independent set in the complementary graph $\overline{G}$ where vertices i and j are connected whenever $\{i, j\} \notin E$ and vice versa.

2.2 Random graphs

In order to speak of median or average cases, and of phase transitions, we have to introduce a probability distribution over graphs. This can be done best by using the concept of *random graphs* as already introduced about 40 years ago by Erdös and Rényi [15]. A random graph $G_{N,p}$ is a graph with N vertices $V = \{1, ..., N\}$, any pair of vertices is connected randomly and independently by an edge with probability p. So the expected number of edges becomes $p\binom{N}{2} = pN^2/2 + O(N)$, and the average connectivity of a vertex equals $p(N-1)$.

The regime we are interested in are however *finite-connectivity graphs* having $p = c/N$ with constant c in the large-N limit. Then the average connectivity $c+O(N^{-1})$ stays finite. In this case, we also expect the size of minimal vertex covers to depend only on c, $x_c(G) = x_c(c)$ for almost all random graphs $G_{N,c/N}$.

Here we want to review shortly some of the fundamental results on random graphs which were already described in [15], and which are important for the following sections:

The first point we want to mention is the distribution of connectivities (or vertex degrees) d, in the limit $N \to \infty$ it is given by a Poisson-distribution with mean c:

$$Po_c(d) = e^{-c}\frac{(c)^d}{d!} \ . \tag{1}$$

A second point which is important for the understanding of the following is the component structure. For $c < 1$, *i.e.* if the vertices have in average less than one neighbor, the graph $G_{N,c/N}$ is built up from connected components containing up to $O(\ln N)$ vertices. The probability that a component is a specific tree T_k of k vertices is given by

$$\rho(k) = e^{-ck}\frac{(c)^{k-1}}{k!} \ , \tag{2}$$

and is equal for all k^{k-2} distinct trees. As the fraction of vertices which are collected in finite trees is $\sum_{k=1}^{\infty} \rho(k) k^{k-2} k = 1$ for all $c < 1$, in this case almost all vertices are collected in such trees. For $c > 1$ a giant component appears which contains a finite fraction of all vertices. $c = 1$ is therefore called the *percolation threshold.*

For a complete introduction to random graphs see the book by Bollobas [16].

2.3 Rigorously known bounds

In this subsection we are going to present some previously known rigorous bounds on $x_c(c)$. A general one for arbitrary, *i.e.* non-random graphs G was given by Harant [17] who generalized an old result of Caro and Wei [18]. Translated into our notation, he showed that

$$x_c(G) \leq 1 - \frac{1}{N}\frac{\left(\sum_{i\in V}\frac{1}{d_i+1}\right)^2}{\sum_{i\in V}\frac{1}{d_i+1} - \sum_{(i,j)\in E}\frac{(d_i-d_j)^2}{(d_i+1)(d_j+1)}} \tag{3}$$

where d_i is the connectivity (or degree) of vertex i. Using the distribution (1) of connectivities and its generalization to pairs of connected vertices, this can easily be converted into an upper bound on $x_c(c)$ which holds almost surely for $N \to \infty$.

The vertex cover problem or the above-mentioned related problems were also studied in the case of random graphs, and even completely solved in the case of infinite connectivity graphs, where any edge is drawn with finite probability p, such that the expected number of edges is $p\binom{N}{2} = 0(N^2)$. There the minimal VC has cardinality $(N - 2\ln_{1/(1-p)} N - O(\ln\ln N))$

[19]. Bounds in the finite-connectivity region of random graphs with N vertices and cN edges were given by Gazmuri [20]. He showed that

$$x_l(c) < x_c(c) < 1 - \frac{\ln c}{c} \tag{4}$$

where the lower bound is given by the unique solution of

$$0 = x_l(c) \ln x_l(c) + (1 - x_l(c)) \ln(1 - x_l(c)) - \frac{c}{2}(1 - x_l(c))^2 \ . \tag{5}$$

This bound coincides with the so-called annealed bound in statistical physics. The correct asymptotics for large c was given by Frieze [21]:

$$x_c(c) = 1 - \frac{2}{c}(\ln c - \ln\ln c + 1 - \ln 2) + o(\frac{1}{c}) \tag{6}$$

with corrections of $o(1/c)$ decaying faster than $1/c$.

3 Equivalence to a hard-sphere lattice gas

After having introduced the problem in mathematical terms, we are now going to connect it to a statistical-mechanics model, more precisely to a lattice gas of hard spheres of chemical radius one.

Any subset $U \subset V$ of the vertex set can be encoded by a configuration of N binary variables:

$$x_i := \begin{cases} 0 & \text{if} \quad i \in U \\ 1 & \text{if} \quad i \notin U \end{cases} \tag{7}$$

The strange choice of setting x_i to zero for vertices in U becomes clear if we look at the vertex cover constraint: An edge is covered by the elements in U iff at most one of the two end-points has $x = 1$. So the variables x_i can be interpreted as occupation numbers of vertices by the center of a particle. The covering constraint translates into a hard sphere constraint: If a vertex is occupied, *i.e.* $x_i = 1$, then all neighboring vertices have to be empty. We thus introduce an indicator function

$$\chi(x_1, ..., x_N) = \prod_{\{i,j\} \in E} (1 - x_i x_j) \tag{8}$$

which is one whenever $\vec{x} = (x_1, ..., x_N)$ corresponds to a vertex cover, and zero else. Having in mind this interpretation, we may write down the grand partition function

$$\Xi = \sum_{\{x_i = 0,1\}} \exp\left(\mu \sum_i x_i\right) \chi(\vec{x}) \tag{9}$$

with μ being a chemical potential which can be used to control the particle number, or the cardinality of U.

For regular lattices, this model is well studied as a lattice model for the fluid-solid transition, for an overview and the famous corner-transfer matrix solution of the two-dimensional hard-hexagon model by R. Baxter see his book [22]. Recently, lattice-gas models with various kinds of disorder have been considered in connection to glasses [23, 24, 25] and granular matter [26, 27, 28, 29, 30, 31].

Denoting the grand canonical average as

$$\langle f(\vec{x})\rangle_\mu = \Xi^{-1} \sum_{\{x_i=0,1\}} \exp\left(\mu \sum_i x_i\right) \ \chi(\vec{x}) \ f(\vec{x}) \tag{10}$$

we can calculate the average occupation density

$$\nu(\mu) = \frac{1}{N}\langle \sum_i x_i \rangle_\mu = \frac{\partial}{\partial\mu}\frac{\ln \Xi}{N} \ , \tag{11}$$

and the corresponding entropy density is given by a Legendre transform of $\ln \Xi$,

$$s(\nu(\mu)) = \left(1 - \mu\frac{\partial}{\partial\mu}\right)\frac{\ln \Xi}{N}, \tag{12}$$

where the thermodynamic limit $N \to \infty$ is implicitly assumed. The entropy of vertex covers of cardinality xN thus reads

$$s_{vc}(x) = s(1-x) \tag{13}$$

Minimal vertex covers correspond to densest particle packings. Considering the weights in (9), it becomes obvious that the density $\rho(\mu)$ is an increasing function of the chemical potential. Densest packings, or minimal vertex covers, are thus obtained in the limit $\mu \to \infty$:

$$x_c(c) = 1 - \lim_{\mu\to\infty} \nu(\mu) \ . \tag{14}$$

4 Numerical methods

Before explicitly following this strategy in the special case of random graphs, we are going to present our numerical methods. So we can later-on directly compare all analytical results to numerical data.

All numerical results were obtained by exact enumerations. For large average connectivities $c \geq 4$ a branch-and-bound algorithm was applied, while for small average connectivities a divide-and-conquer technique is more appropriate. Since some readers may not be familiar with combinatorial optimization algorithms, the methods are explained in detail. Before presenting the two procedures, we first introduce a fast heuristic, which is used within both methods. The heuristic can applied stand-alone as well. In this case only an approximation of the true minimum vertex cover is calculated, which is found to differ only by a few percent from the exact value. All methods have been implemented via the help of the LEDA library [32] which offers many useful data types and algorithms for linear algebra and graph problems.

The basic idea of the heuristic is to cover as many edges as possible by using as few vertices as necessary. Thus, it seems favorable to cover vertices with a high degree. This step can be iterated, while the degree of the vertices is adjusted dynamically by removing edges and vertices which are covered. This leads to the following algorithm, which returns an approximation of the minimum vertex cover V_{vc}, the size $|V_{vc}|$ is an upper bound of the true minimum vertex-cover size:

algorithm min-cover(G)
begin
  initialize $V_{vc} = \emptyset$;
  while there are uncovered edges **do**
  begin

take one vertex i with the largest current degree d_i;
mark i as covered: $V_{vc} = V_{vc} \cup \{i\}$;
remove all incident edges $\{i, j\}$ from E;
remove vertex i from V;
end;
return(V_{vc});
end

In Fig. 1 a simple example is presented, where the heuristic fails to find the true minimal vertex cover. First the algorithm covers the root vertex of degree 3. Thus, additionally 3 vertices have to be subsequently covered, i.e. the heuristic covers 4 vertices. But, the minimum vertex cover has only size 3, as indicated in Fig. 1.

So far we have presented a simple heuristic to find approximations of minimum vertex covers, which will be part of the exact algorithms, which we have been applied to obtain all numerical results presented in this work. Next, two exact algorithms are explained: divide-and-conquer and branch-and-bound.

The basic idea of both methods is as follows: as each vertex is either covered or uncovered, there are 2^N possible configurations which can be arranged as leafs of a binary (backtracking) tree. At each node, the two subtrees represent the subproblems where the corresponding vertex is either *covered* ("left subtree") or *uncovered* ("right subtree"). Vertices, which have not been touched at a certain level of the tree are said to be *free*. Both algorithms do not descent further into the tree, when a cover has been found, i.e. when all edges are covered. Then the search continues in higher levels of the tree (backtracking) for a cover which has possibly a smaller size. Since the number of nodes in a tree grows exponentially with system size, algorithms which are based on backtracking trees have a running time which may grow exponentially with the system size. This is not surprising, since the minimal-VC problem is NP-hard, so all exact methods exhibit an exponential growing worst-case time complexity.

To decrease running time, both algorithms make use of the fact, that only full vertex covers are to be obtained. Therefore, when a vertex i is marked *uncovered*, all neighboring vertices can be *covered* immediately. Concerning these vertices, only the left subtrees are present in the search tree.

The divide-and-conquer [33] approach is based on the fact that a minimum VC of a graph, which consists of several independent connected components, can be obtained by combining the minimum covers of the components. Thus, the full task can be split into several independent tasks. This strategy can be repeated at all levels of the backtracking tree. At each level, the edges which have been covered can be removed from the graph, so the graph may split into further components. As a consequence, below the percolation threshold, where the size of the largest components is of the order $O(\ln N)$, the algorithm exhibits a polynomial running time. Summarizing, the divide-and-conquer approach reads as follows, the given subroutine is called for each component of the graph separately, it returns the size of the minimum vertex cover. Initially all vertices have state *free*:

algorithm divide_and_conquer(G)
begin
take one *free* vertex i with the largest current degree d_i;

mark i as *covered*; **comment** left subtree
$size_1 := 1$;
remove all incident edges $\{i, j\}$ from E;
calculate all connected components $\{C_i\}$ of graph built by *free* vertices;

```
    for all components C_i do
        size_1 := size_1+ divide_and_conquer(C_i);
    insert all edges {i, j} which have been removed;

    mark i as uncovered; comment right subtree;
    size_2 := 0;
    for all neighbors j of i do
    begin
        mark j as covered
        remove all incident edges {j, k} from E;
    end
    calculate all connected components {C_i};
    for all components C_i do
     size_2 := size_2+ divide_and_conquer(C_i);
    for all neighbors j of i do
        mark j as free
    insert all edges {j, k} which have been removed;

    mark i as free;
    if size_1 < size_2 then
        return(size_1);
    else
        return(size_2);
end
```

The algorithm can be easily extended to record the cover sets as well or to calculate the degeneracy. In Fig. 2 an example of the operation is given. The algorithm is able to treat large graphs deep in the percolating regime. We have calculated for example minimum vertex covers for graphs of size $N = 560$ with average connectivity $c = 1.3$

For average connectivities larger than 4, the divide-and-conquer algorithm is too slow, because the graph only rarely splits into several components. Then a branch-and-bound approach [34, 35, 36] is favorable. It differs from the previous method by the fact that no independent components of the graph are calculated. Instead, some subtrees of the back-tracking tree are omitted by introducing a *bound:* This is achieved by storing always the size *best* of the smallest vertex cover found so far (initially N) and recording the number X of vertices which have bee *covered* in higher levels of the tree. Additionally, always a table of *free* vertices ordered in descending current degree d_i is kept. Thus, to achieve a better solution, at most $F = best - X$ vertices can be *covered*. This means, it is not possible to cover more edges, than given by the sum $D = \sum_{l=1}^{F} d_l$ of the F highest degrees in the table of vertices, *i.e.* if some edges remain uncovered, the corresponding subtree can be omitted for sure. Please note that in the case that some edges are running between the F vertices of highest current degree, then a subtree may be entered, even if it contained no smaller cover.

The algorithm can be summarized as follows below. The size of the smallest covered is stored in *best*, which is passed by reference (i.e. the variable, not its value is passed). The number of covered vertices is stored in variable X, please remember $G = (V, E)$:

```
algorithm branch_and_bound(G, best, X)
begin
    if all edges are covered then
    begin
```

```
        if X < best then
            best := X;
        return;
    end;
    calculate F = best − X; D = ∑_{l_1}^F d_l;
    if D < number of uncovered edges then
        return;          comment bound;
    take one free vertex i with the largest current degree d_i;

    mark i as covered; comment left subtree
    X := X + 1;
    remove all incident edges {i, j} from E;
    branch_and_bound(G, best, X);
    insert all edges {i, j} which have been removed;
    X := X − 1;

    if (X > number of current neighbors) then
    begin                comment right subtree;
        mark i as uncovered;
        for all neighbors j of i do
        begin
            mark j as covered;
            X := X + 1;
            remove all incident edges {j, k} from E;
        end;
        branch_and_bound(G, best, X);
        for all neighbors j of i do
        begin
            mark j as free;
            X := X − 1;
        end;
        insert all edges {j, k} which have been removed;
    end;

    mark i as free;
    return;
end
```

For every calculation of the bound one has to access the F vertices of largest current connectivity. Thus it it favorable to implement the table of vertices as two arrays v_1, v_2 of sets of vertices. The arrays are indexed of the current degree of the vertices. The sets of the first array v_1 contains the F *free* vertices of largest current degree, while the other array contains all other *free* vertices. Every time a vertex changes its degree, it is moved to another set, and eventually even changes the array. Also, in case the mark of a vertex changes it may be entered in or removed from an array, possibly the smallest degree vertex of v_1 is moved to v_2 or vice versa. Since we are treating graphs of finite average connectivity, this implementation ensures that the running time spent in each node of the graph is almost constant. For the sake of clarity, we have omitted the update operation for both arrays from the algorithmic presentation.

Although our algorithm is very simple, in the regime $4 < c < 10$ random graphs up to size $N = 140$ could be treated. It is difficult to compare the branch-and-bound algorithm to state-of-the-art algorithms [36, 37] because they are usually tested on a different graph ensemble where each edge appears with a certain probability, independently of the graph size (high-connectivity regime). Nevertheless, in the literature usually graphs with up to 200 vertices are treated, which is slightly larger than the systems considered here. But our algorithm has the advantage that it is easy to implement. Furthermore, it can be easily modified to study more general questions, see [12]

5 Replica approach

After having introduced our numerical methods, we go back to the statistical-mechanics approach displayed in section 3. The main problem in handling the grand partition function (9) is caused by the disorder due to the random structure of the underlying graph, *i.e.* of the edge set E. To calculate typical properties we therefore have to evaluate the disorder average of $\ln \Xi$ over the random graph ensemble. This can be achieved by the replica trick [38],

$$\overline{\ln \Xi} = \lim_{n\to 0} \frac{\overline{\Xi^n} - 1}{n} \tag{15}$$

where the over-bar denotes the disorder average over the random-graph ensemble. Taking n to be a positive integer at the beginning, we may replace the original system by n identical copies (including identical edge sets). In this case, the disorder average is easily obtained, and the $n \to 0$ limit has to be achieved later by analytically continuing in n. We may thus write, with n being a natural number,

$$\overline{\Xi^n} = \sum_{\{x_i^a\}} \exp\left(\mu \sum_{i,a} x_i^a\right) \overline{\prod_{\{i,j\}\in E} \prod_{a=1}^{n} (1 - x_i^a x_j^a)} \tag{16}$$

with a denoting the replica index which runs from 1 to n. Putting edges independently with probability c/N results in

$$\begin{aligned}
\overline{\Xi^n} &= \sum_{\{x_i^a\}} \exp\left(\mu \sum_{i,a} x_i^a\right) \prod_{i<j} \left[1 - \frac{c}{N} + \frac{c}{N} \prod_a (1 - x_i^a x_j^a)\right] \\
&= \sum_{\{x_i^a\}} \exp\left(\mu \sum_{i,a} x_i^a - \frac{cN}{2} + \frac{c}{2N} \sum_{i,j} \prod_a (1 - x_i^a x_j^a) + O(1)\right) .
\end{aligned} \tag{17}$$

Following the ideas of [39], we introduce 2^n order parameters

$$c(\vec{\xi}) = \frac{1}{N} \sum_i \prod_a \delta_{\xi^a, x_i^a} \tag{18}$$

as the fraction of vertices showing the replicated variable $\vec{\xi} \in \{0,1\}^n$. The exponent in the last line of (17) obviously depends only on this quantity. Using Stirlings formula for the number $N!/\prod_{\vec{\xi}}(c(\vec{\xi})N)!$ of configurations of the $\{x_i^a\}$ having the same $c(\vec{\xi})$, we find

$$\overline{\Xi^n} = \int \mathcal{D}c(\vec{\xi}) \exp\left\{ N \left(-\sum_{\vec{\xi}} c(\vec{\xi}) \ln c(\vec{\xi}) - \frac{c}{2} + \mu \sum_{\vec{\xi},a} c(\vec{\xi}) \xi^a + \frac{c}{2} \sum_{\vec{\xi},\vec{\zeta}} c(\vec{\xi}) c(\vec{\zeta}) \prod_a (1 - \xi^a \zeta^a) \right) \right\} \tag{19}$$

where the integration is over all normalized distributions $c(\vec{\xi})$, *i.e.* $\sum_{\vec{\xi}} c(\vec{\xi}) = 1$. In the large-$N$ limit, the integration can be solved by the saddle-point method. The saddle-point equation can be obtained by variation of the exponent in (19) with respect to all allowed $c(\vec{\xi})$:

$$c(\vec{\xi}) = \exp\left\{-\lambda + \mu \sum_a \xi^a + c \sum_{\vec{\zeta}} c(\vec{\zeta}) \prod_a (1 - \xi^a \zeta^a)\right\} . \tag{20}$$

λ is a Lagrange multiplier introduced in order to guarantee the normalization of $c(\vec{\xi})$. For $n \to 0$, it will tend to the connectivity c. Before we can however calculate this limit, we have to introduce some ansatz for $c(\vec{\xi})$ as even the dimensionality of $c(\xi)$ still depends on n. In the next section, we are going to use the simplest-possible, *i.e.* the replica-symmetric ansatz. As this ansatz is found to be valid only for a finite connectivity range, we also include one step of replica-symmetry breaking in the over-next section.

6 The replica-symmetric solution

6.1 The replica-symmetric ansatz

As already explained, we are now using the so-called replica-symmetric ansatz, which in our case assumes that the order parameter $c(\vec{\xi})$ depends on $\vec{\xi}$ only via $\sum_a \xi_a$, cf. also [39, 12]. In this case we are able to write

$$c(\vec{\xi}) = \int dh\ P(h) \frac{\exp(h \sum_a \xi_a)}{(1 + e^h)^n} \tag{21}$$

with $P(h)$ being a probability distribution to guarantee the normalization of $c(\vec{\xi})$. The physical interpretation of $P(h)$ is simple: Take any vertex i, then its average local occupation number $\langle x_i \rangle_\mu$ in the presence of the chemical potential μ can be written as $e^{h_i}/(1 + e^{h_i})$ using an effective chemical potential h_i accounting for all interactions on i. $P(h)$ can now be constructed as the histogram of these effective chemical potentials.

Plugging this ansatz into equations (19) and (20), the replica number n appears as a mere parameter, and the limit $n \to 0$ can be calculated. Details of this calculation are given in appendix A. The saddle-point equation (20) now reads

$$\int dh\ P(h) e^{hy} = \exp\left\{-c + \mu y + c \int dh\ P(h)(1 + e^h)^{-y}\right\} \tag{22}$$

and has to be valid for arbitrary y. According to equations (12,13) we find the entropy density of vertex covers using a fraction $x = 1 - \int dh\ P(h)/(1 + e^{-h})$ of vertices

$$\begin{aligned} s_{VC}(x) &= \int \frac{dh\ dk}{2\pi} e^{ikh} P_{FT}(k)[1 - \ln P_{FT}(k)]\ \ln(1 + e^h) \\ &\quad + \frac{c}{2} \int dh_1\ dh_2 P(h_1) P(h_2) \ln\left(1 - \frac{1}{(1 + e^{-h_1})(1 + e^{-h_2})}\right) \end{aligned} \tag{23}$$

where

$$P_{FT}(k) := \int dh\ e^{-ikh}\ P(h) \tag{24}$$

denotes the Fourier transform of $P(h)$.

6.2 Size of minimal vertex covers

It is however to complicated to directly solve (22) for arbitrary μ. But we are interested in the properties of minimal vertex covers, which, according to section 3 can be described by the limit $\mu \to \infty$ of infinitely large chemical potential. In this case, we also expect the effective chemical potentials h to scale as $z\mu$, with z being a random variable with finite mean and variance. The rescaled probability distribution is denoted by $\tilde{P}(z)$. Please note that a negative z now corresponds to vertices having always $x_i = 0$, whereas positive z indicate vertices with fixed $x_i = 1$. All vertices being occupied in some ground-states and empty in others are collected in $z = 0$. This picture has to be refined for the calculation of the vertex-cover entropy in section 6.4: There also contributions of $O(\mu^0)$ have to be taken into account. For the present purpose the dominant terms are however sufficient. To obtain a well-defined limit $\mu \to \infty$ of eq. (22), we also have to rescale y by $k := \mu y$. Thus eq. (22) becomes

$$\int dz\ \tilde{P}(z) e^{zk} = \exp\left\{ -c + z + c\left(\int_{-\infty}^{+0} dz\ \tilde{P}(z) + \int_{+0}^{\infty} dz\ \tilde{P}(z) e^{-zk} \right) \right\} \ . \tag{25}$$

The interpretation of this equation in terms of the cavity approach, see [38], becomes evident if we Fourier-transform it and develop the last part of the exponential on the right-hand side,

$$\begin{aligned} \tilde{P}(z) &= \sum_{d=0}^{\infty} e^{-c} \frac{c^d}{d!} \left[\delta(\cdot + 1) * \tilde{P}_-^{*d}(\cdot) \right](z) \ , \\ \tilde{P}_-(z) &= \delta(z) \int_{-\infty}^{+0} dz\ \tilde{P}(z) + \Theta(-z) \tilde{P}(-z) \ . \end{aligned} \tag{26}$$

$*$ denotes the convolution product. This equation describes the effective chemical potentials of a vertex which is the linear superposition of the exterior chemical potential $1 \cdot \mu$ and the contributions of all neighbors. The contribution $\tilde{P}_-(z)$ of a neighbor depends on the effective potentials the neighbors would have without the presence of the central vertex, for details on this cavity interpretation see [38], and reflects the hard-sphere condition. A neighbor with positive potential would be occupied, $x_i = 1$, and thus forces a negative field for the central term. Neighbors with non-negative chemical potential do not impose anything as they would have $x_i = 0$. At the end, the resulting distribution is averaged over the connectivity distribution of random graphs.

This saddle-point equation has the simple solution

$$\tilde{P}(z) = \sum_{l=-1}^{\infty} \frac{W(c)^{l+2}}{(l+1)!\ c} \delta(z + l) \tag{27}$$

where the Lambert-W function $W(c)$ is defined as the real solution of

$$c = W(c)\ e^{W(c)} \ . \tag{28}$$

We already mentioned that vertices having negative fields are frozen to $x_i = 0$, vertices with positive fields to $x_i = 1$. At the present level, we are however not able to calculate the average magnetization of the vertices belonging to $z = 0$. This will be done in the next section. Here we only use the result: Half of $(z = 0)$-vertices are occupied, half are empty. We therefore find an average occupation density of hard spheres,

$$\nu(\mu \to \infty) = \frac{1}{N} \langle \sum_i x_i \rangle_{\mu \to \infty} = \frac{2W(c) + W(c)^2}{2c} \ , \tag{29}$$

which translates to a minimal vertex-cover size given by

$$x_c(c) = 1 - \frac{2W(c) + W(c)^2}{2c} \ . \tag{30}$$

In figure 3, this result is compared to numerical simulations. Extremely good coincidence is found for small connectivities c. For non-percolating graphs, *i.e.* for $c < 1$, our result was recently proven to be exact [40].

Systematic deviations show up later. For large c, eq. (30) even violates the bounds given in section 2.3 and the exactly known asymptotics (6). As we will see later, this can be explained within our approach: Replica symmetry breaks down at $c = e \simeq 2.718$, see the following sections. Up to this value however, we expect the replica-symmetric result to be exact. This is astonishing, as the solution does not show any particular signature of the percolation transition of the underlying random graph at $c = 1$.

6.3 The backbone

The distribution $\tilde{P}(z)$ contains much more statistical information on minimal vertex covers than simply its size. One important effect is a partial freezing which can be observed: There are vertices, which are always uncovered ($x_i = 1$) in all minimal vertex covers, others are always covered ($x_i = 0$). We call these spins uncovered (resp. covered) *backbone*. The fractions of vertices belonging to these two backbone types are given by

$$\begin{aligned} b_{uncov}(c) &= \frac{W(c)}{c} \\ b_{cov}(c) &= 1 - \frac{W(c) + W(c)^2}{c} \ . \end{aligned} \tag{31}$$

The remaining $W(c)^2 N/c$ vertices are unfrozen, their covering state changes from ground state to ground state. These different freezing properties can already be seen in simple finite graphs: A graph consisting of two connected vertices has two minimal vertex covers, and the state of the two vertices is not uniquely determined. They thus do not belong to the backbone. The situation is different for graphs with three vertices and two edges. The central vertex is covered in the unique minimal vertex cover, thus belonging to the covered backbone. The other two vertices form the negative backbone.

Let us now investigate the influence of the close environment of a vertex on its behavior, more precisely the influence of its connectivity. The total connectivity distribution is given by the Poisson law (1), but we can distinguish three distinct contributions:

- The joint probability $P(d, \langle x \rangle = 1)$ that a vertex has connectivity d and belongs to the uncovered backbone of all minimal vertex covers.
- $P(d, \langle x \rangle = 0)$ gives the probability that a vertex has connectivity d and belongs to the covered backbone of all minimal vertex covers.
- The remaining vertices are not in the backbone, thus being described by $P(d, 0 < \langle x \rangle < 1)$.

These quantities can be easily computed from $\tilde{P}(z)$: according to the interpretation of the self-consistent equation (26) we can calculate the effective-field distribution for a vertex of connectivity d which, in average, has typical neighbors:

$$\tilde{P}_d(z) = \left[\delta(\cdot + 1) * \tilde{P_-}^{*d}(\cdot)\right](z) \tag{32}$$

where $\tilde{P}_-(z)$ is exactly the quantity given in (26). Plugging our solution (27) into this equation, we find

$$\begin{aligned} P(d, \langle x\rangle = 1) = \tilde{P}_d(z>0) Po_c(d) &= e^{-c} \frac{[c - W(c)]^d}{d!} \\ P(d, 0 < \langle x\rangle < 1) = \tilde{P}_d(z=0) Po_c(d) &= e^{-c} \frac{W(c)[c - W(c)]^{d-1}}{(d-1)!} \\ P(d, \langle x\rangle = 0) = \tilde{P}_d(z<0) Po_c(d) &= e^{-c} \frac{[c + (d-1)W(c)][c - W(c)]^{d-1}}{(d-1)!} \end{aligned} \tag{33}$$

The results for $c = 2$ are displayed in Fig. 5 along with numerical data for $N = 17, 35, 70$. Please note that the numerical results seem to converge towards the analytical one, thus showing an excellent coincidence of both approaches. The curves are easily understood: A vertex of connectivity 0 has no neighbors. Therefore, it does not appear in any optimum cover and we obtain $P(0, \langle x\rangle = 1) = Po_c(0)$, $P(0, \langle x\rangle < 1) = 0$. With increasing connectivity the probability that a vertex is covered increases, thus the contribution of $P(d, 0 < \langle x\rangle \le 1)$ to $\mathrm{Po}_c(d)$ increases as well. For large connectivities it is very probable that a vertex belongs to all VCs but even a finite fraction of vertices with $\langle x\rangle = 1$ remains. These results justify *a posteriori* the application of a greedy heuristic within the algorithm: Vertices having large connectivity are at first included into the cover set.

6.4 Entropy of minimal vertex covers

In order to calculate the entropy of minimal vertex covers, we have to go beyond the leading terms in the effective chemical potentials. If we consider *e.g.* the non-backbone spins, the order μ of the effective fields is vanishing, but the order μ^0 determines the actual average occupation. We thus have to decompose the effective potentials according to $h = \mu z + \tilde{z}$, and write the order parameter as

$$c(\vec{\xi}) = \int dz\; d\tilde{z}\; P(z, \tilde{z}) \frac{\exp((\mu z + \tilde{z}) \sum_a \xi^a)}{(1 + e^{\mu z + \tilde{z}})^n} \tag{34}$$

where both z and $\tilde{z}$ stay finite in the limit $\mu \to \infty$. In this sense, we have $\int d\tilde{z}\; P(z, \tilde{z}) = \tilde{P}(z)$, the effective distribution calculated in (27). We thus write (for $\mu \to \infty$)

$$\begin{aligned} P(z, \tilde{z}) &= \sum_{l=-1}^{\infty} p_l \delta(z + l) \rho^{(l)}(\tilde{z}) \\ p_l &= \frac{W(c)^{l+2}}{(l+1)!\; c} \end{aligned} \tag{35}$$

where $\rho^{(l)}(\tilde{z})$ describes the still unknown sub-dominant contributions to the effective potential $-\mu l$. Plugging ansatz (35) into equation (19) for the grand partition function, we see that the dominating part in $\ln \Xi$ is linear in μ, but vanishes finally once the saddle-point condition is used. As is shown in appendix B, the term of $O(\mu^0)$ can be calculated and leads to the entropy of minimal vertex covers,

$$\begin{aligned} s_{VC}(x_c(c)) &= \int \frac{dz\; dk}{2\pi} \int \frac{d\tilde{z}\; d\tilde{p}}{2\pi}\; e^{izk + i\tilde{z}\tilde{k}}\; P_{FT}(k, \tilde{k})\; [1 - \ln P_{FT}(k, \tilde{k})]\; \Phi(z, \tilde{z}) \\ &\quad + \frac{c}{2}\; p_{-1}^2 \int d\tilde{z}_1\; d\tilde{z}_2\; \rho^{(-1)}(\tilde{z}_1)\; \rho^{(-1)}(\tilde{z}_2)\; \ln(e^{-\tilde{z}_1} + e^{-\tilde{z}_1}) \end{aligned}$$

$$+cp_0 p_1 \int d\tilde{z}\ \rho^{(0)}(\tilde{z})\ \ln\left(1-\frac{1}{1+e^{-\tilde{z}}}\right)$$
$$+\frac{c}{2}p_0^2 \int d\tilde{z}_1\ d\tilde{z}_2\ \rho^{(0)}(\tilde{z}_1)\ \rho^{(0)}(\tilde{z}_2)\ \ln\left(1-\frac{1}{(1+e^{-\tilde{z}_1})(1+e^{-\tilde{z}_2})}\right) \quad (36)$$

where $P_{FT}(k,\tilde{k})$ signifies the two-dimensional Fourier transform of distribution $P(z,\tilde{z})$, and

$$\Phi(z,\tilde{z}) = \begin{cases} 0 & \text{if} \quad z<0 \\ \ln(1+e^{\tilde{z}}) & \text{if} \quad z=0 \\ \tilde{z} & \text{if} \quad z>0 \end{cases} \quad (37)$$

is the $O(\mu^0)$-term in $\ln(1+e^{\mu z+\tilde{z}})$. The corresponding saddle-point equations for the densities $\rho^{(l)}(\tilde{z})$, which are also calculated in appendix B, read

$$\begin{aligned} \rho_{FT}^{(-1)}(\tilde{k}) &= \exp\left\{-cp_0 + cp_0 \int d\tilde{z}\ \rho^{(0)}(\tilde{z})\ (1+e^{\tilde{z}})^{-i\tilde{k}}\right\} \\ \rho_{FT}^{(l)}(\tilde{k}) &= \rho_{FT}^{(-1)}(\tilde{k})\ \rho_{FT}^{(-1)}(-\tilde{k})^{l+1}\ . \end{aligned} \quad (38)$$

We now easily see that $\rho^{(0)}$ is an even distribution, and exactly half of the non-backbone vertices are covered in all minimal vertex covers. Within the region of validity of the replica-symmetric ansatz, this result is also verified numerically, see figure 8. Our argument for deriving equation (30) for the minimal vertex-cover size is thus completed. Using these saddle-point equations, we can eliminate all but one $\rho^{(l)}$ in the entropy (36). After a lengthy calculation which is again delegated to the appendix, we finally get

$$\begin{aligned} s_{VC}(x_c(c)) &= \frac{p_0}{2}\int \frac{d\tilde{z}\ d\tilde{k}}{2\pi} e^{i\tilde{z}\tilde{k}} \rho_{FT}^{(0)}(\tilde{k})[1-\ln\rho_{FT}^{(0)}(\tilde{k})]\ \ln(1+e^{\tilde{z}}) \\ &\quad +\frac{cp_0^2}{2}\int d\tilde{z}_1\ d\tilde{z}_2\ \rho^{(0)}(\tilde{z}_1)\ \rho^{(0)}(\tilde{z}_2)\ \ln\left(1-\frac{1}{(1+e^{-\tilde{z}_1})(1+e^{-\tilde{z}_2})}\right) \end{aligned} \quad (39)$$

which formally equals the expression (23) for the vertex-cover entropy for finite chemical potential μ, with c replaced by $2cp_0 = 2W(c)^2$. The main difference is however the self-consistent equation

$$\rho_{FT}^{(0)}(\tilde{k}) = \exp\left\{-2cp_0 + cp_0\int d\tilde{z}\rho^{(0)}(\tilde{z})\left[(1+e^{\tilde{z}})^{i\tilde{k}} + (1+e^{\tilde{z}})^{-i\tilde{k}}\right]\right\} \quad (40)$$

which can be obtained from (39) by optimization with respect to all *even* distributions $\rho^{(0)}(\tilde{z})$. We are however not able to solve the last equation, and are therefore restricted to variational approaches similar to [7]. A first upper estimate would be

$$s_{VC}(x_c(c)) \simeq \frac{p_0}{2}\ln 2 + \frac{cp_0^2}{2}\ln\frac{3}{4} \quad (41)$$

resulting from $\rho_{var}^{(0)} = \delta(\tilde{z})$. This results can be slightly improved by using a Gaussian variational ansatz for $\rho^{(0)}$, but the difference is only up to about 1%. For a comparison with numerical results see figure 6.

From equation (40) we are also able to read off analytically some limitations of the replica-symmetric solution (27). By developing (40) two second order in $\tilde{k}$, we find

$$\begin{aligned} \Delta^2 &:= \int d\tilde{z}\rho^{(0)}(\tilde{z})\tilde{z}^2 \\ &= 2cp_0\int_0^\infty d\tilde{z}\rho^{(0)}(\tilde{z})\left\{\ln(1+e^{-\tilde{z}})^2 + \ln(1+e^{\tilde{z}})^2\right\}\ . \end{aligned} \quad (42)$$

Rescaling $\tilde{z} = \Delta z$, we find

$$1 = \frac{2cp_0}{\Delta^2} \int_0^\infty dz \; \tilde{\rho}(z) \left\{ \ln(1+e^{-\Delta z})^2 + \ln(1+e^{\Delta z})^2 \right\} \tag{43}$$

with $\tilde{\rho}(z) = \rho^{(0)}(z/\Delta)/\Delta$ being of unit variance. For any $\tilde{\rho}$, the right-hand side is an monotonously decreasing function of Δ ranging from $+\infty$ for $\Delta_0 = 0$ to $cp_0 = W(c)^2$ for $\Delta \to \infty$. Identity (43) can thus be satisfied if and only if $W(c)^2 < 1$, which is valid for $c < e \simeq 2.718$. We thus have to conclude that our replica-symmetric solution (27) becomes inconsistent beyond average connectivity e, which is again in perfect agreement with the systematic deviations of numerical data beyond this point, cf. figures 3 and 4. Note however, that this point is far beyond the percolation threshold of the random graph. After the next subsection we will come back to this point, and consider more involved replica-symmetric and one-step broken saddle points.

6.5 The structure of the non-backbone subgraph

Before doing this, we will complete the discussion of the structure of minimal vertex covers in the region $0 < c < e$, where the described solution is expected to be exact. In this subsection we will concentrate on the structure of the non-backbone subgraph, *i.e.* the subgraph composed of all vertices which are not in the backbone, and all edges from E connecting these vertices.

The first intuition on the structure of the non-backbone component can be drawn from the saddle-point equation (26) for the distribution of effective chemical potentials. We consider an arbitrary vertex, and call the graph reduced which is obtained from the original graph by deleting the considered vertex and all its incident edges. According to the cavity interpretation of (26), the vertex is not in the backbone iff exactly one of its neighbors would be in the uncovered backbone of the reduced graph. The meaning of this becomes evident if we consider the non-backbone graphs in figure 7: Take *e.g.* the graph consisting of four vertices and three edges. All its vertices belong to the non-backbone. Deleting a boundary vertex, the reduced graph becomes backbone, in particular the neighbor of the boundary vertex belongs to the uncovered backbone. Deleting instead a central vertex, the reduced subgraph becomes disconnected into an isolated vertex, being uncovered backbone, and a connected vertex pair, being non-backbone.

Iterating this argument, we conclude that the non-backbone can be partitioned into $p_0 N/2$ pairs of vertices, every pair containing an edge and being eventually connected to other pairs or to covered backbone vertices. The supplementary edges connecting different pairs are conjectured to be drawn randomly and independently with the original probability c/N between any non-backbone vertices.

Even if we are not able to prove this conjecture, we may give strong arguments to support it:

- Looking at the non-backbone subgraphs of small tree-like graphs, the predicted structure is found. A cluster expansion up to connected clusters of four vertex pairs provides lower and upper bounds for the entropy which are in good agreement with numerical findings (*e.g.* in the first four non-zero digits for $c = 0.1$).

- We can apply the statistical-mechanics formalism to a restricted random graph ensemble having exactly the properties described above. This directly leads to expressions (39) and (40) for the entropy and the effective-potential distribution.

- The proposed structure results in an even distribution of effective potentials for connectivities $c < e$, whereas the average occupation density is expected to exceed $1/2$ for $c > e$. This is verified numerically, see figure 8.

- The average connectivity of a vertex pair to other vertex pairs in the restricted ensemble is $2cp_0 = 2W(c)^2$, the percolation threshold would therefore be at $W(c) = 1/\sqrt{2}$, *i.e.* at $c = \exp(1/\sqrt{2})/\sqrt{2} \simeq 1.434$. We have checked this numerically by calculating the fraction of non-backbone vertices in the largest connected component of the non-backbone subgraph, see Fig. 9. This quantity clearly extrapolates to 0 for connectivities below the percolation point, and saturates at a finite value for larger connectivities. The reason why this transition is shifted to higher connectivity compared to graph percolation, becomes obvious by considering the action of covered backbone vertices. They "cut" the graph into smaller pieces. Please remember also that high-connectivity vertices are more frequently found in the covered backbone, making this cutting mechanism more effective.

- We should add the remark that we have performed a similar numerical study for the backbone subgraph. We found that the percolation threshold of the backbone subgraph is identical to the percolation threshold $c = 1.0$ of the whole graph.

The percolation however does not bother the validity of the replica-symmetric result, which is valid even for percolated non-backbone subgraphs as long as $c < e$. The proposed structure also allows for a very simple interpretation of approximation (41) of the entropy of minimal VCs: An isolated pair contributes an entropy $\ln 2$ as it has two possible minimal VCs, thus explaining the first term in (41). This entropy is decreased by the insertion of supplementary edges. The simplest structure are chains of four vertices, every one having only three minimal VCs – leading directly to the second term in (41) as two pairs are connected with probability $4c/N$. More complicated non-backbone graphs lead to corrections, and may be included by a non-trivial $\rho^{(0)}$.

6.6 Unphysical replica-symmetric saddle points

We have seen, that solution (27) of the replica-symmetric saddle-point equation (26) is correct only up to average connectivity $c = e$. Before searching for replica-symmetry-broken saddle points, we should however exclude the existence of further replica-symmetric saddle points. We therefore consider again equation (26). It is consistent with any ansatz

$$\tilde{P}_m(z) = \sum_{l=-m}^{\infty} p_l^{(m)} \delta\left(z + \frac{l}{m}\right) \ . \tag{44}$$

One can easily write down the self-consistent conditions for the probabilities $p_l^{(m)}$, and find out that, for $m > 1$, these have non-trivial solutions with positive $p_l^{(m)}$s only for $c > e$. We will show this explicitly only for $m = 2$. The saddle-point equations read

$$\begin{aligned} p_{-2} &= \exp\{-c(p_{-2} + p_{-1})\} \\ p_{-1} &= \exp\{-c(p_{-2} + p_{-1})\} c p_{-1} \\ p_0 &= \exp\{-c(p_{-2} + p_{-1})\}(c p_{-2} + \frac{c}{2} p_{-1}^2) \\ &\cdots \end{aligned} \tag{45}$$

The only solution with non-zero p_{-1} is

$$\begin{aligned} p_{-2} &= \frac{1}{c} \\ p_{-1} &= \frac{\ln c - 1}{c} \\ p_0 &= \frac{2 + (\ln c - 1)^2}{2c} \\ &\cdots \end{aligned} \tag{46}$$

p_{-1} is obviously positive only if $c > e$.

The corresponding threshold $x_c(c)$ would be larger than the old one resulting from $m = 1$, which is correct compared to the systematic deviation of the numerical data. The multi-peak solutions (44) are however unphysical due to the existence of effective potentials of *e.g.* the value μ/m. This positive potential would force the corresponding vertex to be in the uncovered backbone for large μ, but the only physical mechanism for this is given by the global chemical potential μ. The influence of neighbors results only in negative or zero potentials. Positive fractions of μ are consequently unphysical.

We can however interpret multi-peak solutions as a kind of hidden replica-symmetry breaking. This will become clear in the following section.

7 The simplest one-step replica symmetry broken solution

This section is dedicated to the appearance of replica-symmetry breaking (RSB) in VC. Despite several efforts, the question of how to handle RSB in finite-connectivity systems is still open. Most attempts [41, 42, 43, 44] try to apply the first step of Parisi's RSB scheme [38] which however was originally developped for infinite-connectivity spin glasses. Due to the more complex structure of the order parameter a complete solution is however still missing. Recently, based on the connection to combinatorial optimization, the interest in this question was renewed [39], and some promising approximation schemes [7, 45] have been developed. Here we closely follow the approach proposed in [39] which allows to construct a simple one-step RSB solution.

In case of one-step RSB, the full permutation symmetry of the order parameter corresponding to the equivalence of all n replicas breaks down. According to Parisi's scheme, the replicas can be grouped into n/m blocks of equal size m, where the symmetry is now restricted to permutations of replicas within every block, or to permutations of full blocks. We therefore introduce a new numbering of replicas by index pairs (a, α), with $a = 1, ..., n/m$ denoting the block number, and $\alpha = 1, ..., m$ counting the replicas within block a. Due to the described symmetry, the order parameters $c(\vec{\xi})$ thus depend on $\vec{\xi}$ only via the block quantities $s^a = \sum_{\alpha=1}^{m} \xi^{a\alpha}$, or even more precisely, on the number of blocks having $s^a = ym$, which can be described by

$$\nu(y) = \sum_{a=1}^{n/m} \delta(y - s^a/m) \ . \tag{47}$$

y stands for the average occupation number of a block and ranges from 0 to 1. Its discrete nature present for natural n vanishes in the analytical continuation needed for the replica limit $n \to 0$.

Following the cavity-like argumentation of Monasson [39], the order parameter can be expressed as

$$\begin{aligned} c(\vec{\xi}) &= \int \mathcal{D}\rho\ \mathcal{P}[\rho] \prod_{a=1}^{n/m} \int dh\ \rho(h)\ \frac{e^{hs^a}}{(1+e^h)^m} \\ &= \int \mathcal{D}\rho\ \mathcal{P}[\rho]\ \exp\left\{ \int_0^1 dy\ \nu(y) \ln\left[\int dh\ \rho(h)\ \frac{e^{hmy}}{(1+e^h)^m} \right] \right\} \\ &=: c[\nu] \end{aligned} \tag{48}$$

where $\mathcal{P}[\rho]$ is a histogram of the local distributions $\rho_i(h)$, which themselves are histograms of local effective potentials over all thermodynamically relevant pure states, see [38] for a detailed discussion of this interpretation. In the second line, the analytic continuation in n has already been made, m is now considered as usually as a parameter in the interval $[0,1]$ which has to be optimized in the saddle-point solution. The only requirement to $\nu(y)$ is, that

$$\int_0^1 dy\ \nu(y) = \frac{n}{m} \to 0 \tag{49}$$

vanishes in the replica limit $n \to 0$.

This ansatz can be plugged into the saddle-point equation (20)

$$c(\vec{\xi}) = \exp\left\{ -c + \mu \sum_{a\alpha} \xi^{a\alpha} + c \sum_{\vec{\zeta}} c(\vec{\zeta}) \prod_{a\alpha} (1 - \xi^{a\alpha}\zeta^{a\alpha}) \right\} . \tag{50}$$

Proceeding term by term on the right-hand side, we find

$$\sum_{a\alpha} \xi^{a\alpha} = \sum_a s^a = m \int_0^1 dy\ v(y)\ y \tag{51}$$

and

$$\begin{aligned} \sum_{\vec{\zeta}} c(\vec{\zeta}) \prod_{a\alpha} (1 - \xi^{a\alpha}\zeta^{a\alpha}) &= \int \mathcal{D}\rho\ \mathcal{P}[\rho] \prod_{a=1}^{n/m} \int dh\ \rho(h)\ \prod_{\alpha=1}^{m} \left[\sum_{\zeta=\pm 1} \frac{e^{h\zeta}(1-\xi^{a\alpha}\zeta)}{1+e^h} \right] \\ &= \int \mathcal{D}\rho\ \mathcal{P}[\rho] \prod_{a=1}^{n/m} \int dh\ \rho(h)\ \left[\frac{1}{1+e^h} \right]^{s^a} \\ &= \int \mathcal{D}\rho\ \mathcal{P}[\rho]\ \exp\left\{ \int_0^1 dy\ \nu(y)\ \ln\left[\int dh\ \rho(h)\ (1+e^h)^{-my} \right] \right\} \end{aligned}$$

Plugging this results together with equations (48) and (51) into (50), we obtain for $n = 0$ a closed equation for $\mathcal{P}[\rho]$ which has to be fulfilled for every $\nu(y)$ satisfying condition (49).

This saddle-point equation is still valid for any chemical potential. In the limit of minimal vertex covers, *i.e.* for $\mu \to \infty$, this equation simplifies again. For the $\rho(h)$ we assume an ansatz similar to the replica-symmetric value for $P(h)$ in (27)

$$\rho(h) = \omega_\mu \sum_{l=l_-}^{l_+} \rho_l e^{\mu m |l|/2} \delta(h + \mu l) \tag{52}$$

where the support of ρ is now restricted by l-values between (l_-, l_+) with $l_- \geq -1$. The constant prefactor ω_μ becomes irrelevant for $n \to 0$ due to condition (49). The exponential factor is inspired by its appearance in infinite-connectivity models, cf. [39]. Please note that the replica-symmetric case can be obtained by $l_- = l_+$. Introducing weights $\mathcal{P}_{l_-,l_+}$ as the integrated weight of all $\rho(h)$ having the same (l_-, l_+), the order parameter simplifies to

$$\lim_{\mu\to\infty} c[\nu/\mu] = \sum_{-1\le l_-\le l_+} \mathcal{P}_{l_-,l_+} \exp(\nu_- l_- + \nu_+ l_+) \tag{53}$$

where $\nu_- = \int_{1/2}^{1} dy\ \nu(y)\ y$ and $\nu_+ = \int_0^{1/2} dy\ \nu(y)\ y$. Details of this calculation are delegated to appendix C. Our saddle-point equation thus becomes

$$\sum_{-1\le l_-\le l_+} \mathcal{P}_{l_-,l_+} \exp(\nu_- l_- + \nu_+ l_+) = \exp\left\{-c + \mu(\nu_- + \nu_+) + c\mathcal{P}_{-1,-1} e^{\nu_- + \nu_+} + c\mathcal{P}_{-1,0} e^{\nu_-}\right\} \tag{54}$$

which has to be fulfilled for all ν_-, ν_+. Please note that the m-dependence is completely disappeared [46]. This equation can be easily solved:

$$\begin{aligned}
\mathcal{P}_{-1,-1} &= \frac{1}{c} \\
\mathcal{P}_{-1,0} &= \frac{\ln(c)-1}{c} \\
\mathcal{P}_{l_-,l_+} &= \frac{c^{l_+ +1}}{(l_- + 1)!(l_+ - l_-)!} {\mathcal{P}_{-1,-1}}^{l_- +2} {\mathcal{P}_{-1,0}}^{l_+ - l_-} \\
&= \frac{(\ln(c)-1)^{l_+ - l_-}}{(l_- + 1)!(l_+ - l_-)!c}
\end{aligned} \tag{55}$$

Let us discuss this solution:

- At first we realize that $\mathcal{P}_{-1,0}$ is positive only for connectivities $c > e$. This is consistent with our previous finding that replica symmetry is restricted to smaller c.

- Introducing p_l as the sum over all $\mathcal{P}_{l_-,l_+}$ having $l = l_- + l_+$, saddle-point equation (55) reduces for $\nu_- = \nu_+$ to equations (45) for the unphysical replica-symmetric saddle point showing half-integer valued effective potentials. This underlines the interpretation of these solutions as hidden-RSB solutions.

- As we do not know the non-backbone magnetization in the RSB solution, we are only able to give lower and upper estimates for $x_c(c)$. The upper one, $x_c(c) < 1 - \mathcal{P}_{-1,-1} - \mathcal{P}_{-1,0} = 1 - \ln(c)/c$ coincides with the rigorous upper bound of Gazmuri [20]. The lower one would be $x_c(c) > 1 - \mathcal{P}_{-1,-1} - \mathcal{P}_{-1,0} - \mathcal{P}_{-1,+1} - \mathcal{P}_{0,0}$. Having in mind the numerical result, that non-backbone effective potentials have a positive bias, we can however conclude $x_c(c) > 1 - \mathcal{P}_{-1,-1} - \mathcal{P}_{-1,0} - \mathcal{P}_{-1,+1}/2 - \mathcal{P}_{0,0}/2$ which is slightly better than the replica-symmetric result. In figure 3 both results are nearly indistinguishable, so we have omitted the RSB data from the figure.

- Also the evaluation of the backbone-size is slightly subtle. In principle we would expect that backbone vertices have $\rho(h)$ which are supported either only on positive or only on negative fields. This would result in

$$\begin{aligned}
b^{(1)}_{uncov}(c) &= \mathcal{P}_{-1,-1} = \frac{1}{c} \\
b^{(1)}_{cov}(c) &= \sum_{1\le l_-\le l_+} \mathcal{P}_{l_-,l_+} = 1 - \frac{2}{e}\ .
\end{aligned} \tag{56}$$

Due to the existence of the exponential factors in ansatz (52) also $\mathcal{P}_{l_-,l_+}$ with $l_- \neq -l_+$ lead to average occupation numbers zero and one, and thus contribute to the backbone:

$$\begin{aligned} b^{(2)}_{uncov}(c) &= \mathcal{P}_{-1,-1} + \mathcal{P}_{-1,0} = \frac{\ln c}{c} \\ b^{(2)}_{cov}(c) &= \sum_{l_- \geq -1; l_+ > |l_-|} \mathcal{P}_{l_-,l_+} \\ &= 1 - \frac{1}{c}\left(1 + \ln c + \frac{(1-\ln c)^2}{4}\right) . \end{aligned} \tag{57}$$

Both values do not coincide with numerical findings, see also figure 4. Probably this could be cured by assuming $m \sim \mu^{-1}$, cf. [46], instead of $m \sim \mu^0$. This would remove the exponential dominance of fields of largest absolute value for $\mu \to \infty$. We could however construct no solution to this case.

We may conclude that the presented one-step saddle point improves the replica-symmetric findings for $x_c(c)$, but is still plagued by certain problems. It remains an open question, if these problems can be cured already by including a different scaling of m, or if finally more than one step of RSB is required.

8 Conclusion and outlook

In this paper, we have presented a detailed analysis of size and structure of minimal vertex covers on random graphs. In particular, we have calculated the size dependence of minimal VCs on the average connectivity, and we ahve shown that those VCs are exponentially numerous. Many statistical properties, as *e.g.* the partial freezing into backbone and non-backbone vertices, could be characterized. All our results are based on exact numerical enumerations as well as replica calculations. We have found that replica-symmetric results appear to be exact up to graph connectivities $c = e \simeq 2.718$, whereas replica-symmetry breaking has to be included for an understanding of higher-connectivity graphs. This is however a complicated task: Even if there has been some recent progress on the question of one-step replica-symmetry breaking in finite connectivity systems based on various approximation schemes [39, 7, 45], a definite technical approach is still missing. Due to the simplicity of its replica-symmetric solution, as compared e.g. to satisfiability problems [5], vertex cover could be a good model for further progress into this direction.

In our paper, we have only considered finite-connectivity random graphs. These show however a very simple geometrical structure. They are locally tree-like, and loops are of length $O(\ln N)$. It would be interesting to considered therefore restricted graph ensembles which include non-trivial local structures. The question of such topological influences on the solution structure of combinatorial optimization problems still remains an interesting open question, as also other studied problems include mainly locally tree-like problems [1, 2]. Restricted graph ensembles could therefore provide a possible starting point for further research.

A last comment concerns the interpretation of vertex covers as packings of hard spheres on random lattices. We were able to describe the maximally dense packings, which were found to show very interesting properties due to the disorder present in the graph: There where backbone sites having the same occupation state in all densest packings, whereas others are found to be free in some packings, occupied in others. This effect resembles the existence of blocked and unblocked particles in real packings. With some modifications, the hard-sphere lattice gas can therefore be understood as a possible mean-field model of granular packings, compare also [30]. Work is in progress along these lines.

Acknowledgements

The authors are deeply indebted to R. Monasson and R. Zecchina for many fruitful discussions. AKH acknowledges financial support by the DFG (*Deutsche Forschungsgemeinschaft*) under grant Zi209/6-1.

A The replica-symmetric limit $n \to 0$

Starting from equations (19) and (20) we will present the calculation of the replica limit $n \to 0$ under the replica-symmetric ansatz

$$c(\vec{\xi}) = \int dh P(h) \frac{\exp(h \sum_a \xi^a)}{(1+e^h)^n} \tag{58}$$

The procedure is very similar to the one presented in [39] for Ising-spin-glass models. We start with the grand partition function as given in (19):

$$\lim_{N\to\infty} \frac{1}{N} \overline{\ln \Xi} = \lim_{n\to 0} \frac{1}{n} \left(-\sum_{\vec{\xi}} c(\vec{\xi}) \ln c(\vec{\xi}) - \frac{c}{2} + \mu \sum_{\vec{\xi},a} c(\vec{\xi}) \xi^a + \frac{c}{2} \sum_{\vec{\xi},\vec{\zeta}} c(\vec{\xi}) c(\vec{\zeta}) \prod_a (1 - \xi^a \zeta^a) \right) \tag{59}$$

where $c(\vec{\xi})$ takes its saddle-point value. At first, we consider the combinatorial entropy, and use again a replica trick:

$$\sum_{\vec{\xi}} c(\vec{\xi}) \ln c(\vec{\xi}) = \left[\frac{\partial}{\partial l} \sum_{\vec{\xi}} c(\vec{\xi})^l \right]_{l=1} \tag{60}$$

Assuming positive integer l at the beginning, and plugging in the replica-symmetric ansatz for $c(\vec{\xi})$, we write

$$\begin{aligned}
\sum_{\vec{\xi}} c(\vec{\xi})^l &= \int dh_1 \cdots dh_l P(h_1) \cdots P(h_l) \sum_{\vec{\xi}} \frac{\exp(\sum_{m=1}^l h_m \sum_a \xi^a)}{\prod_{m=1}^l (1+e^{h_l})^n} \\
&= \int dh_1 \cdots dh_l P(h_1) \cdots P(h_l) \left(\sum_{\xi=0,1} \frac{\exp(\sum_l h_l \xi)}{\prod_{m=1}^l (1+e^{h_l})} \right)^n \\
&= 1 + n \int dh_1 \cdots dh_l P(h_1) \cdots P(h_l) \ln(1 + \exp\{\sum_m h_m\}) \\
&\quad -nl \int dh\ P(h) \ln(1+e^h) + O(n^2)
\end{aligned}$$

Introducing new variables $H_k = \sum_{m=1}^k h_m$, the last expression becomes

$$\begin{aligned}
\sum_{\vec{\xi}} c(\vec{\xi})^l &= 1 + n \int dH_1 \cdots dH_l P(H_1) P(H_2 - H_1) \cdots P(H_l - H_{l-1}) \ln(1 + e^{H_l}) \\
&\quad -nl \int dh\ P(h) \ln(1+e^h) + O(n^2) \\
&= 1 + n \int dH_l \int \frac{dk}{2\pi} e^{iH_l k} P_{FT}(k)^l \ln(1+e^{H_l}) - nl \int dh\ P(h) \ln(1+e^h) + O(n^2)
\end{aligned}$$

In the last step we have used the fact, that the l-fold convolution of $P(h)$ with itself can be express as the Fourier-back transformation of the lth power of its Fourier transform P_{FT}. Now the differentiation with respect to l can be carried out, and according to (60) we find

$$\sum_{\vec{\xi}} c(\vec{\xi}) \ln c(\vec{\xi}) = n \int \frac{dh\, dk}{2\pi} e^{ihk} P_{FT}(k) \left[-1 + \ln P_{FT}(k)\right] \ln(1+e^h) \ . \tag{61}$$

The other terms in (59) can be evaluated directly,

$$\begin{aligned}
\sum_{\vec{\xi},a} c(\vec{\xi})\ \xi^a &= n \int dh\ P(h)\ \frac{e^h (1+e^h)^{n-1}}{(1+e^h)^n} \\
&= n \int dh\ P(h)\ \frac{1}{(1+e^{-h})} \\
\sum_{\vec{\xi},\vec{\zeta}} c(\vec{\xi}) c(\vec{\zeta}) \prod_a (1-\xi^a\zeta^a) &= \int dh_1 dh_2\ P(h_1)\ P(h_2) \sum_{\vec{\xi},\vec{\zeta}} \prod_{a=1}^{n} \frac{(1-\xi^a\zeta^a)\exp(h_1\xi^a + h_2\zeta^a)}{(1+e^{h_1})(1+e^{h_2})} \\
&= \int dh_1 dh_2\ P(h_1)\ P(h_2) \left[1 - \frac{e^{h_1+h_2}}{(1+e^{h_1})(1+e^{h_2})}\right]^n \\
&= 1 + n \int dh_1 dh_2\ P(h_1)\ P(h_2) \ln\left[1 - \frac{e^{h_1+h_2}}{(1+e^{h_1})(1+e^{h_2})}\right] + O(n^2)
\end{aligned}$$

Putting these results together, we find

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N} \overline{\ln \Xi} &= \int \frac{dh\, dk}{2\pi} e^{ihk} P_{FT}(k) \left[1 - \ln P_{FT}(k)\right] \ln(1+e^h) + \mu \int dh\ P(h)\ \frac{1}{(1+e^{-h})} \\
&\quad - \frac{c}{2} \int dh_1 dh_2\ P(h_1)\ P(h_2) \ln\left[1 - \frac{e^{h_1+h_2}}{(1+e^{h_1})(1+e^{h_2})}\right]
\end{aligned} \tag{62}$$

which finally results in equation (23) for the vertex-cover entropy.

For the saddle-point equation

$$c(\vec{\xi}) = \exp\left\{ -\lambda + \mu \sum_a \xi^a + c \sum_{\vec{\zeta}} c(\vec{\zeta}) \prod_a (1-\xi^a\zeta^a) \right\} \ . \tag{63}$$

we proceed analogously. Obviously both side depend on $\vec{\xi}$ only via $y = \sum_a \xi^a$. The left-hand side thus simplifies for $n \to 0$:

$$c(\vec{\xi}) \to_{n\to 0} \int dh\ P(h)\ e^{hy} \ , \tag{64}$$

whereas the right-hand side (rhs) gives

$$rhs = \exp\left\{ -\lambda + \mu y + c \int dh\ P(h) \left(\frac{1}{1+e^h}\right)^y \right\} \tag{65}$$

We now can determine the Lagrange multiplier from the normalization of $P(h)$. For $y = 0$, the left-hand side equals one, whereas the right-hand side equals $\exp(-\lambda + c)$, which results directly in $\lambda = c$, and thus in the replica-symmetric saddle-point equation (22).

The same saddle-point equation can of course be derived by varying equation (62) directly with respect to $P(h)$. Note however that the result given here is stronger: We have shown that the original saddle-point equation for $c(\vec{\xi})$ is closed under our replica-symmetric ansatz, thus leading to a real saddle point of the free energy. The second procedure would however be important if we would use a variational ansatz which does not close the $c(\vec{\xi})$-equation.

B Calculation of the entropy

For calculating the entropy of minimal vertex covers, we start again with equation (19) for the disorder-averaged grand partition function, but now we plug in the refined ansatz (34), *i.e.*

$$c(\vec{\xi}) = \int dz\ d\tilde{z}\ P(z,\tilde{z}) \frac{\exp((\mu z + \tilde{z})\sum_a \xi^a)}{(1+e^{\mu z+\tilde{z}})^n} \tag{66}$$

where $P(z,\tilde{z})$ is assumed to stay a well-behaved probability distribution in the limit $\mu \to \infty$ of minimal vertex covers. Consistency with the dominant behavior discussed in section 6.2 requires

$$\int d\tilde{z} P(z,\tilde{z}) = \sum_{l=-1} p_l \delta(z+l) \tag{67}$$

with

$$p_l = \frac{W(c)^{l+2}}{(l+1)!\ c} \tag{68}$$

for $\mu \to \infty$, cf. (27). We therefore may write

$$P(z,\tilde{z}) = \sum_{l=-1} p_l \delta(z+l)\rho^{(l)}(\tilde{z}) \tag{69}$$

with probability distributions $\rho^{(l)}(\tilde{z})$ which still have to be determined. By plugging ansatz (66) into $\overline{\ln \Xi}$ as given in (19) and following the same procedure as in the last section, we find for finite μ:

$$\begin{aligned}
\lim_{N\to\infty} \frac{\overline{\ln \Xi}}{N} \;=\; & \int \frac{dz\ dk}{2\pi} \int \frac{d\tilde{z}\ d\tilde{k}}{2\pi} e^{izk+i\tilde{z}\tilde{k}} P_{FT}(k,\tilde{k}) \left[1 - \ln P_{FT}(k,\tilde{k})\right] \ln(1+e^{\mu z+\tilde{z}}) \\
& + \mu \int dz\ d\tilde{z}\ P(z,\tilde{z})\ \frac{1}{(1+e^{-\mu z-\tilde{z}})} \\
& - \frac{c}{2} \int dz_1 dz_2 d\tilde{z}_1 d\tilde{z}_2 P(z_1,\tilde{z}_1) P(z_2,\tilde{z}_2) \ln\left[1 - \frac{1}{(1+e^{-\mu z_1-\tilde{z}_1})(1+e^{-\mu z_2-\tilde{z}_2})}\right]
\end{aligned} \tag{70}$$

with $P_{FT}(k,\tilde{k}) = \int dz\ d\tilde{z}\ P(z,\tilde{z}) \exp\{-izk - i\tilde{z}\tilde{k}\}$ being the 2d Fourier-transform of $P(z,\tilde{z})$. For $\mu \to \infty$, the dominant behavior seems to be of $O(\mu)$, but its coefficient has to vanish at the saddle point as the entropy stays finite. This has been checked explicitly, without presenting those details we therefore concentrate on the second term of $O(\mu^0)$ which will give the entropy of minimal vertex covers,

$$s_{VC}(x_c(c)) = \lim_{\mu\to\infty} \left(\lim_{N\to\infty} \frac{1}{N} \overline{\ln \Xi} - \mu \int dz\ d\tilde{z}\ P(z,\tilde{z})\ \frac{1}{(1+e^{-\mu z-\tilde{z}})} \right) . \tag{71}$$

Starting with the first term in (70), we have to leading orders

$$\begin{aligned}
\ln(1+e^{\mu z+\tilde{z}}) \;&\to\; \mu z \Theta(z) + \Phi(z,\tilde{z}) \\
\Phi(z,\tilde{z}) \;&=\; \begin{cases} 0 & \text{if } z<0 \\ \ln(1+e^{\tilde{z}}) & \text{if } z=0 \\ \tilde{z} & \text{if } z>0 \end{cases}
\end{aligned} \tag{72}$$

At the moment, replacing $\ln(1+e^{\mu z+\tilde{z}})$ by $\Phi(z,\tilde{z})$ is all we can do in the first term without using the saddle-point equations for the $\rho^{(l)}$s. The situation is better for the last term in

(70). Having in mind that z can take only integer values smaller or equal to $+1$, we use

$$\ln\left[1-\frac{1}{(1+e^{-\mu z_1-\tilde{z}_1})(1+e^{-\mu z_2-\tilde{z}_2})}\right]\to\begin{cases}-\mu+\ln(e^{-\tilde{z}_1}+e^{-\tilde{z}_2}) & \text{if} \quad z_1=z_2=1\\ \ln\left[1-\frac{1}{1+e^{-\tilde{z}_1}}\right] & \text{if} \quad z_1=0,\ z_2=1\\ \ln\left[1-\frac{1}{1+e^{-\tilde{z}_2}}\right] & \text{if} \quad z_1=1,\ z_2=0\\ \ln\left[1-\frac{1}{(1+e^{-\tilde{z}_1})(1+e^{-\tilde{z}_2})}\right] & \text{if} \quad z_1=z_2=0\\ 0 & \text{if} \quad z_1,z_2<0\end{cases} \tag{73}$$

where all terms are dropped which are exponentially small in μ. Plugging this result into (70), we find

$$\begin{aligned} s_{VC}(x_c(c)) \quad = \quad & \int\frac{dz\ dk}{2\pi}\int\frac{d\tilde{z}\ d\tilde{p}}{2\pi}\ e^{izk+i\tilde{z}\tilde{k}}\ P_{FT}(k,\tilde{k})\ [1-\ln P_{FT}(k,\tilde{k})]\ \Phi(z,\tilde{z})\\ & +\frac{c}{2}\ p_{-1}^2\int d\tilde{z}_1\ d\tilde{z}_2\ \rho^{(-1)}(\tilde{z}_1)\ \rho^{(-1)}(\tilde{z}_2)\ \ln(e^{-\tilde{z}_1}+e^{-\tilde{z}_1})\\ & +cp_0p_1\int d\tilde{z}\ \rho^{(0)}(\tilde{z})\ \ln\left(1-\frac{1}{1+e^{-\tilde{z}}}\right)\\ & +\frac{c}{2}p_0^2\int d\tilde{z}_1\ d\tilde{z}_2\ \rho^{(0)}(\tilde{z}_1)\ \rho^{(0)}(\tilde{z}_2)\ \ln\left(1-\frac{1}{(1+e^{-\tilde{z}_1})(1+e^{-\tilde{z}_2})}\right) \end{aligned} \tag{74}$$

which is equation (36).

We continue again with the derivation of the saddle-point equation, again we start from the original equation for $c(\vec{\xi})$ as given in eq. (20),

$$c(\vec{\xi})=\exp\left\{-\lambda+\mu\sum_a\xi^a+c\sum_{\vec{\zeta}}c(\vec{\zeta})\prod_a(1-\xi^a\zeta^a)\right\}\ . \tag{75}$$

Plugging in the replica-symmetric ansatz and continuing analogously to the previous appendix for $n\to 0$, we find

$$\int dz\ d\tilde{z}P(z,\tilde{z})e^{\mu zk+\tilde{z}k}=\exp\left\{-c+\mu k+c\int dz\ d\tilde{z}P(z,\tilde{z})\left[1+e^{\mu z+\tilde{z}}\right]^{-k}\right\} \tag{76}$$

For $k=O(\mu^{-1})$ in the limit $\mu\to\infty$, we find back the old saddle-point equation for the dominant effective chemical potentials. The saddle-point equations for the sub-dominant corrections $\rho^{(l)}(\tilde{z})$ are however obtained for $k=O(\mu^0)$. The corresponding limit $\mu\to\infty$ is not obvious due to the existence of terms like μk. We have to use equation (69). The left-hand side of the last equation thus reads

$$\int dz\ d\tilde{z}P(z,\tilde{z})e^{\mu zk+\tilde{z}k}=\sum_{l=-1}^{\infty}p_le^{-\mu kl}\int d\tilde{z}\rho^{(l)}(\tilde{z})e^{\tilde{z}k} \tag{77}$$

The dominant contribution for large μ and positive k is given by the term having $l=-1$, and diverges exponentially as $e^{\mu k}$. Multiplying equation (76) with $e^{-\mu k}$ thus yields a well-defined limit $\mu\to\infty$, we find

$$\rho_{FT}^{(-1)}(k)=\exp\left\{-cp_0+cp_0\int d\tilde{z}\rho^{(0)}(\tilde{z})\left(1+e^{\tilde{z}}\right)^{ik}\right\}\ . \tag{78}$$

We now proceed by subtracting the dominant contributions $\sim e^{\mu k}$ on both sides of eq. (76), and find for $\mu \to \infty$

$$\begin{aligned} \rho_{FT}^{(0)}(k) &= \exp\left\{ -cp_0 + cp_0 \int d\tilde{z} \rho^{(0)}(\tilde{z}) \left(1+e^{\tilde{z}}\right)^{ik} \right\} \rho_{FT}^{(-1)}(-k) \\ &= \rho_{FT}^{(-1)}(k)\ \rho_{FT}^{(-1)}(-k) \end{aligned} \tag{79}$$

where we have used (78) in the last line. Continuing by iteration, we finally find

$$\rho_{FT}^{(l)}(k) = \rho_{FT}^{(-1)}(k)\ \rho_{FT}^{(-1)}(-k)^l\ . \tag{80}$$

So it is very simple to solve all but one of these saddle point equations. We can consequently express $s_{VC}(x_c(c))$ in terms of $\rho^{(0)}(\tilde{z})$, as is done in eq. (39) in section 6.4. $\rho^{(0)}(\tilde{z})$ itself is described by (40) which follows directly from equations (78,79). The corresponding calculations are lengthy but straight-forward, so we do not present it here. The only trick which has to be used is the following: Using (80) we may write

$$\begin{aligned} P_{FT}(k,\tilde{k}) &= \sum_{l=-1}^{\infty} \frac{W(c)^{l+2}}{(l+1)!\ c} e^{ikl} \rho_{FT}^{(l)}(\tilde{k}) \\ &= \frac{W(c)}{c} e^{-ik} \rho_{FT}^{(-1)}(\tilde{k}) \sum_{l=-1}^{\infty} \frac{1}{(l+1)!} \left[W(c) e^{ik} \rho_{FT}^{(-1)}(-\tilde{k}) \right]^{l+1} \\ &= \frac{W(c)}{c} e^{-ik} \rho_{FT}^{(-1)}(\tilde{k})\ \exp\left\{ W(c) e^{ik} \rho_{FT}^{(-1)}(-\tilde{k}) \right\}\ . \end{aligned} \tag{81}$$

This expression helps to simplify $\ln P_{FT}(k,\tilde{k})$ in equation (74).

C Evaluation of the RSB saddle-point equation

This last appendix shows how the $\mu \to \infty$ limit can be taken in the one-step RSB saddle-point equation. We start with the order paramter as given in (48)

$$c[\nu] = \int \mathcal{D}\rho\ \mathcal{P}[\rho]\ \exp\left\{ \int_0^1 dy\ \nu(y) \ln\left[\int dh\ \rho(h)\ \frac{e^{hmy}}{(1+e^h)^m} \right] \right\} \tag{82}$$

and plug in ansatz (52),

$$\rho(h) = \omega_\mu \sum_{l=l_-}^{l_+} \rho_l e^{\mu m |l|/2} \delta(h+\mu l)\ . \tag{83}$$

We assume in particular that $\rho_{l_\pm} \neq 0$ for uniqueness of the definition of l_- and l_+. Setting $\rho_l = 0$ for all $l < l_-$ and all $l > l_+$, we find for the exponent in (82):

$$\begin{aligned} \{...\} &= \int_0^1 dy\ \nu(y) \ln\left[\int dh\ \rho(h)\ \frac{e^{hmy}}{(1+e^h)^m} \right] \\ &= \int_0^1 dy\ \nu(y) \ln\left[\omega_\mu \rho_{-1} e^{\mu m(y-1/2)} + \omega_\mu \frac{\rho_0}{2^m} + \omega_\mu \sum_{l>0} \rho_l e^{-\mu m(y-1/2)} \right] \end{aligned} \tag{84}$$

ω_μ can be skipped in the last line, because $\nu(y)$ has to have zero integral due to (49). For large μ this is exponentially dominated by only one term which depends on y: If $y < 1/2$ the term with $l = l_+$ dominates, for $y > 1/2$ the l_--term becomes exponentially larger than all others. Introducing

$$\begin{aligned} \nu_- &= \int_{1/2}^{1} dy\ \nu(y)\ y \\ \nu_- &= \int_{0}^{1/2} dy\ \nu(y)\ y \end{aligned} \tag{85}$$

we conclude

$$\lim_{\mu\to\infty} \mu^{-1}\{...\} = \nu_- l_- + \nu_+ l_+\ , \tag{86}$$

and

$$\lim_{\mu\to\infty} c[\nu/\mu] = \sum_{-1\le l_- \le l_+} \mathcal{P}_{l_-,l_+} e^{\nu_- l_- + \nu_+ l_+} \tag{87}$$

which is the left-hand side of the saddle-point equation. On the right-hand side, an integral similar to (84) has to be determined. Following exactly the same scheme as above, we find the expression given in equation (54).

References

[1] T. Hogg, B. A. Huberman, and C. Williams (eds.), Frontiers in problem solving: phase transitions and complexity, Artif. Intell. 81 (I+II) (1996)

[2] O. Dubois, R. Monasson, B. Selman, and R. Zecchina, special issue of Theor. Comp. Sci., to appear (2001)

[3] M. R. Garey and D. S. Johnson, Computers and Intractability, (Freeman, San Francisco, 1979)

[4] D. Mitchell, B. Selman, and H. Levesque, in Proc. 10th Natl. Conf. Artif. Intell. (AAAI-92), 440 (AAAI Press/MIT Press, Cambridge, Massachusetts, 1992)

[5] R. Monasson, R. Zecchina, Phys. Rev. E56, 1357 (1997)

[6] R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, and L. Troyansky, Nature 400, 133 (1999)

[7] G. Biroli, R. Monasson, and M. Weigt, Europ. Phys. J. B14, 551 (2000)

[8] F. Ricci-Tersenghi, M. Weigt, and R. Zecchina, to appear in Phys. Rev. E (01 Feb 2001)

[9] S. Mertens, Phys. Rev. Lett 81, 4281 (1998)

[10] M. Mézard, G. Parisi, J. Phys. 47, 1285 (1986)

[11] J. Houdayer, J.H. Boutet de Monvel, and O.C. Martin, Eur. Phys Jour. B6, 383 (1998)

[12] M. Weigt and A. K. Hartmann, Phys. Rev. Lett. 84, 6118 (2000)

[13] S. Cocco and R. Monasson, subm. to Phys. Rev. Lett., also preprint cond-mat/0009410

[14] M. Weigt and A. K. Hartmann, subm. to Phys. Rev. Lett., also preprint cond-mat/0009417

[15] P. Erdös and A. Rényi, Publ. Math. Inst. Hung. Acad. Sci. 5,17 (1960)

[16] B. Bollobas, Random Graphs (Academic Press 1985)

[17] J. Harant, Discr. Math. 188, 239 (1998)

[18] C. Caro, Technical Report, Tel Aviv University (1979)
V. K. Wei, Bell Lab. Technical Memorandum 81-11217-9 (1981)

[19] B. Bollobás and P. Erdös, Math. Proc. Cambridge Philos. Soc. 80, 419 (1976)

[20] P. G. Gazmuri, Networks 14, 367 (1984)

[21] A. M. Frieze, Discr. Math. 81, 171 (1990)

[22] R.J. Baxter, Exactly Solved Models in Statistical Mechanics, (Academic Press London, 1982)

[23] F.M. Russo, Phys. Lett. A239, 17 (1998)

[24] S. Scarpetta, A. de Candidia and A. Coniglio, Phys. Rev. E55, 4943 (1997)

[25] F. Ricci-Tersenghi, D.A. Stariolo, J.J. Arenzon, Phys. Rev. Lett. 84, 4473 (2000), Phys. Rev. E62, 5978 (2000)

[26] M. Nicodemi, A. Coniglio, and H.J. Herrmann, Phys. Rev. E55, 3962 (1997); J. Phys. A30, L379 (1997)

[27] M. Nicodemi J. Phys. I7, 1365 (1998)

[28] E. Caglioti, H.J. Herrmann, V. Loreto, and M. Nicodemi, Phys. Rev. Lett. 79, 1575 (1997)

[29] M. Nicodemi and A. Coniglio, Phys. Rev. Lett. 82, 916 (1999), J. Phys. C12, 6601 (2000)

[30] R. Monasson and O. Pouliquen, Physica A 236, 395 (1997)

[31] A. Barrat and V. Loreto, J. Phys. A33, 4401 (2000)

[32] K. Mehlhorn and St. Näher, The LEDA Platform of Combinatorial and Geometric Computing, (Cambridge University Press, 1999);
see also http://www.mpi-sb.mpg.de/LEDA/leda.html

[33] A.V. Aho, J.E. Hopcroft and J.D. Ullman: The design and analysis of computer algorithms, Addison-Wesley, Readin (MA) 1974

[34] E.L. Lawler and D.E. Wood, Oper. Res. 14, 699 (1966)

[35] R.E. Tarjan and A.E. Trojanowski, SIAM J. Comp. 6, 537 (1977)

[36] M. Shindo and E. Tomita, Syst. Comp. Jap. 21, 1 (1990)

[37] R. Lüling and B. Monien, Sixth International Parallel Processing Symposium, IEEE Comput. Soc. Press, Los Alamitos, USA, 543 (1992)

[38] M. Mézard, G. Parisi, and Virasoro, Spin glasses and beyond, (World Scientific, Singapore 1987)

[39] R. Monasson, J. Phys. A31, 513 (1998)

[40] M. Bauer and O. Golinelli, preprint cond-mat/0006472

[41] C. De Dominicis and P. Mottishaw, J. Phys. A20, L1267 (1987)

[42] P. Mottishaw and C. De Dominicis, J. Phys. A20, L375 (1987)

[43] K.Y.M. Wong and D. Sherrington, J. Phys. A21, L459 (1988)

[44] Y.Y. Goldschmidt and P.Y. Lai, J. Phys. A23, L775 (1990)

[45] M. Mezard and G. Parisi, preprint cond-mat/0009418

[46] In principle, also a scaling $m \sim \mu^{-1}$ can be plugged in. This would be analogous to infinite-connectivity spin-glass models in the zero-temperature limit, and was also considered in the variational approach [7] to 3-satisfiability. The inclusion into the presented formulation leads however to saddle-point equations which could not be solved by the authors.

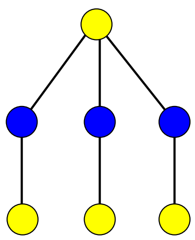

Figure 1: A small sample graph with minimum vertex cover of size 3. The vertices belonging to the minimum V_{vc} are dark. For this graph the heuristic fails to find the true minimum cover, because is starts by covering the root vertex, which has the highest degree 3.

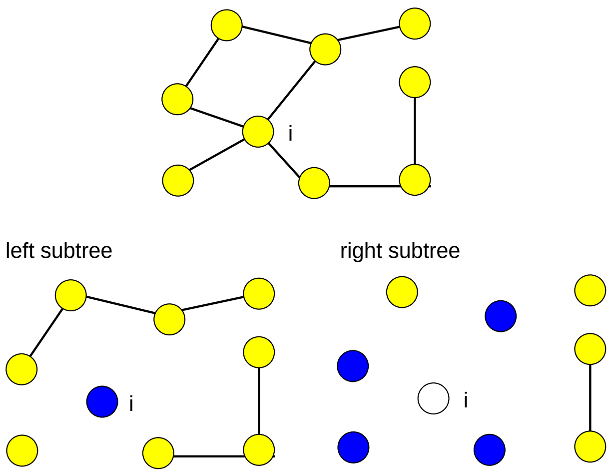

Figure 2: Example how the divide-and-conquer algorithm operates. Above the graph is shown. The vertex i with the highest degree is considered. In the case it is *covered* (left subtree), all incident edges can be removed. In case it is *uncovered*, (right subtree) all neighbors have to be *covered* and all edges incident to the neighbors can be removed. In both case, the graph may split into several components, which can be treated independently by recursive calls of the algorithm.

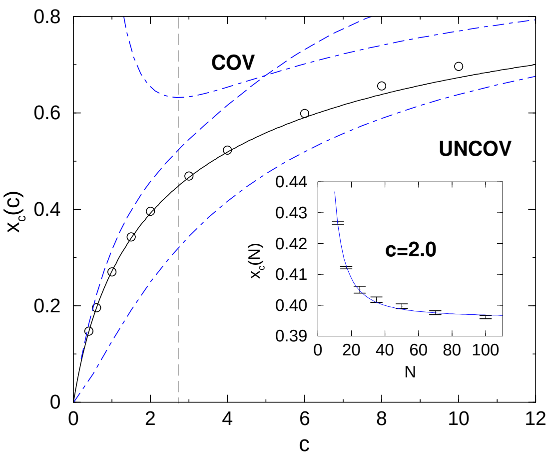

Figure 3: Phase diagram: Fraction $x_c(c)$ of vertices in a minimal vertex cover as function of the average connectivity c. For $x > x_c(c)$, almost all graphs have covers with xN vertices, while they have almost surely no cover for $x < x_c(c)$. The solid line shows the replica-symmetric result. The circles represent the results of the numerical simulations. Error bars are much smaller than symbol sizes. The upper bound of Harant is given by the dashed line, the bounds of Gazmuri by the dash-dotted lines. The vertical line is at $c = e$. Inset: All numerical values were calculated from finite-size scaling fits of $x_c(N, c)$ using functions $x_c(N) = x_c + aN^{-b}$. We show the data for $c = 2.0$ as an example.

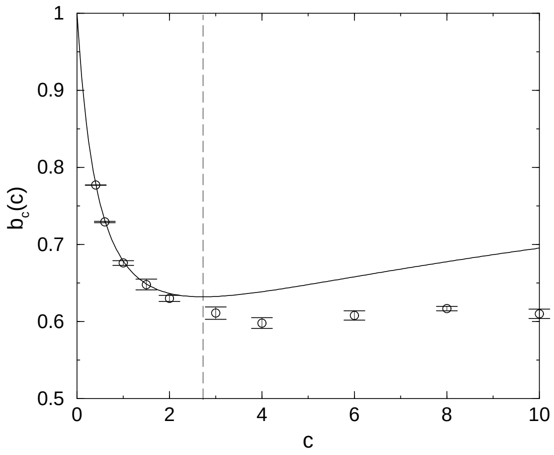

Figure 4: The total backbone size $b_{uncov}(c) + b_{cov}(c)$ of minimal vertex covers as a function of c. The solid line shows the replica-symmetric result, the dotted ones are the two results of one-step RSB. Numerical data are represented by the error bars. They were obtained from finite-size scaling fits similar to the calculation for $x_c(c)$. The vertical line is at $c = e$ where replica symmetry breaks down.

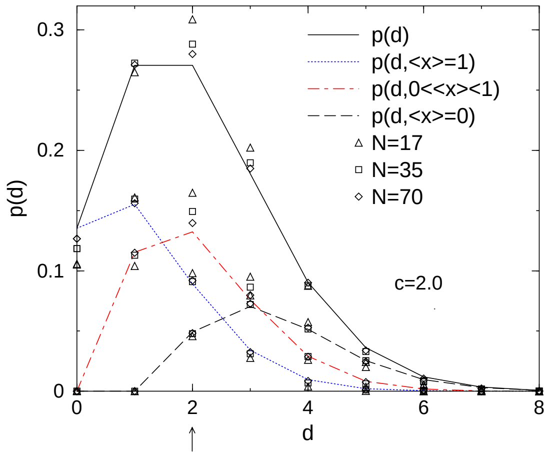

Figure 5: Distribution of connectivities d for $c = 2.0$. We show the total connectivity distribution, given by a Poissonian of mean c, as well as results describing the minimal vertex covers. The total distribution is divided into three contributions arising from the vertices which either are not in the backbone $(0 < \langle x \rangle < 1)$ or which are in the covered/uncovered backbone $(\langle x \rangle = 0/1)$. Analytical predictions are represented by the lines (which are guides to the eyes only, connecting the results for integer arguments), while the numerical results for $N = 17, 35, 70$ are displayed using the symbols.

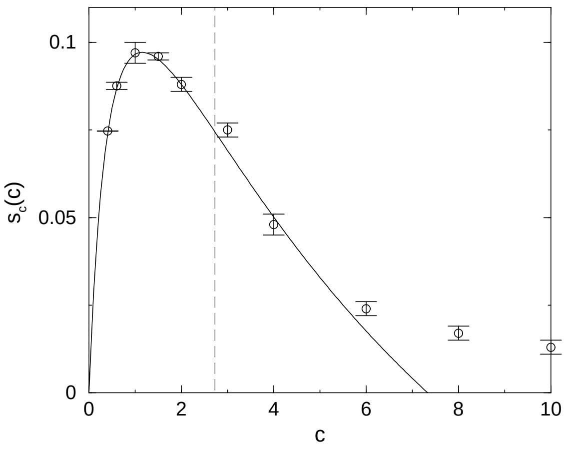

Figure 6: Entropy of minimal vertex as function of the average connectivity c. The solid line results from the Gaussian approximation described in section 6.4 while numerical data are given by the symbols with error bars. Each numerical result was obtained by an extrapolation $N \to \infty$ via fitting a function $s_c(N) = s_c + \alpha N^{-\beta}$ to the data for each c. The vertical line is at $c = e$.

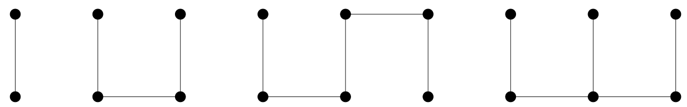

Figure 7: Examples of smallest non-backbone graphs. Note that all graphs can be divided into connected vertex pairs and some supplementary edges connecting different pairs. A similar structure is found also for the full non-backbone subgraph at connectivities $c < e$.

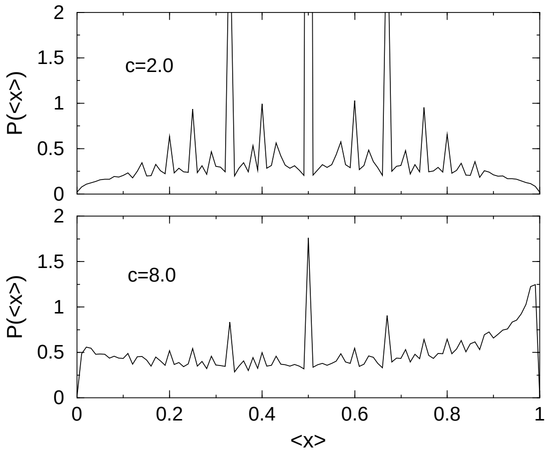

Figure 8: Numerical histograms of local average occupation numbers $\langle x \rangle_{\mu\to\infty}$ of non-backbone vertices for average connectivities $c = 2.0$ and 8.0. The upper distribution is perfectly symmetric as predicted by theory for all $c < e$. The lower one shows an obvious bias towards higher occupation. The effect becomes stronger with increasing connectivity. Please note also the existence of pronounced peaks in both distributions. These result from small non-backbone components or dangling ends of the giant cluster, *e.g.* the peaks at $m = 1/3, 2/3$ appear in chains of four vertices connected by three edges as given by the second graph in the previous figure. The weight of these peaks decreases with increasing size of the giant non-backbone component.

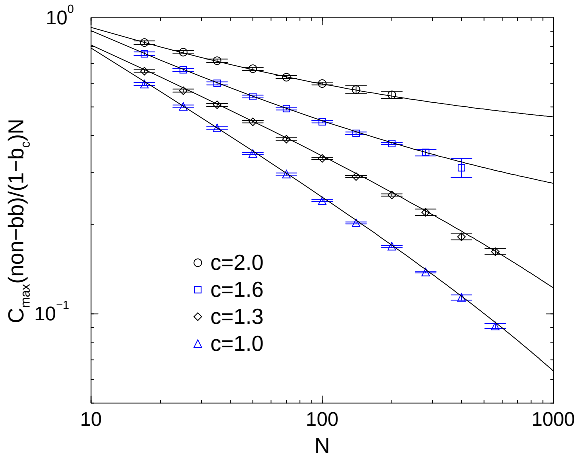

Figure 9: Fraction $f_{\max} = C_{\max}/(1-b_c)N$ of the largest component of the non-backbone subgraph from numerical calculations as a function of graph sizes N up to size $N = 560$. For $b_c(c)$ the numerical values were taken. In a double logarithmic plot, for connectivities smaller than the predicted threshold $c \simeq 1.434$, the function $f_{\max}(N)$ has a negative curvature, indicating that $f_{\max}$ converges towards zero. Thus, for small connectivities, the non-backbone does not percolate. For larger connectivities, $f_{max}(N)$ has a positive curvature, and fits of the form $f(N) = f_\infty + bN^c$ result in stricly positive values, here $f_\infty = 0.17(1)$ $(c = 1.6)$ resp. $f_\infty = 0.37(3)$ $(c = 2.0)$. Hence, the non-backbone percolates.