# On Percolation and $\mathcal{NP}$-Hardness*

**Huck Bennett†[1], Daniel Reichman[2], and Igor Shinkar‡[3]**

1 **Department of Computer Science, Courant Institute of Mathematical Sciences, New York University, New York, USA**
`hbennett@cs.nyu.edu`
2 **Department of Cognitive and Brain Sciences, University of California, Berkeley, CA, USA**
`daniel.reichman@gmail.com`
3 **Department of Computer Science, Courant Institute of Mathematical Sciences, New York University, New York, USA**
`ishinkar@cims.nyu.edu`

## Abstract

The edge-percolation and vertex-percolation random graph models start with an arbitrary graph $G$, and randomly delete edges or vertices of $G$ with some fixed probability. We study the computational hardness of problems whose inputs are obtained by applying percolation to worst-case instances. Specifically, we show that a number of classical $\mathcal{NP}$-hard graph problems remain essentially as hard on percolated instances as they are in the worst-case (assuming $\mathcal{NP} \nsubseteq \mathcal{BPP}$). We also prove hardness results for other $\mathcal{NP}$-hard problems such as Constraint Satisfaction Problems, where random deletions are applied to clauses or variables.

We focus on proving the hardness of the Maximum Independent Set problem and the Graph Coloring problem on percolated instances. To show this we establish the robustness of the corresponding parameters $\alpha(\cdot)$ and $\chi(\cdot)$ to percolation, which may be of independent interest. Given a graph $G$, let $G'$ be the graph obtained by randomly deleting edges of $G$. We show that if $\alpha(G)$ is small, then $\alpha(G')$ remains small with probability at least 0.99. Similarly, we show that if $\chi(G)$ is large, then $\chi(G')$ remains large with probability at least 0.99.

**1998 ACM Subject Classification** F.2.2 Nonnumerical Algorithms and Problems

**Keywords and phrases** percolation, NP-hardness, random subgraphs, chromatic number

**Digital Object Identifier** 10.4230/LIPIcs.ICALP.2016.80

## 1 Introduction

The theory of $\mathcal{NP}$-hardness suggests that we are unlikely to find optimal solutions to $\mathcal{NP}$-hard problems in polynomial time. This theory applies to the worst-case setting where one considers the worst running-time over all inputs of a given length. It is less clear whether these hardness results apply to "real-life" instances. One way to address this question is to examine to what extent known $\mathcal{NP}$-hardness results are stable under random perturbations, as it seems reasonable to assume that a given instance of a problem may be subjected to noise originating from multiple sources.

Recent work has studied the effect of random perturbations of the input on the runtime of algorithms. In their seminal paper Spielman and Teng [28] introduced the idea of *smoothed*

* The full version of this paper is available at `http://eccc.hpi-web.de/report/2015/132`.
† Huck Bennett was supported by NSF grant CCF 1423228.
‡ Igor Shinkar was supported by NSF grants CCF 1422159, 1061938, 0832795, and a Simons Collaboration on Algorithms and Geometry grant.


43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016).
Editors: Ioannis Chatzigiannakis, Michael Mitzenmacher, Yuval Rabani, and Davide Sangiorgi;
Article No. 80; pp. 80:1–80:14
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

*analysis* to explain the superior performance of algorithms in practice compared with formal worst-case bounds. Roughly speaking, smoothed analysis studies the running time of an algorithm on a perturbed worst-case instance. In particular, they showed that subjecting the weights of an arbitrary linear program to Gaussian noise yields instances on which the simplex algorithm runs in expected polynomial time, despite the fact that there are pathological linear programs for which the simplex algorithm requires exponential time. Since then smoothed analysis has been applied to a number of other problems [10, 29].

In contrast to smoothed analysis, we study when worst-case instances of problems remain hard under random perturbations. Specifically, we study to what extent $\mathcal{NP}$-hardness results are robust when instances are subjected to random deletions. Previous work is mainly concerned with *Gaussian* perturbations of *weighted* instances. Less work has examined the robustness of hardness results of unweighted instances with respect to discrete noise.

We focus on two forms of percolation on graphs. Given a graph $G = (V, E)$ and a parameter $p \in (0, 1)$, we define $G_{p,e} = (V, E')$ as the probability space of graphs on the same set of vertices, where each edge $e \in E$ is contained in $E'$ independently with probability $p$. We say that $G_{p,e}$ is obtained from $G$ by *edge percolation*. We define $G_{p,v} = (V', E')$ as the probability space of graphs, in which every vertex $v \in V$ is contained in $V'$ independently with probability $p$, and $G_{p,v}$ is the subgraph of $G$ induced by the vertices $V'$. We say that $G_{p,v}$ is obtained from $G$ by *vertex percolation*. We also study appropriately defined random deletions applied to instances of other $\mathcal{NP}$-hard problems, such as 3-SAT and Subset-Sum.

Throughout we refer to instances that are subjected to random deletions as *percolated instances*. Our main question is whether such percolated instances remain hard to solve by polynomial-time algorithms assuming $\mathcal{NP} \nsubseteq BPP$.

## 1.1 A first example – 3-Coloring

Consider the 3-Coloring problem, where given a graph $G = (V, E)$ we need to decide whether $G$ is 3-colorable. Suppose that given a graph $G$ we sample a random subgraph $G'$ of $G$, by deleting each edge of $G$ independently with probability $p = \frac{1}{2}$, and ask whether the resulting graph is 3-colorable. Is there a polynomial time algorithm that can decide with high probability whether $G'$ is 3-colorable?

We demonstrate that a polynomial-time algorithm for deciding whether $G'$ is 3-colorable is impossible assuming $\mathcal{NP} \nsubseteq \mathcal{BPP}$. We show this by considering the following polynomial time reduction from the 3-Coloring problem to itself.

Given an $n$-vertex graph $H$ the reduction outputs a graph $G$ that is an $R$-blow-up of $H$ for $R = C\sqrt{\log(n)}$, where $C > 0$ is large enough. That is, replace each vertex of $H$ by a cloud of $R$ vertices that form an independent set in $G$, and for each edge in $H$ place a complete $R \times R$ bipartite graph in $G$ between the corresponding clouds in $G$. It is easy to see that $H$ is 3-colorable if and only if $G$ is 3-colorable.

In fact, the foregoing reduction satisfies a stronger *robustness* property for random subgraphs $G'$ of $G$. Namely, if $H$ is 3-colorable, then $G$ is 3-colorable, and hence $G'$ is also 3-colorable with probability 1. On the other hand, if $H$ is not 3-colorable, then $G$ is not 3-colorable, and with high probability $G'$ is not 3-colorable either.

Indeed, for any edge $(v_1, v_2)$ in $H$ let $U_1, U_2$ be two clouds in $G$ corresponding to $v_1$ and $v_2$. Fixing two arbitrary sets $U_1' \subseteq U_1$ and $U_2' \subseteq U_2$ each of size $R/3$, the probability that there is no edge connecting a vertex from $U_1$ to a vertex in $U_2$ is $2^{-R^2/9}$. By union bounding over the $|E| \cdot \binom{R}{R/3}^2 \ll 2^{R^2/9}$ choices of $U_1', U_2'$ we get that there is at least one edge between $U_1'$ and $U_2'$ with high probability. When this holds we can decode any 3-coloring of $G'$ to a 3-coloring of $H$ by coloring each vertex $v$ of $H$ with the color that appears the largest number of times in the coloring of the corresponding cloud in $G'$, breaking ties arbitrarily. Therefore a polynomial time algorithm for deciding the 3-colorability of $G$ implies a polynomial time algorithm for determining the 3-colorability of $H$ with high probability. It follows that unless $\mathcal{NP} \subseteq \text{co}\mathcal{RP}$ there is no polynomial time algorithm that given a 3-colorable graph $G$ finds a 3-coloring of a random subgraph of $G$.[1]

### Toward a stronger notion of robustness

The example above raises the question of whether the blow-up described above is really necessary. Naïvely, one could hope for stronger hardness of the 3-Coloring problem, namely, that for any graph $H$ if $H$ is not 3-colorable, then with high probability a random subgraph $H'$ of $H$ is not 3-colorable either. However, this is not true in general, as $H$ can be a 3-critical graph, i.e., a 3-colorable graph such that deletion of *any* edge of $H$ decreases its chromatic number (consider for example the case of an odd cycle).

Nonetheless, if random deletions do not decrease the chromatic number of a graph by much, then one could use hardness of approximation results for chromatic number to deduce hardness results for coloring percolated graphs. In this paper we show that the chromatic number of a graph is indeed robust to random deletions. We show that if we delete each edge of a graph with probability $\frac{1}{2}$, then (with probability 0.99) the chromatic number does not drop by much.

We also consider the question of robustness for other graph parameters. For independent sets we demonstrate that if the independence number of $G$ is small, then with high probability the independence number of a random subgraph of $G$ is small as well. Similarly, we show that for a $k$-SAT formula that is sufficiently dense, randomly deleting its clauses does not change the maximum possible fraction of clauses that can be satisfied simultaneously. In particular, this implies that these problems remain essentially as hard on percolated instances as they are on worst-case instances.

▶ Remark. It is worth noting that there are graph parameters for which percolated instances differ significantly from the original instance. For example, standard results in random graph theory imply that for every $n$-vertex graph $G$, with high probability the size of the largest clique in the graph $G'$ obtained by edge percolation with $p = \frac{1}{2}$ is $O(\log n)$. In particular, a maximum clique in $G'$ can be found in time $n^{O(\log n)}$, which is significantly faster than the fastest known algorithm for finding a maximum clique in the worst-case.

## 1.2 Robustness of $\mathcal{NP}$-hard problems under percolation

In proving hardness results for percolated instances we use the concept of *robust reductions* which we explain below. It will be convenient to consider promise problems[2]. We start by introducing the following definition.

▶ **Definition 1.** Let $A = (A_{YES}, A_{NO})$ and $B = (B_{YES}, B_{NO})$ be two promise problems. For each $y \in \{0,1\}^*$ (an instance of the problem $B$) let $\mathsf{noise}(y)$ be a distribution on $\{0,1\}^*$, that is samplable in time $\text{poly}(|y|)$.

[1] Note that in the foregoing example, if we start with a bounded degree graph $H$, we can reduce it to a bounded degree graph $G$ by using an $R \times R$ bipartite expander instead of the complete bipartite graph.

[2] Recall, that a promise problem is a generalization of a decision problem, where for the problem $L$ there are two disjoint subsets $L_{YES}$ and $L_{NO}$, such that an algorithm that solves $L$ must accept all the inputs in $L_{YES}$ and reject all inputs in $L_{NO}$. If the input does not belong to $L_{YES} \cup L_{NO}$, there is no requirement on the output of the algorithm.

- A polynomial time reduction $\mathsf{R}$ from $A$ to $B$ is said to be *noise-robust* if
  1. For all $x \in A_{YES}$ it holds that $\mathsf{R}(x) \in B_{YES}$, and $\Pr[\mathsf{noise}(\mathsf{R}(x)) \in B_{YES}] > 0.99$.
  2. For all $x \in A_{NO}$ it holds that $\mathsf{R}(x) \in B_{NO}$, and $\Pr[\mathsf{noise}(\mathsf{R}(x)) \in B_{NO}] > 0.99$.
- If in the first item we have $\Pr[\mathsf{noise}(\mathsf{R}(x)) \in B_{YES}] = 1$, then we say that $\mathsf{R}$ is a noise-robust co$\mathcal{RP}$-reduction. Similarly, if in the second item we have $\Pr[\mathsf{noise}(\mathsf{R}(x)) \in B_{NO}] = 1$, then we say that $\mathsf{R}$ is a noise-robust $\mathcal{RP}$-reduction.
- The problem $B = (B_{YES}, B_{NO})$ is said to be *$\mathcal{NP}$-hard under a noise-robust reduction* if there exists a *noise-robust* reduction from an $\mathcal{NP}$-hard problem to $B$.
- We say that the problem $A$ is *strongly-noise-robust* to $B$ if
  1. For all $x \in A_{YES}$ it holds that $x \in B_{YES}$, and $\Pr[\mathsf{noise}(x) \in B_{YES}] > 0.99$.
  2. For all $x \in A_{NO}$ it holds that $x \in B_{NO}$, and $\Pr[\mathsf{noise}(x) \in B_{NO}] > 0.99$.

Note that in the last item of Definition 1 there is no reduction involved. Instead, we think of the problem $A$ as a relaxation of $B$ with $A_{YES} \subseteq B_{YES}$ and $A_{NO} \subseteq B_{NO}$, and hence any algorithm that solves $B$ in particular solves $A$. However, it is a relaxed problem in a stronger sense, namely, after applying noise to a YES-instance (resp. NO-instance) of $A$, it stays a YES-instance (resp. NO-instance) of $B$ with high probability.

We use the term noise-robust to avoid confusion with other notions of robust reductions that have appeared in the literature. In order to ease readability, we will often write robust reductions instead, always referring to noise-robust reductions as defined above.

▶ **Proposition 2.** *Let $L = (L_{YES}, L_{NO})$ be a promise problem, and for each $y$ instance of $L$, let noise$(y)$ be a distribution on instances of $L$ that is samplable in time* $\mathrm{poly}(|y|)$*.*

*If $L$ is $\mathcal{NP}$-hard under a noise-robust reduction, then there is no polynomial time algorithm that when given an input $y$ decides with high probability whether noise$(y) \in L_{YES}$ or $noise(y) \in L_{NO}$, unless $\mathcal{NP} \subseteq \mathcal{BPP}$.*

Indeed, the example given in Section 1.1 gives a noise-robust reduction from the 3-Coloring problem to itself, where noise refers to random deletions of the edges in a given graph. Therefore, the 3-Coloring problem is $\mathcal{NP}$-hard under a noise-robust reduction.

## 1.3 Our results

In this paper we show that a number of $\mathcal{NP}$-hard problems remain hard to solve even after random deletions, i.e., they are $\mathcal{NP}$-hard under noise-robust reductions. Furthermore, we show that some gap $\mathcal{NP}$-hard problems are, in fact, strongly-noise-robust to the same problems with a smaller gap. Specifically, we focus on showing these results for the gap versions of the maximum independent set and chromatic number problems. As technical tools, we prove a number of combinatorial results about the independence number and the chromatic number of percolated graphs that might be of independent interest.

### Maximum Independent Set and Percolation

▶ **Theorem 3.** *Let $G = (V, E)$ be an $n$-vertex graph. Then, with high probability $\alpha(G_{p,e}) \leq O\left(\frac{\alpha(G)}{p}\log(np)\right)$.*

We observe that in general, the upper bound above cannot be improved, as it is well known that the independence number of $G(n, p)$ is $\Omega\left(\frac{\log(np)}{p}\right)$ with high probability (see, e.g., [4]).

In the Coloring-vs-MIS$(q, a)$ problem, given an $n$-vertex graph $G$ such that $q \cdot a \geq n$, the goal is to distinguish between the YES-case where $\chi(G) \leq q$ and the NO-case where

$\alpha(G) \leq a$. By using Theorem 3 together with the inapproximability results of Feige and Kilian [11] saying that for every $\varepsilon > 0$ it is $\mathcal{NP}$-hard to decide whether a given $n$-vertex graph $G$ satisfies $\chi(G) \leq n^{\varepsilon}$ or $\alpha(G) \leq n^{\varepsilon}$ we obtain the following hardness result.

▶ **Theorem 4.** *For any $q, a$ the Coloring-vs-MIS$(q, a)$ problem is strongly-noise-robust to Coloring-vs-MIS$(q, O\left(\frac{a}{p}\log(np)\right))$, where $n$ denotes the number of vertices in the given graph, and noise is the $p$-edge-percolation of this graph.*

*In particular, for any constant $\varepsilon > 0$, unless $\mathcal{NP} \subseteq \mathcal{BPP}$ there is no polynomial time algorithm that given an $n$-vertex graph $G$ approximates either $\alpha(G_{p,e})$ or $\chi(G_{p,e})$ within a $\frac{1}{pn^{1-2\varepsilon}}$ (resp. $pn^{1-2\varepsilon}$) factor for any $p > \frac{1}{n^{1-2\varepsilon}}$.*

We also prove analogous theorems for vertex percolation.

## Graph Coloring and Percolation

Theorem 3 says that it is hard to approximate the chromatic number of a percolated graph within a $n^{1-\varepsilon}$ factor, but says nothing about hardness of coloring percolated graphs with small (constant) chromatic number. We address this question below by proving lower bounds[3] on the chromatic number of percolated graphs. To do this we use results from additive combinatorics and discrete Fourier analysis.

▶ **Theorem 5.** *Let $G = (V, E)$ be an $n$-vertex graph. Then, for every $\alpha \in (0, 1)$ it holds that* $\Pr[\chi(G_{\frac{1}{2},v}) \geq \max\{\chi(G)/3 - O_{\alpha}(1), \chi(G)/2 - O_{\alpha}(\sqrt{n})\}] > 1 - \alpha$.

▶ **Theorem 6.** *Let $G = (V, E)$ be an $n$-vertex graph with $m$ edges. Then, for every $\alpha \in (0, 1)$ it holds that* $\Pr[\chi(G_{\frac{1}{2},e}) \geq \max\{\Omega_{\alpha}(\chi(G)^{1/3}), \Omega_{\alpha}(\chi(G)/m^{1/4})\}] > 1 - \alpha$.

For $G_{\frac{1}{2},v}$ the $\chi(G)/2 - O_{\alpha}(\sqrt{n})$ lower bound is better when $\chi(G) = \omega(\sqrt{n})$, and the $\chi(G)/3 - O_{\alpha}(1)$ lower bound is better when $\chi(G) = o(\sqrt{n})$. For $G_{\frac{1}{2},e}$ the $\Omega_{\alpha}(\chi(G)/m^{1/4})$ lower bound is better when $\chi(G) = \omega(m^{3/8})$, and the $\Omega_{\alpha}(\chi(G)^{1/3})$ lower bound is better when $\chi(G) = o(m^{3/8})$.

Note that this result also gives lower bounds on the chromatic number of $G_{p,v}, G_{p,e}$ where $p \neq \frac{1}{2}$ by composing the bounds in Theorems 5 and 6 $\lceil \log_2(1/p) \rceil$ times.

▶ Remark. Bukh [7] has considered coloring edge-percolated graphs, and states the question of whether $\mathbf{E}[\chi(G_{\frac{1}{2},e})] = \Omega(\chi(G)/\log(\chi(G)))$ as an "interesting problem." Bukh observed that the chromatic number of $G_{\frac{1}{2},e}$ has the same distribution as the chromatic number of the complement of $G_{\frac{1}{2},e}$, and therefore $\mathbf{E}[\chi(G_{\frac{1}{2},e})] \geq \sqrt{\chi(G)}$. However, it is not clear how to leverage the lower bound on the expectation to obtain a lower bound on $\chi(G_{\frac{1}{2},e})$ *with high probability*, which is required for our noise robust reductions. Moreover, for $k \ll \sqrt{n}$ standard martingale methods do not seem to work for showing high probability estimates.

In the Gap-Coloring$(q, Q)$ problem we are given an $n$-vertex graph $G$ and the goal is to distinguish between the YES-case where $G$ is $q$-colorable, and the NO-case where the chromatic number of $G$ is at least $Q$. There is a large body of work proving hardness results for this problem [14, 20, 18] including stronger results assuming variants of the Unique Games Conjecture [8, 9]. Using the $\mathcal{NP}$-hardness of the Gap-Coloring$(q, \exp(\Omega(q^{1/3})))$ problem of Huang [18] we obtain an analogous hardness result under noise-robust reductions for this problem.

[3] The notation $O_{\alpha}(f(n))$ means that $O(f(n))$ holds for fixed $\alpha$.

▶ **Theorem 7.** *For all $q < Q$ the Gap-Coloring$(q, Q)$ problem is strongly-noise-robust to the Gap-Coloring$(q, \Omega(Q^{1/3}))$ problem, where noise is $\frac{1}{2}$-edge-percolation applied to the graph.*

*In particular, for any sufficiently large constant $q$ given a $q$-colorable graph $G$ it is $\mathcal{NP}$-hard to find a $2^{\Omega(q^{1/3})}$-coloring of $G_{\frac{1}{2},e}$.*

### Satisfiability and Other Problems

We also state a hardness result for approximating the value of a clause-percolated instance of $k$-SAT. A $k$-SAT formula $\Phi$ is a collection of $m$ clauses on $n$ Boolean variables, where each clause is an OR of $k$-literals. Given a formula $\Phi$, and an assignment $\sigma$ to its variables, denote by $val_\sigma(\Phi)$ the fraction of constraints of $\Phi$ satisfied by $\sigma$. The value of $\Phi$ is defined as $val(\Phi) = \max_\sigma val_\sigma(\Phi)$. If $val(\Phi) = 1$ we say that $\Phi$ is satisfiable.

Given an instance $\Phi$ of $k$-SAT its *clause percolation* is a random formula $\Phi_p^c$ over the same set of variables, obtained from $\Phi$ by keeping each clause of $\Phi$ independently with probability $p$.

▶ **Theorem 8.** *Let $\varepsilon, \delta \in (0, 1)$ be fixed constants. Then, unless $\mathcal{NP} \subseteq \mathrm{co}\mathcal{RP}$, there is no polynomial time algorithm that when given a satisfiable instance $\Phi$ over $n$-variables of 3-SAT, finds an assignment $\sigma$ to $\Phi_p^c$ such that $val_\sigma(\Phi_p^c) > 7/8 + \varepsilon$ for all $p > \frac{1}{n^{2-\delta}}$.*

One ingredient of the proof of Theorem 8, that may be of independent interest, is establishing that $k$-SAT does not admit a non-trivial approximation on dense formulas that contain $n^{k-\eta}$ clauses, where $\eta > 0$ is an arbitrary small positive constant.

We prove analogous theorems also for other CSP's as well as other graph theoretic problems such as Vertex Cover and Directed Hamiltonian Cycle. We also prove hardness results for the percolated Subset Sum problem. The exact statements and complete proofs, including of Theorem 8, appear in the full version of the paper.

## 1.4 Preliminaries

An *independent set* in a graph $G = (V, E)$ is a set of vertices that spans no edges. The *independence number* $\alpha(G)$ denotes the maximum size of an independent set in $G$. A *legal coloring* of a graph $G$ is an assignment of colors to vertices such that no two adjacent vertices share the same color. The *chromatic number* $\chi(G)$ denotes the minimum number of colors needed for a legal coloring of $G$. Note that in a legal coloring of $G$ each color class forms an independent set, and hence $\chi(G) \cdot \alpha(G) \geq n$.

We will always measure the running time of algorithms in terms of the size of the percolated instance. Since $G$ and $G_{p,e}$ have the same number of vertices, this generally does not affect the size of the instance by more than a polynomial factor. On the other hand, $G_{p,v}$ may be much smaller than $G$ for very small values of $p$. However, in this work we will be only dealing with the case where $p = \frac{1}{n^{1-\Omega(1)}}$, hence with high probability the size of the vertex percolated and original graphs are polynomially related as well.

We will use the following version of the Chernoff bound.

▶ **Lemma 9** (Chernoff bound, Theorem 7.3.2 in [17]). *Let $x_1, \ldots, x_n$ be independent Bernoulli trials with $\Pr[x_i = 1] = p$, and let $\mu = \mathbf{E}[\sum_{i=1}^n x_i] = pn$. Let $r \geq e^2$. Then $\Pr[\sum_{i=1}^n x_i > (1 + r)\mu] < \exp(-(\mu r/2) \ln r)$.*

## 1.5 Related Work

There is a wide body of work on random discrete structures that has produced a wide range of mathematical tools [4, 13, 15, 23]. Randomly subsampling subgraphs by including each edge independently in the sample with probability $p$ has been studied extensively in works concerned with cuts and flows (e.g., [19]). More recently, sampling subgraphs has been used to find independent sets [12]. The effect of subsampling variables in mathematical relaxations of constraint satisfaction problems on the value of these relaxations was studied in [2].

Edge-percolated graphs have been also used to construct hard instances for specific algorithms. For example, Kučera [21] proved that the well known greedy coloring algorithm performs poorly on the complete $r$-partite graph in which every edge is removed independently with probability $1/2$ and $r = n^{\varepsilon}$ for $\varepsilon > 0$. Namely, for this graph $G$, even if vertices are considered in a random order by the greedy algorithm, with high probability $\Omega(\frac{n}{\log n})$ colors are used to color the percolated graph whereas $\chi(G) \leq n^{\varepsilon}$.

Misra [24] studies edge percolated instances of the Max-Cut problem. He proves that in graphs of fixed maximal degree $d$ it is impossible (assuming $\mathcal{NP} \neq \mathcal{BPP}$) to compute the size of the maximum cut in $G_{p,e}$ in polynomial time whenever $p = \frac{1+\varepsilon}{d-1}$. The techniques used in [24] differ from ours and rely on the recent hardness result for counting independent sets in sparse graphs [27].

The chromatic number of Erdős-Rényi random graphs $G(n,p)$ has been studied extensively. Grimmett and McDiarmid [16] showed that for a fixed $p$ with high probability it holds that $\chi(G(n,p)) = \Theta(\log(1/1-p)\frac{n}{\log n})$. Bollobás [3] later determined the right constant, proving that $\chi(G(n,p)) \sim \log(1/(1-p))\frac{n}{2\log(n)}$ for every $p \in (0,1)$. Łuczak [22] further improved the previous result by showing that it holds for subconstant values of $p$. In this paper we study the independence number and chromatic number of general percolated graphs. A recent paper by Bollobás et al. [5] studied a special case of this, namely the independence number of edge percolated Kneser graphs.

## 2 Maximum Independent Set and Percolation

In this section we demonstrate the hardness of approximating $\alpha(G)$ and $\chi(G)$ in both edge percolated and vertex percolated graphs. We base our results on a theorem of Feige and Kilian, saying that for every fixed $\varepsilon > 0$ the problem Coloring-vs-MIS$(n^{\varepsilon}, n^{\varepsilon})$ is $\mathcal{NP}$-hard.

▶ **Theorem 10** ([11]). *For every $\varepsilon > 0$ it is $\mathcal{NP}$-hard to decide whether a given $n$-vertex graph $G$ satisfies $\chi(G) \leq n^{\varepsilon}$ or $\alpha(G) \leq n^{\varepsilon}$.*

### Edge percolation

Below we prove Theorem 3. We will use the following lemma, due to Turan (see, e.g. [1]).

▶ **Lemma 11.** *Every graph $H$ with $l$ vertices and $e$ edges contains an independent set of size at least $\frac{l^2}{2e+l}$.*

As a corollary we observe that if a graph contains no large independent sets, then it can also cannot contain large subsets of the vertices that span a small number of edges.

▶ **Corollary 12.** *Let $G = (V,E)$ be an $n$-vertex graph satisfying $\alpha(G) < k$. Then every set of vertices of size $l \geq k$ spans at least $l(l-k)/2k$ edges.*

**Proof.** Let $H$ be a subgraph of $G$ induced by $l$ vertices, and suppose that $H$ spans $e$ edges. Then, by Lemma 11 we have $\alpha(H) \geq \frac{l^2}{2e+l}$. On the other hand, $\alpha(H) \leq \alpha(G) \leq k$, and hence $\frac{l^2}{2e+l} \leq k$, as required. ◀

We are now ready to prove Theorem 3 saying that for any $n$-vertex graph $G = (V,E)$ it holds that with high probability $\alpha(G_{p,e}) \leq O\left(\frac{\alpha(G)}{p}\log(np)\right)$.

**Proof of Theorem 3.** For a given graph $G$, let $k = \alpha(G)+1$ and let $C > 0$ be a large enough constant. By Corollary 12, every subset of size $l = C\frac{\alpha(G)}{p}\log(np)$ spans at least $\frac{l(l-k)}{2k}$ edges in $G$. Hence, by taking union bound over all subsets of size $l$, the probability there exists a set of size $l$ in $G_{p,e}$ that spans no edge is at most

$$\binom{n}{l} \cdot (1-p)^{\frac{l(l-k)}{2k}} < \left(\frac{en}{l}\right)^l \cdot \exp\left(-p \cdot \frac{l(l-k)}{2k}\right) < (np)^{-\Omega(l)},$$

where the last inequality uses the choices of $l$ and $k$, implying that $\left(\frac{en}{l}\right)^l < (np)^l$ and $\exp(-p\frac{l(l-k)}{2k}) < \exp(-\Omega(l \cdot \log(np))) = (np)^{-\Omega(l)}$. Therefore, with high probability $\alpha(G_{p,e}) \leq C\frac{\alpha(G)}{p}\log(np)$. ◀

Theorem 4 follows immediately from Theorem 3.

**Proof of Theorem 4.** Let $G$ be an instance $G$ of the Coloring-vs-MIS$(q,a)$ problem. Note that for the YES-case if $\chi(G) \leq q$, then clearly $\chi(G_{p,e}) < q$. For the NO-case by Theorem 3 if $\alpha(G) \leq a$, then with high probability $\alpha(G_{p,e}) \leq O\left(\frac{a}{p}\log(np)\right)$ which implies the strongly-noise-robust hardness.

The "in particular" part follows immediately from Theorem 10. ◀

▶ Remark. Note that for constant $p > 0$ (e.g., $p = 1/2$) this theorem establishes inapproximability for the independence number of $G_{p,e}$ that matches the inapproximability for the worst case.

▶ Remark. Note also that for $p > \frac{1}{n^{1-\varepsilon}}$ (in fact, for $p > \frac{\log(n)}{n}$) such random percolated graphs have maximal degree at most $O(pn)$ with high probability. Therefore, such graphs $G_{p,e}$ can be colored efficiently using $O(pn)$ colors. In particular, with high probability $G_{p,e}$ contains an independent set of size $\Omega(1/p)$ and hence, the independence number can be approximated within a factor of $1/pn$ on $p$-percolated instances.

## Vertex percolation

Next we handle vertex percolation. We show that approximating $\alpha(G)$ and $\chi(G)$ on percolated instances is essentially as hard as worst-case instances, where the vertices remain with probability $p > \frac{1}{n^{1-\delta}}$, where $n$ is the number of vertices in the graph for any constant $\delta \in (0,1)$. We do it again by relying on the hardness of the gap problem Coloring-vs-MIS for percolated instances.

Note that in the case of vertex percolation, the (in)approximability guarantee should depend on the number of vertices in the percolated graph $G_{p,v}$, and not on the number in the original graph.

▶ **Theorem 13.** *The Coloring-vs-MIS$(q,a)$ problem is strongly-noise-robust to itself, where noise is the vertex percolation with parameter any $p > 0$.*

*In particular, for any $\delta, \varepsilon > 0$ unless $\mathcal{NP} \subseteq \mathcal{BPP}$ there is no polynomial time algorithm that approximates either $\alpha(G_{p,v})$ or $\chi(G_{p,v})$ within a factor $m^{1-\varepsilon}$ for constant any $\varepsilon > 0$, where $m$ denotes the number of vertices in $G_{p,v}$, and any $p > \frac{1}{n^{1-\delta}}$.*

**Proof.** The strong robustness of Coloring-vs-MIS$(q,a)$ is clear, since for any graph $G$ if $G'$ is a vertex induced subgraph of $G$, then $\chi(G') \leq \chi(G)$, and $\alpha(G') \leq \alpha(G)$, which is, in particular, true for $G' = G_{p,v}$.

For the "in particular" part, for a given $p > \frac{1}{n^{1-\delta}}$ let $c = \frac{\log(pn)}{\log(n)} \in (\delta, 1)$ so that $p = \frac{1}{n^{1-c}}$, and let $\eta = \varepsilon \cdot c$.

Let $G$ be an $n$-vertex graph, let $G_{p,v}$ be it vertex-percolated subgraph, and let $m$ be the number of vertices in $G_{p,v}$. By the Chernoff bound in Lemma [17], with high probability we have $|m - pn| < 0.1pn$, and so, we assume from now on that $n^\eta < 2m^\varepsilon$.

By Theorem 10 it is $\mathcal{NP}$-hard to decide whether a given $n$-vertex graph $G$ satisfies $\chi(G) \le n^\eta$ or $\alpha(G) \le n^\eta$. By the choice of parameters, if $\chi(G) \le n^\eta$ then $\chi(G_{p,v}) \le n^\eta < 2m^\varepsilon$, and similarly, if $\alpha(G) \le n^\eta$ then $\alpha(G_{p,v}) < n^\eta < 2m^\varepsilon$. This completes the proof of the theorem. ◀

## 3 Graph Coloring and Percolation

We present our results in terms of the *maximum coverage problem* [30], which is a variant of the set cover problem, and show later how graph coloring is related to maximum coverage.

### 3.1 Maximum Coverage

In the maximum coverage problem we are given a family of sets $\mathcal{F} = \{S_1, \dots, S_m\}$ with $S_i \subseteq [n]$ and a number $c$. The goal is to find $c$ sets in $\mathcal{F}$ such the cardinality of the union of these $c$ sets is as large as possible. We will make use of the representation of a set $S$ in terms of its incidence vector $x(S) \in \{0,1\}^n$. In this way, we can reformulate the maximum coverage problem as follows. Given $A \subseteq \mathbb{F}_2^n$, find elements $y_1, \dots, y_c \in A$ that maximize $\|\vee_{i=1}^c y_i\|_1$, the Hamming weight of the bitwise-OR of the vectors.

We will prove two existential results saying that if $A$ is of constant density $\alpha > 0$, then there exists a good cover using only 2 or 3 vectors.

▶ **Lemma 14.** *Let $A \subseteq \mathbb{F}_2^n$ with $|A| \ge \alpha 2^n$. Then there exist $y_1, y_2, y_3 \in A$ such that $\|y_1 \vee y_2 \vee y_3\|_1 \ge n - 4/\alpha^3$.*

▶ **Lemma 15.** *Let $A \subseteq \mathbb{F}_2^n$ with $|A| \ge \alpha 2^n$. Then there exist $y_1, y_2 \in A$ such that $\|y_1 \vee y_2\|_1 \ge n - (1 + r)\sqrt{n}$, where $r = \max\{e^2, 2\ln 1/\alpha\}$.*

### 3.2 Proof of Lemma 14 using additive combinatorics

Lemma 14 follows almost immediately from a result about sumsets. Recall that the Minkowski sum of two sets $A, B$ is defined as $A + B = \{x + y : x \in A, y \in B\}$.

▶ **Lemma 16** (Corollary 3.5 in [26]). *Let $A \subseteq \mathbb{F}_2^n$ with $|A| \ge \alpha 2^n$. Then $A + A + A$ contains an affine subspace of dimension at least $n - 4/\alpha^3$.*

Because an affine subspace of dimension at least $n - 4/\alpha^3$ must contain an element of Hamming weight at least $n - 4/\alpha^3$, Lemma 14 follows from Lemma 16 and the observation that $\|\sum_{i=1}^c y_i\|_1 \le \|\bigvee_{i=1}^c y_i\|_1$.

### 3.3 Proof of Lemma 15 using Fourier analysis

We use an inequality from Fourier analysis to give a proof of Lemma 15 via the probabilistic method.

▶ **Definition 17.** Given $x \in \mathbb{F}_2^n$, define $y \sim N_\rho(x)$ by letting each $y_i$ be equal to $x_i$ with probability $\frac{1+\rho}{2}$, and be equal to $1 - x_i$ with probability $\frac{1-\rho}{2}$.

Let $\mathrm{Uni}(S)$ denote the uniform distribution on a set $S$, and let $U_n$ denote $\mathrm{Uni}(\mathbb{F}_2^n)$. The following lemma is a corollary of the reverse Bonami-Beckner inequality.

▶ **Lemma 18** (Corollary 3.5 in [25])**.** *Let $A, B \subseteq \mathbb{F}_2^n$ with $|A| = |B| = \alpha 2^n$. Then*

$$\Pr_{\substack{x \leftarrow Uni(A) \\ y \leftarrow N_\rho(x)}}[y \in B] \geq \alpha^{(1+\rho)/(1-\rho)}.$$

**Proof of Lemma 15.** Let $A \subseteq \mathbb{F}_2^n$ with $|A| \geq \alpha 2^n$, and let $B = A + \vec{1} = \{x + \vec{1} : x \in A\}$, where $\vec{1}$ is the $n$-dimensional all 1s vector. Note that to prove Lemma 15 it suffices to show that there exist $x \in A, y \in B$ such that $\|x + y\|_1 = (1 + r) \cdot \sqrt{n}$, since then $y + \vec{1} \in A$ and $\|x + (y + \vec{1})\|_1 = n - (1 + r) \cdot \sqrt{n}$.

Let $\varepsilon = 1/\sqrt{n}$ and let $\rho = 1 - 2\varepsilon$. By Lemma 18,

$$\Pr_{\substack{x \leftarrow U_n \\ y \leftarrow N_\rho(x)}}[x \in A, y \in B] = \Pr_{\substack{x \leftarrow \mathrm{Uni}(A) \\ y \leftarrow N_\rho(x)}}[y \in B] \cdot \Pr_{x \leftarrow U_n}[x \in A] \geq \alpha^{2/(1-\rho)} = \alpha^{\sqrt{n}}. \tag{1}$$

Set $r = \max\{e^2, 2\ln(1/\alpha)\}$. Note that by definition of $y \sim N_\rho(x)$ we have that $\Pr[x_i \neq y_i] = 1/\sqrt{n}$ for each $i$ independently. Therefore, by the Chernoff bound in Lemma 9,

$$\Pr_{\substack{x \leftarrow U_n \\ y \leftarrow N_\rho(x)}}[\|x + y\|_1 \leq (1 + r)\sqrt{n}] \geq 1 - e^{-(r/2\ln r)\sqrt{n}} \geq 1 - \alpha^{2\sqrt{n}}. \tag{2}$$

Since the sum of the probabilities in Equations (1) and (2) is strictly greater than 1, the corresponding events cannot be disjoint. Hence there exist $x \in A, y \in B$ such that $\|x + y\|_1 \leq (1 + r)\sqrt{n}$. ◀

## 3.4 Coloring Using Subgraphs

We now show how to apply the results in the previous subsection to the graph coloring problem. Throughout this section we let $G = (V, E)$ with $n = |V|, m = |E|$. We will identify the elements of $[n]$ with vertices $V$ in the vertex percolation case and the elements of $[m]$ with edges $E$ in the edge percolation case. Let $G_{|U}$ denote the subgraph of $G$ induced by $U \subseteq V$.

▶ **Lemma 19.** *Let $G = (V, E)$ and let $V_1, V_2 \subseteq V$ with $V_1 \cup V_2 = V$. If $\chi(G_{|V_1}) \leq k_1$ and $\chi(G_{|V_2}) \leq k_2$ then $\chi(G) \leq k_1 + k_2$.*

**Proof.** Assume that $V_1 \cap V_2 = \emptyset$ (if not, replace $V_1$ with $V_1 \setminus V_2$ in the following argument). Color $G_{|V_1}$ with $k_1$ colors and color $G_{|V_2}$ with $k_2$ fresh colors. Because $G_{|V_1}$ and $G_{|V_2}$ are colored with separate colors any edges between $V_1$ and $V_2$ have endpoints with distinct colors. ◀

▶ **Lemma 20.** *Let $G = (V, E)$, let $E_1, E_2 \subseteq E$ with $E_1 \cup E_2 = E$, and let $G_1 = (V, E_1), G_2 = (V, E_2)$. If $\chi(G_1) \leq k_1$ and $\chi(G_2) \leq k_2$ then $\chi(G) \leq k_1 k_2$.*

**Proof.** Let $c_1$ be a coloring of $G_1$ with $k_1$ colors, and let $c_2$ be the coloring of $G_2$ with $k_2$ colors. We claim that the coloring as $c(v) = (c_1(v), c_2(v))$ is a legal coloring of $G$ with $k_1 k_2$ colors. Consider an edge $e = (u, v) \in E$. If $e \in E_1$ then $c(u)$ differs from $c(v)$ in the fist coordinate. Otherwise $e \in E_2$ in which case $c(u)$ differs from $c(v)$ in the second coordinate. ◀

## 3.5 Lower Bounding the Chromatic Number

We now prove lower bounds on the chromatic number of percolated graphs. We will consider both vertex and edge percolation with $p = \frac{1}{2}$. This choice of $p$ is important because $G_{\frac{1}{2},v}$,

$G_{\frac{1}{2},e}$ become the distributions of graphs induced by uniformly random subsets of $V$ and $E$, respectively. However, we also obtain bounds for $p < \frac{1}{2}$ by composing the bounds for $p = \frac{1}{2}$. When stating bounds based on Lemma 15 we set $r = \max\{e^2, 2\ln(1/\alpha)\}$.

The idea will be to argue that if many subgraphs of a graph $G$ are $k$-colorable then $G$ is colorable with $f(k)$ colors for relatively small $f(k)$. To see how this idea works, consider the following easy case. Suppose that $\Pr[\chi(G_{\frac{1}{2},v}) \le k] > \frac{1}{2}$. Then there exists $V' \subseteq V$ such that $G_{|V'}$ and $G_{|\overline{V'}}$ are both $k$-colorable. It follows that $G$ is $2k$-colorable by Lemma 19. We now consider the case where the density of $k$ colorable subgraphs $\alpha$ is less than $\frac{1}{2}$.

### Proof of Theorem 5

We now prove Theorem 5, saying that if $G$ is an $n$-vertex graph, then for every $\alpha \in (0,1)$ it holds that $\Pr[\chi(G_{\frac{1}{2},v}) \ge \max\{\chi(G)/3 - O_\alpha(1), \chi(G)/2 - O_\alpha(\sqrt{n})\}] > 1 - \alpha$. The proof relies on the following two lemmas.

▶ **Lemma 21.** $\Pr[\chi(G_{\frac{1}{2},v}) \le k] \ge \alpha \Rightarrow \chi(G) \le 3k + 4/\alpha^3$.

**Proof.** Identify subsets of vertices $V$ with vectors in $\mathbb{F}_2^n$. Because $\Pr[\chi(G_{\frac{1}{2},v}) \le k] \ge \alpha$ by Lemma 14 there exist $V_1, V_2, V_3 \subseteq V$ such that each $G_{|V_i}$ is $k$-colorable and $|V_1 \cup V_2 \cup V_3| \ge n - 4/\alpha^3$. Using Lemma 19, we can then color $G_{|V_1 \cup V_2 \cup V_3}$ with $3k$ colors. Coloring the remaining $4/\alpha^3$ nodes each with a different, new color implies the lemma. ◀

▶ **Lemma 22.** $\Pr[\chi(G_{\frac{1}{2},v}) \le k] \ge \alpha \Rightarrow \chi(G) \le 2k + (1+r)\sqrt{n}$.

**Proof.** Identify subsets of vertices $V$ with vectors in $\mathbb{F}_2^n$. Because $\Pr[\chi(G_{\frac{1}{2},v}) \le k] \ge \alpha$ by Lemma 15 there exist $V_1, V_2 \subseteq V$ such that $G_{|V_1}, G_{|V_2}$ are $k$-colorable and $|V_1 \cup V_2| \ge n - (1+r)\sqrt{n}$. Using Lemma 19, we can then color $G_{|V_1 \cup V_2}$ with $2k$ colors. Coloring the remaining $(1+r)\sqrt{n}$ nodes each with a different, new color implies the lemma. ◀

Lemmas 21 and 22 imply Theorem 5. Taking the contrapositive of Lemma 21 we get

$$\chi(G) > 3k + 4/\alpha^3 \Rightarrow \Pr[\chi(G_{\frac{1}{2},v}) > k] \ge 1 - \alpha,$$

which is equivalent to

$$\chi(G) > \ell \Rightarrow \Pr[\chi(G_{\frac{1}{2},v}) > (\ell - 4/\alpha^3)/3] \ge 1 - \alpha.$$

Similarly, taking the contrapositive of Lemma 22 we get the bound

$$\chi(G) > \ell \Rightarrow \Pr[\chi(G_{\frac{1}{2},v}) > \ell/2 - O_\alpha(\sqrt{n})] \ge 1 - \alpha.$$

### Proof of Theorem 6

Next we prove Theorem 6, saying that if $G$ is an $n$-vertex graph with $m$ edges, then for every $\alpha \in (0,1)$ it holds that $\Pr[\chi(G_{\frac{1}{2},e}) \ge \max\{\Omega_\alpha(\chi(G)^{1/3}), \Omega_\alpha(\chi(G)/m^{1/4})\}] > 1 - \alpha$. The techniques are similar to those used for proving Theorem 5, but with several additional observations we push the techniques further.

▶ **Lemma 23.** $\Pr[\chi(G_{\frac{1}{2},e}) \le k] \ge \alpha \Rightarrow \chi(G) \le k^3 + 8/\alpha^3$.

**Proof.** Identify subsets of edges $E$ with vectors in $\mathbb{F}_2^m$. Because $\Pr[\chi(G_{\frac{1}{2},e}) \le k] \ge \alpha$ by Lemma 14 there exist $E_1, E_2, E_3 \subseteq E$ such that each $G_i = (V, E_i)$ is $k$-colorable and $|E_1 \cup E_2 \cup E_3| \ge m - 4/\alpha^3$. Using Lemma 20, we can then color $G(V, E_1 \cup E_2 \cup E_3)$ with $k^3$ colors. Color the endpoints of the remaining $E \setminus (E_1 \cup E_2 \cup E_3)$ edges using $8/\alpha^3$ new colors to achieve a $(k^3 + 8/\alpha^3)$-coloring of $G$. ◀

The next lemma gives an unconditional upper bound on chromatic number.

▶ **Lemma 24.** *Let $G = (V, E)$ be a graph with $|E| = m$. Then $\chi(G) \leq 3\sqrt{m} + 1$.*

**Proof.** Partition $V$ into sets $V_0 = \{v \in V : \deg(v) < \sqrt{m}\}$ and $V_1 = \{v \in V : \deg(v) \geq \sqrt{m}\}$. By Brooks' Theorem [6], $\chi(G_{|V_0}) \leq \max_{v \in V_0} \deg(v) + 1 \leq \sqrt{m} + 1$. Furthermore, because $\sum_{v \in V_1} \deg(v) \leq 2m$, it follows that $|V_1| \leq 2\sqrt{m}$, and in particular $\chi(G_{|V_1}) \leq 2\sqrt{m}$. The result follows by Lemma 19. ◀

We use a variant of the same partitioning trick in the following lemma.

▶ **Lemma 25.** $\Pr[\chi(G_{\frac{1}{2},e}) \leq k] \geq \alpha \Rightarrow \chi(G) \leq (4 + 2r)km^{1/4}$.

**Proof.** Note first that if $k \geq m^{1/4}$, then the claimed bound holds by Lemma 24. So we assume henceforth that $k < m^{1/4}$.

Identify subsets of edges $E$ with vectors in $\mathbb{F}_2^m$. Because $\Pr[\chi(G_{\frac{1}{2},e}) \leq k] \geq \alpha$ by Lemma 15 there exist $E_1, E_2 \subseteq E$ such that $G_1 = (V, E_1), G_2 = (V, E_2)$ are $k$-colorable and $|E_1 \cup E_2| \geq m - (1 + r)\sqrt{m}$.

Let $E_3 = E \setminus (E_1 \cup E_2)$ be the set of edges that are not in $E_1 \cup E_2$, and define the graph $G_3 = (V, E_3)$. Define a partition $U, \overline{U}$ of $V$, where $U = \{v \in V : \deg_{G_3}(v) < m^{1/4}/k\}$ and $\overline{U} = \{v \in V : \deg_{G_3}(v) \geq m^{1/4}/k\}$. We claim (1) that $\chi(G_{|U}) \leq 2km^{1/4}$ and (2) that $\chi(G_{|\overline{U}}) \leq 2(1 + r)km^{1/4}$. By Lemma 19 we then get the upper bound $\chi(G) \leq \chi(G_{|U}) + \chi(G_{|\overline{U}}) \leq (4 + 2r)km^{1/4}$.

To prove (1) note that by Brooks' Theorem [6] we have $\chi((G_3)_{|U}) \leq 2m^{1/4}/k$, and thus by Lemma 20 $\chi(G_{|U}) \leq \chi(G_1) \cdot \chi(G_2) \cdot \chi((G_3)_{|U}) \leq 2km^{1/4}$. For (2) note that $\sum_{v \in \overline{U}} \deg_{G_3}(v) \leq 2(1 + r)\sqrt{m}$, and hence $\chi(G_{|\overline{U}}) \leq |\overline{U}| \leq 2(1 + r)km^{1/4}$, as required. ◀

Taking the contrapositive of Lemmas 23 and 25 implies Theorem 6.

## Proof of Theorem 7

Finally, we use Theorem 6 to prove the strong robustness result for Gap-Coloring. Let $G$ be an instance of the Gap-Coloring$(q, Q)$ problem. We claim the following:

**YES-case:** If $\chi(G) \leq q$, then $\chi(G_{\frac{1}{2},e}) < q$.

**NO-case:** If $\chi(G) \geq Q$, then $\chi(G_{\frac{1}{2},e}) \geq \Omega(Q^{1/3})$ with probability at least 0.99.

The YES-case is clear, since removing edges cannot increase the chromatic number. The NO-case follows from Theorem 6. Thus, the Gap-Coloring$(q, Q)$ problem is strongly-noise-robust to the Gap-Coloring$(q, \Omega(Q^{1/3}))$ problem. The "in particular" part of the theorem follows from the result of Huang [18] showing that Gap-Coloring$(q, 2^{\Omega(q^{1/3})})$ is $\mathcal{NP}$-hard.

**Acknowledgements.** We thank Itai Benjamini, Uri Feige, Sam Hopkins for useful discussions. We are grateful to the anonymous referees for helpful comments and turning our attention to [7]. Part of this work was done during the visit of the second author at NYU. The second author wishes to thank Subhash Khot for his hospitality.

## References

1 N. Alon and J. H. Spencer. *The Probabilistic Method*. Wiley-Interscience series in discrete mathematics and optimization. J. Wiley & Sons, New York, 2000.

2 B. Barak, M. Hardt, T. Holenstein, and Steurer D. Subsampling mathematical relaxations and average-case complexity. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, SODA, San Francisco, California, USA*, pages 512–531, 2011. doi:10.1137/1.9781611973082.41.

3 B. Bollobás. The chromatic number of random graphs. *Combinatorica*, 8(1):49–55, 1988. doi:10.1007/BF02122551.

4 B. Bollobás. *Random graphs*. Springer, 1998.

5 B. Bollobás, B. P. Narayanan, and A. M. Raigorodskii. On the stability of the erdős-ko-rado theorem. *J. Comb. Theory, Ser. A*, 137:64–78, 2016. doi:10.1016/j.jcta.2015.08.002.

6 R. L. Brooks. On colouring the nodes of a network. *Mathematical Proceedings of the Cambridge Philosophical Society*, 37:194–197, 4 1941. doi:10.1017/S030500410002168X.

7 B. Bukh. Interesting problems that I cannot solve. Problem 2. http://www.borisbukh.org/problems.html.

8 I. Dinur, E. Mossel, and O. Regev. Conditional hardness for approximate coloring. *SIAM J. Comput.*, 39(3):843–873, 2009. doi:10.1137/07068062X.

9 I. Dinur and I. Shinkar. On the conditional hardness of coloring a 4-colorable graph with super-constant number of colors. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, 13th International Workshop, APPROX 2010, and 14th International Workshop, RANDOM 2010, Barcelona, Spain, September 1-3, 2010. Proceedings*, pages 138–151, 2010. doi:10.1007/978-3-642-15369-3_11.

10 M. Etscheid and H. Röglin. Smoothed analysis of local search for the maximum-cut problem. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA, Portland, Oregon, USA*, pages 882–889, 2014. doi:10.1137/1.9781611973402.66.

11 U. Feige and J. Kilian. Zero knowledge and the chromatic number. *Journal of Computer and System Sciences*, 57(2):187–199, 1998. doi:10.1006/jcss.1998.1587.

12 U. Feige and Daniel Reichman. Recoverable values for independent sets. *Random Struct. Algorithms*, 46(1):142–159, 2015. doi:10.1002/rsa.20492.

13 A. M. Frieze and C. McDiarmid. Algorithmic theory of random graphs. *Random Struct. Algorithms*, 10(1-2):5–42, 1997. doi:10.1002/(SICI)1098-2418(199701/03)10:1/2<5::AID-RSA2>3.0.CO;2-Z.

14 M. R. Garey and D. S. Johnson. The complexity of near-optimal graph coloring. *J. ACM*, 23(1):43–49, January 1976. doi:10.1145/321921.321926.

15 G. Grimmett. *Percolation*. Springer, 1999.

16 G. R. Grimmett and C. J. H. McDiarmid. On colouring random graphs. *Mathematical Proceedings of the Cambridge Philosophical Society*, 77:313–324, 3 1975. doi:10.1017/S0305004100051124.

17 S. Har-Peled. Concentration of Random Variables – Chernoff's Inequality. Available at http://sarielhp.org/teach/13/b_574_rand_alg/lec/07_chernoff.pdf.

18 S. Huang. Improved hardness of approximating chromatic number. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques – 16th International Workshop, APPROX 2013, and 17th International Workshop, RANDOM 2013, Berkeley, CA, USA, August 21-23, 2013. Proceedings*, pages 233–243, 2013. doi:10.1007/978-3-642-40328-6_17.

19 D. R. Karger. Random sampling in cut, flow, and network design problems. In *Proceedings of the Twenty-Sixth Annual ACM Symposium on Theory of Computing, Montréal, Québec, Canada*, pages 648–657, 1994. doi:10.1145/195058.195422.

20 S. Khot. Improved inapproximability results for maxclique, chromatic number and approximate graph coloring. In *42nd Annual Symposium on Foundations of Computer Science, FOCS 2001, 14-17 October 2001, Las Vegas, Nevada, USA*, pages 600–609, 2001. doi:10.1109/SFCS.2001.959936.

21 L. Kucera. The greedy coloring is a bad probabilistic algorithm. *J. Algorithms*, 12(4):674–684, 1991.

22 T. Łuczak. The chromatic number of random graphs. *Combinatorica*, 11(1):45–54, 1991. doi:10.1007/BF01375472.

23 M. Mezard and A. Montanari. *Information, physics, and computation.* Oxford University Press, 2009.

24 S. Misra. *Decay of Correlation and Inference in Graphical Models.* PhD thesis, MIT, 2014.

25 E. Mossel, R. O'Donnell, O. Regev, J. E. Steif, and B. Sudakov. Non-interactive correlation distillation, inhomogeneous Markov chains, and the reverse Bonami-Beckner inequality. *Israel Journal of Mathematics*, 154:299–336, 2006.

26 O. Sisask. Discrete fourier analysis. Available at https://people.kth.se/~sisask/Shillong/DFA_2.pdf.

27 A. Sly. Computational transition at the uniqueness threshold. In *51th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2010, Las Vegas, Nevada, USA*, pages 287–296, 2010. doi:10.1109/FOCS.2010.34.

28 D. A. Spielman and S-H. Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *J. ACM*, 51(3):385–463, 2004. doi:10.1145/990308.990310.

29 D. A. Spielman and S-H. Teng. Smoothed analysis: an attempt to explain the behavior of algorithms in practice. *Commun. ACM*, 52(10):76–84, 2009. doi:10.1145/1562764.1562785.

30 V. V. Vazirani. *Approximation Algorithms.* Springer-Verlag New York, Inc., New York, NY, USA, 2001.