



On the kernel size of clique cover reductions for random intersection graphs



Tobias Friedrich^{a,*}, Christian Hercher^b

^a Hasso Plattner Institute, Potsdam, Germany

^b Friedrich-Schiller-Universität Jena, Germany

ARTICLE INFO

Article history:

Received 30 November 2014

Received in revised form 21 April 2015

Accepted 27 May 2015

Available online 9 June 2015

Keywords:

Parameterized algorithms

Clique cover

Kernelization

Average case

Random intersection graphs

Erdős–Rényi graphs

ABSTRACT

Covering all edges of a graph by a minimum number of cliques is a well known NP-hard problem. For the parameter k being the maximal number of cliques to be used, the problem becomes fixed parameter tractable. However, assuming the Exponential Time Hypothesis, there is no kernel of subexponential size in the worst-case.

We study the average kernel size for random intersection graphs with n vertices, edge probability p , and clique covers of size k . We consider the well-known set of reduction rules of Gramm, Guo, Hüffner, and Niedermeier (2009) [17] and show that with high probability they reduce the graph completely if p is bounded away from 1 and $k < c \log n$ for some constant $c > 0$. This shows that for large probabilistic graph classes like random intersection graphs the expected kernel size can be substantially smaller than the known exponential worst-case bounds.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

In the past few years, many results on upper and lower bounds on the kernel sizes of parameterized problems have been shown. Nearly all of them only consider the worst-case. We take a different route and consider the average kernel size for a probabilistic graph model and present a tight characterization depending on the graph density.

Our study problem is the NP-hard problem CLIQUE COVER, which aims to covering the edges of a graph with a minimum number of cliques. The problem arises in studies of the interaction of entities in real-world networks [18] and in protein–protein interaction networks [3]. It also has applications in compiler optimization [29], computational geometry [1], and computational statistics [16,27]. In different domains the problem has been described by varying names, other variants are KEYWORD CONFLICT [22], COVERING BY CLIQUES [15], and INTERSECTION GRAPH BASIS [14].

Formal definition of the problem. A clique in an undirected graph $G = (V, E)$ is a subgraph, where any two vertices are connected by an edge. For an undirected graph G and an integer $k \geq 0$, the decision problem (EDGE) CLIQUE COVER answers yes iff there is a set of at most k cliques $\{G_1, G_2, \dots, G_k\}$ in G such that each edge in G has both its endpoints in at least one of the cliques G_i .

* Corresponding author.

E-mail address: surname@hpi.de (T. Friedrich).

Previous results for special graph classes. CLIQUE COVER has been studied from many perspectives. It remains NP-hard even when restricting to planar input graphs [6] or graphs with maximum degree 6 [21]. However there are polynomial-time algorithms when restricting to graphs with maximum degree 5 [21], chordal graphs [24] or line graphs [26]. Back in the general setting Lund and Yannakakis have shown that the optimization variant of CLIQUE COVER is not approximable within a factor of $|V|^\varepsilon$ for some $\varepsilon > 0$ (unless $P = NP$) [23]. In fact, it remains APX-hard even when restricting to biconnected graphs of maximum degree 7 [20]. Hence there is no hope of good and fast approximation algorithms.

Previous parametrized results. Another popular approach on handling the problem is studying its parametrized version. Fixing the output size as a parameter k , Gyarfas [19] presents a set of simple reduction rules. Successively applying such reduction rules leads to a smaller instance and/or parameter. This process is called a *kernelization* as it results in a so-called kernel, which is a yes-instance iff the original input was a yes-instance, too. The reduction rules considered in [19] leave a kernel of size (number of vertices) at most 2^k . Computing this kernelization can be done in $\text{poly}(n)$ time. As this kernel size does not depend on $|V| = n$, CLIQUE COVER with this parametrization is therefore in FPT, see Gramm et al. [17]. Performing these reduction rules, we get an algorithm with runtime $\mathcal{O}(n^4 + f(2^k))$. Recent results of Cygan et al. [7,8] show that there is no guaranteed polynomial sized kernel if not $NP \subseteq \text{coNP/poly}$ and there is no subexponential sized kernel, unless the Exponential Time Hypothesis fails. In fact, unless $P = NP$, the kernel has to be of exponential size in worst case [28]. Thus we do not expect a $2^{o(k)} \cdot \text{poly}(n)$ time algorithm – a double exponential runtime is required. These results are not depending on the set of reduction rules used for kernelization. Moreover, kernel sizes (whether polynomial or not) are of great interest because they lead to a finer structural analysis of FPT problems, see Bodlaender et al. [4].

Previous average case results. CLIQUE COVER has also been studied on random graphs. Bollobas et al. [5] give lower and upper bounds on the number of cliques required to cover the entire graph, which hold with high probability (meaning with probability $1 - o(1)$ as n tends to infinity, w.h.p.) for random graphs (Erdős–Renyi) with edge probability $p = \frac{1}{2}$. This is equivalent to a uniform distribution over all graphs with set of vertices $|V| = \{1, 2, \dots, n\}$. This result was improved to $\Theta(n^2 / \log^2 n)$ for all constant p with $0 < p < 1$ by Frieze and Reed [13]. Also on random intersection graphs the problem was studied. Behrisch and Taraz [2] give algorithms, which find w.h.p. a clique cover of minimal size in polynomial time for certain probability functions, if the underlying feature set (and thus expected size of the clique cover) is n^α for a constant $0 < \alpha < 1$.

Our results. We study the kernel size in a probabilistic graph model called random intersection graphs similar to Erdős–Renyi that ensures that the drawn graphs are coverable by at most k cliques (for details see Section 3). First we study a set of reduction rules (defined in Section 4), which intuitively arises from the definition of the graph model. In Section 5 we show that for sparse graphs (edge probabilities p decreasing at least polynomially to 0) we get w.h.p. polynomial sized kernels with respect to the parameter k . Therefore the worst case instances there are rare. If the graphs are dense (i.e., edge probabilities p not decreasing at least polynomially to 0, e.g., constant p with $0 < p < 1$) the set of reduction rules w.h.p. reduces the instances only to kernels with exponential size. We give a full characterization of the (w.h.p.) kernel sizes with respect to the edge probability in Theorem 5.1. The situation of stepwise decreasing kernel size is visualized in Fig. 1 on page 132. At probabilities $p(n)$ which are asymptotically $n^{-2/i}$, with $1 < i \leq k$ (more precisely, $\lim_{n \rightarrow \infty} \frac{\log p}{\log n} = -\frac{2}{i} =: -a$), the behavior of the kernel size cannot be described with full certainty by the parameter a alone and only an interval can be stated.

In a second step we generalize the previous set of reduction rules and study the original set of reduction rules given by Gramm et al. [17]. For this we can use the same techniques to show in Section 6 that for an edge probability of $p < 1 - r^k$ for a constant $0 < r < 1$ and $k < c(r) \log n$ for a positive constant $c(r)$ depending on r this set of rules reduces the graph w.h.p. completely. For random intersection graphs the reduction rules of Gramm et al. [17] can therefore be seen as an ‘optimal set’ of reduction rules for kernelization.

2. Erdős–Renyi random graphs

The well known observation that searching for a minimal clique cover is NP-complete is a pure worst case statement. For a better understanding, we want to study probabilistic graph models and perform an average analysis. One such popular model was constructed by Erdős and Renyi [9] and is defined as follows.

Definition 2.1 (*Erdős–Renyi graphs*). Let n be a positive integer and $p(n)$ a function $\mathbb{N} \rightarrow [0, 1]$. Then $G(n, p)$ is the probability space on the set of graphs with n vertices, where each edge is in the graph with probability $p = p(n)$ and every two edges are drawn independently.

This model was used in the analyses of CLIQUE COVER by Bollobas et al. [5] and Frieze and Reed [13]. For a constant edge probability $0 < p < 1$ w.h.p., the number of cliques needed to cover the graph tends to infinity. Consequently, if we fix the number k of such cliques, most instances will be no-instances. This can be stated more rigorously as follows.

Lemma 2.2. Let k be a fixed positive integer, $G(n, p)$ the Erdős–Rényi model with edge probability $p(n)$ and number of nodes n tending to infinity. If $p(n)$ is not tending to 1 and in $\omega\left(\frac{1}{n^2}\right)$ a graph is w.h.p. a no-instance.

Proof. In $G(n, p)$ with p not tending to 1, the largest clique has a size of $\mathcal{O}(\log n)$. Hence at most $\mathcal{O}(k \cdot \log^2 n)$ edges could be covered with k cliques. As for $p = \omega(\log^2(n)/n^2)$ there are more edges to cover, such graphs have to be no-instances w.h.p. If $p = \mathcal{O}\left(\frac{\log^2 n}{n^2}\right) = o\left(\frac{1}{n}\right)$, almost all vertices have degree at most 1. To cover these edges, we need a unique clique. This is only possible if there are at most $\Theta(k)$ edges on average in the graph. Hence $p = \mathcal{O}\left(\frac{1}{n^2}\right)$. \square

This shows that model of Erdős–Rényi graphs does not give a good insight in the process of kernelization of CLIQUE COVER as w.h.p. we only get no-instances. We could now try to (probabilistically) condition on yes-instances, but this results in very hard to control dependencies between the edges. We therefore study the alternative model of random intersection graphs, which also forces the graphs to be yes-instances.

3. Random intersection graphs

In Gramm et al. [17] simple reduction rules (first introduced in Gyárfás [19]) for kernelization of CLIQUE COVER were studied. Experimental results showed that these rules are very effective. However, the test cases (random graphs in an Erdős–Rényi model with edge probability of $p = 0.15, 0.1$ and $\log(n)/n$) were very small (graph sizes of $n = 85, 150$ and 1000 , respectively). This is the reason why most instances in their study were yes-instances. However, we observed in Lemma 2.2 that with the number of nodes growing to infinity, this does no longer hold.

This motivates studying a random graph model for graphs which are known to be yes-instances. We study the well-known model of random intersection graphs, which is a probabilistic model of graphs that inherits a clique cover of size at most k for a function $k(n): \mathbb{N} \rightarrow \mathbb{N} \setminus \{0\}$. We use the following definition.

Definition 3.1 (Random intersection graphs). Let $n > 0$ be an integer. We construct the probability space $CC_k(n, p)$ on the set of graphs with n vertices as follows:

- Let $q := \left(1 - (1 - p)^{\frac{1}{k}}\right)^{\frac{1}{2}}$.
- For each vertex v , choose independently at random a vector $c_v \in \{0, 1\}^k$.
- Every such entry is chosen independently at random to be 1 with probability q and 0 otherwise.
- Add an edge between two different vertices v and w if and only if there is an index $1 \leq \ell \leq k$ such that both vectors c_v and c_w have an entry 1 at this index ℓ .
- All other graphs are drawn with probability 0.

The idea of assigning a binary vector to each vertex which is coding the membership of the given vertex in the covering cliques first appeared in the proof of bounded kernel sizes by Gyárfás [19]. Random intersection graphs have been widely studied, e.g. by Michal et al. [25] for subgraphs and by Behrisch and Taraz [2] for clique cover. For a survey on recent results for random intersection graphs, see Zhao et al. [30].

It should be noted that the usual definition of random intersection graphs uses the ‘vector probability’ q directly. However, our definition based on the ‘edge probability’ p is equivalent: If k is given, Definition 3.1 gives a direct conversion from p to q and vice versa. For a better comparison with the standard Erdős–Rényi model $G(n, p)$ with edge probability $p(n)$, we prefer this way of defining random intersection graphs.

With similar arguments as Gyárfás [19], we can show a maximal kernel size of $2^k - 1$ (not only 2^k) because of the elimination of the zero vector (see Theorem 4.1 below). We now first observe that the model of random intersection graphs only leaves yes-instances.

Theorem 3.2. For $0 < p < 1$, a graph has positive probability if and only if it inherits a clique cover of size at most k .

Proof. First observe that if $p \in (0, 1)$, also $q \in (0, 1)$.

Let us first consider the only-if-part. If a graph has a cover consisting of the k cliques C_1 to C_k (where some might be empty), we can assign vectors as in the model to the vertices as follows: If a vertex v is in clique C_i , the i -th index of its vector c_v is set to 1, otherwise to 0. Such a collection of vectors has a strict positive probability.

We now consider the if-part. As every graph with positive probability has such a vector assignment, all vertices with entry 1 at index i in its vector form a clique C_i and by definition all edges are covered by these cliques. \square

We observe that the random intersection model behaves very similar to Erdős–Rényi graphs conditioned on yes-instances:

Theorem 3.3. Let $n > 0$, $0 < p < 1$ and $CC_k(n, p)$ be the probability space defined above. Let G be a randomly chosen graph out of $CC_k(n, p)$. Then:

- All potential edges have the same probability of p , but are not independently of another.
- The expected degree of all vertices is identical. Thus we have a homogeneous graph model.

Proof. Consider two different vertices. An edge between them exists if and only if there is one of the k indices where both vectors have an entry 1. The counterevent means that in all k components there are no two 1s. This happens with probability $1 - q^2 = (1 - p)^{\frac{1}{k}}$ and hence the counterevent has probability $\left((1 - p)^{\frac{1}{k}}\right)^k = 1 - p$ as the components are independent of each other. Therefore every potential edge occurs with probability p . As the definition is invariant under permutation of the vertices, all vertices have the same expected degree. \square

In random intersection graphs, different edges are not independent of each other. This would destroy the property of guaranteeing a k -clique cover. However, we note the dependencies are surely of a different structure than in Erdős–Rényi graphs conditioned on yes-instances. The advantage of random intersection graphs is that (due to the vectors c_v) we have a tool to analyze the edge dependencies, which is not the case on conditioned Erdős–Rényi graphs.

4. Reduction rules

We study two sets of reduction rules. The first set contains the following three rules introduced by Gramm et al. [17].

Reduction Rule Set 1. Apply the following rules:

- (1.1) Delete a vertex if it is isolated or only incident to covered edges.
- (1.2) If an uncovered edge is contained in exactly one maximal clique C , then mark its edges as covered, and decrease k by one.
- (1.3) From two vertices with identical closed neighborhood, delete one.

This set of reduction rules can be performed in time polynomial in n . After they cannot be applied anymore, they result in a kernel of size at most 2^k (or the input graph was a no-instance). This bound on the kernel size and the correctness of the rules (applying them does not change whether the graph are a yes- or no-instance) is proven by Gramm et al. [17]. For studying these rules, it will be beneficial to first consider the following simple set of related reduction rules, which are more specific for the considered random intersection graph model.

Reduction Rule Set 2. Apply the following rules:

- (2.1) Delete a vertex if its vector is a zero vector.
- (2.2) From two vertices with identical vectors, delete one.

This set of rules gives sufficient conditions for the applicability of the rules given by Gramm et al. [17].

Theorem 4.1. Let n, k be integers, $0 < p < 1$ and G be a graph randomly chosen from $CC_k(n, p)$. Let the underlying set of vectors be $\{c_v : v \in V(G)\}$. If one of the rules of [Reduction Rule Set 2](#) could be applied, then the corresponding rules of [Reduction Rule Set 1](#) also applies.

Proof. We check both rules of [Reduction Rule Set 2](#) separately:

- (1) A vertex with zero vector assigned is isolated and therefore could be deleted by rule (1.1).
- (2) If there are two vertices v and w with identical vectors c_v and c_w , both vectors are connected (or both isolated) and have the same set of neighbors. Thus rule (1.3) (or (1.1)) could be used to eliminated one. \square

Our set of reduction rules defined in [Reduction Rule Set 2](#) is weaker than the original [Reduction Rule Set 1](#). Thus the kernel size after [Reduction Rule Set 2](#) is an upper bound of the kernel size given by [Reduction Rule Set 1](#).

5. Analysis of kernelization with [Reduction Rule Set 2](#)

We can now state and prove our first theorem. It shows that in dense graphs (i.e., expected degree = $\Theta(n)$) [Reduction Rule Set 2](#) leaves w.h.p. a kernel of exponential size in k ; while in sparse graphs (expected degree = $\mathcal{O}(n^{1-a})$ for a positive real constant a) we get w.h.p. a kernel whose size is bounded above by a polynomial in k (with degree only depending

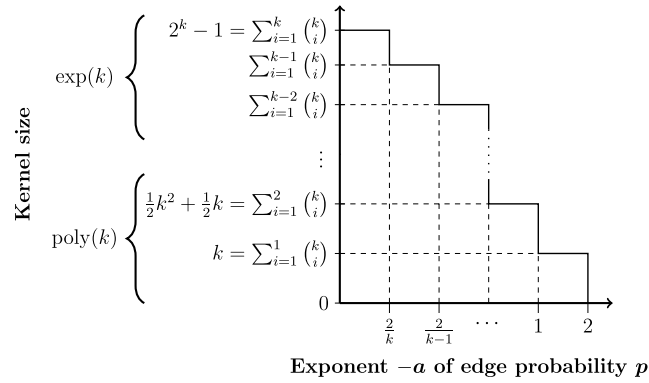


Fig. 1. Illustration of the resulting kernel sizes as a function of the edge probability for Reduction Rule Set 2, as shown in Theorem 5.1: The size of the kernel (w.h.p.) as a function of the edge probability exponent $a := -\lim_{n \rightarrow \infty} \frac{\log p}{\log n}$. The kernels for a given edge probability in the horizontal sections are unique up to isomorphism (w.h.p.). The shown bounds are all tight (w.h.p.). Only at the jump discontinuities the kernel size varies w.h.p. in the given interval.

on a). In all but the last case of the theorem, the given kernels are unique up to isomorphism. An illustration is given in Fig. 1.

Theorem 5.1. Let $k > 2$ be a fixed positive integer and $n \rightarrow \infty$. Let $p(n)$ be a function $\mathbb{N} \rightarrow \left(0, 1 - \left(\frac{3}{4}\right)^k\right)$ with existing limit $a := \lim_{n \rightarrow \infty} -\frac{\log p}{\log n}$ for a nonnegative real number a and $CC_k(n, p)$ the probability space defined in Definition 3.1. Then applying Reduction Rule Set 2 results w.h.p. in a kernel of size

- $2^k - 1$ if $a < \frac{2}{k}$,
- $\sum_{i=1}^{\ell} \binom{k}{i} = \frac{1}{\ell!} k^\ell + \Theta(k^{\ell-1})$ if $\frac{2}{\ell+1} < a < \frac{2}{\ell}$ for a fixed integer $1 \leq \ell < k$,
- 0 if $2 < a$ and
- $\sum_{i=1}^{\ell-1} \binom{k}{i} + c \cdot \binom{k}{\ell} = \frac{c}{\ell!} k^\ell + \Theta(k^{\ell-1})$ with a real constant $0 < c < 1$ (depending on the more detailed behavior of $\frac{\log p}{\log n}$) if $a = \frac{2}{\ell}$ for a positive integer $\ell \leq k$.

Proof. Let a be a positive real constant and $p(n)$ a function which for every $\varepsilon > 0$ is in $\mathcal{O}(n^{-a-\varepsilon}) \cap \Omega(n^{-a+\varepsilon})$. This is equivalent to the statement that $\lim_{n \rightarrow \infty} -\frac{\log p(n)}{\log n}$ exists and is equal to a .

We first prove a similar statement for the “model probability” $q = \left(1 - (1-p)^{\frac{1}{k}}\right)^{\frac{1}{2}}$, that is, $\log q = \frac{1}{2} \cdot \log \left(1 - (1-p)^{\frac{1}{k}}\right)$. For this, we calculate the Taylor expansion of the argument of the logarithm at $p_0 = 0$ and get $1 - (1-p)^{\frac{1}{k}} = \frac{1}{k} \cdot p + \mathcal{O}(p^2)$. Taking the logarithm gives

$$\log q = \frac{1}{2} \cdot (\log p - \log k) + \mathcal{O}(\log(1+p)) = \frac{1}{2} \log p + \mathcal{O}(1)$$

and thus

$$\lim_{n \rightarrow \infty} -\frac{\log q}{\log n} = \lim_{n \rightarrow \infty} -\frac{1}{2} \cdot \frac{\log p + \mathcal{O}(1)}{\log n} = \frac{1}{2} \cdot \lim_{n \rightarrow \infty} -\frac{\log p}{\log n} = \frac{1}{2} \cdot a,$$

which implies

$$q(n) \in \mathcal{O}\left(n^{-\frac{1}{2}(a-\varepsilon)}\right) \cap \Omega\left(n^{-\frac{1}{2}(a+\varepsilon)}\right) \text{ for every } \varepsilon > 0.$$

Let now $0 < \ell \leq k$ be an integer and v be the vector with the value 1 at the first ℓ entries and 0 otherwise. The probability of this vector is $q^\ell \cdot (1-q)^{k-\ell} = \mathcal{O}\left(n^{-\frac{1}{2}(a-\varepsilon)\cdot\ell}\right)$. For simplicity, let us assume that the weight of a vector is the number of nonzero entries. Hence clearly every vector with weight ℓ has the same probability and because of $q = \left(1 - (1-p)^{\frac{1}{k}}\right)^{\frac{1}{2}} < \left(1 - \left(1 - \left(1 - \left(\frac{3}{4}\right)^k\right)^{\frac{1}{k}}\right)\right)^{\frac{1}{2}} = \left(1 - \frac{3}{4}\right)^{\frac{1}{2}} = \frac{1}{2}$ and therefore $q < 1 - q$, the probability of a vector with higher weight is at most this value. As there are only at most 2^k different such vectors, the probability of a vector having weight at least ℓ is in $\mathcal{O}\left(n^{-\frac{1}{2}(a-\varepsilon)\cdot\ell}\right)$, where the constant only depends on k .

If $\frac{1}{2}a \cdot \ell > 1$, there exists an $\varepsilon > 0$ such that also $\frac{1}{2}(a - \varepsilon) \cdot \ell > 1$. Then the probability of such a vector is in $o\left(\frac{1}{n}\right) = \frac{o(1)}{n}$. Thus with probability $1 - \frac{o(1)}{n}$, a randomly chosen vector has weight at most $\ell - 1$. Moreover, with probability $P := \left(1 - \frac{o(1)}{n}\right)^n$ all n chosen vectors have weight at most $\ell - 1$. Now observe that with n tending to infinity, P tends to 1. Therefore w.h.p. there are only vectors with at most $\ell - 1$ nonzero entries and after deleting duplicates with reduction rule (2.2) only at most one vector of every type remains.

On the other hand, if $\frac{1}{2}a \cdot \ell < 1$, there is also an $\varepsilon > 0$ with $\frac{1}{2}(a + \varepsilon) \cdot \ell < 1$ and hence the probability of a vector v with weight ℓ is in $\omega\left(\frac{1}{n}\right) = \frac{\omega(1)}{n}$. Thus with probability $1 - \frac{\omega(1)}{n}$, a randomly chosen vector is not v . Moreover, with probability $P := \left(1 - \frac{\omega(1)}{n}\right)^n$ all n chosen vectors differ from v . As now with n tending to infinity, P tends to 0, we get w.h.p. that v is one of the n chosen vectors.

Summing up, we get the following situation: If $\frac{2}{\ell+1} < a < \frac{2}{\ell}$ for an integer $\ell < k$ then w.h.p. all vectors with weight $\leq \ell$ and none of the vectors with weight $> \ell$ occur under the n chosen vectors. If a is exactly $\frac{2}{\ell}$, also w.h.p. all vectors with weight $< \ell$ occur and all with weight $> \ell$ do not occur. However, vectors with weight exactly ℓ occur with a positive probability < 1 depending on the more detailed behavior where $\frac{\log p}{\log n}$ goes to its limit. Thus only a part of them will contribute to the kernel. If $a < \frac{2}{k}$, all 2^k vectors will occur w.h.p. Therefore only the zero vector will be eliminated there.

All duplicates of vectors will be deleted by reduction rule (2.2) and the zero vector by (2.1). There are exactly $\binom{k}{m} = \mathcal{O}(k^m)$ different vectors with weight m . Summing up over the relevant weights, gives the claimed results.

If not $a = \frac{2}{\ell}$ for an integer $1 < \ell < k$, w.h.p. the type of every remaining vector (which is the type of the node in which cliques he is part of) is unique. Therefore in these cases the remaining kernel is unique up to isomorphism. \square

If p is near one, we get the following similar but slightly different result:

Corollary 5.2. Let $k > 2$ be a fixed positive integer and $n \rightarrow \infty$. Let $p(n)$ be a function $\mathbb{N} \rightarrow \left(1 - \left(\frac{3}{4}\right)^k, 1\right)$ with existing limit $a := \lim_{n \rightarrow \infty} -\frac{\log(1-p)}{\log n}$ for a nonnegative real number a and $CC_k(n, p)$ the probability space defined in Definition 3.1. Then applying Reduction Rule Set 2 results w.h.p. in a kernel of size

- $2^k - 1$ if $a < 1$,
- $\sum_{i=0}^{\ell} \binom{k}{i}$ if $\frac{k}{\ell+1} < a < \frac{k}{\ell}$ for a fixed integer $1 \leq \ell < k$,
- 1 if $k < a$ and
- $\sum_{i=0}^{\ell-1} \binom{k}{i} + c \cdot \binom{k}{\ell}$ with a real constant $0 < c < 1$ (depending on the more detailed behavior of $\frac{\log(1-p)}{\log n}$) if $a = \frac{k}{\ell}$ for a positive integer $\ell \leq k$.

Proof. We want to use the proof of Theorem 5.1, but exchanging the roles of the entries 0 and 1 in the vectors. Therefore we have to work with $\tilde{q} := 1 - q$ instead of q in the proof above. Now if $p = 1 - \mathcal{O}(n^{-a})$, we have $q = \left(1 - (1-p)^{\frac{1}{k}}\right)^{\frac{1}{2}} = \left(1 - \mathcal{O}(n^{-\frac{a}{k}})\right)^{\frac{1}{2}} = 1 - \mathcal{O}(n^{-\frac{a}{k}})$. Hence $\tilde{q}(n) = 1 - q(n) = n^{-\frac{a}{k}}$. This gives in an analogous way as in the proof of Theorem 5.1: $q(n) \in \mathcal{O}\left(n^{-\frac{1}{k}(a-\varepsilon)}\right) \cap \Omega\left(n^{-\frac{1}{k}(a+\varepsilon)}\right)$ for every $\varepsilon > 0$. Now all calculations, which weights of vectors occur, could be done in the exact same way.

We therefore get that: If $\frac{k}{\ell+1} < a < \frac{k}{\ell}$ for an integer $\ell < k$ then w.h.p. all vectors with weight $\geq k - \ell$ and none with smaller weight occurs under the n chosen vectors. Using $\binom{k}{k-\ell} = \binom{k}{\ell}$ this leads to the results stated above. \square

We now observe that our high-probability results also translate to bounds for the average case.

Corollary 5.3. Let the parameter k , the edge probability p , and the probability space $CC_k(p, n)$ be defined as in Theorem 5.1. The expected kernel size of Reduction Rule Set 2 is then exponential in k if $p(n)$ is constant and at most polynomial in k if $p(n)$ or $1 - p(n)$ is in $\mathcal{O}(n^{-a})$.

Proof. If $p(n)$ is constant we have with probability $1 - o(1)$ a kernel of $2^k - 1$ and with probability $o(1)$ a smaller one. Thus with $n \rightarrow \infty$ the average case tends to $2^k - 1$.

If $p(n)$ tends at least polynomially to zero the proof of Theorem 5.1 shows that there is only an exponentially decreasing probability of an occurrence of a vector with weight $> \frac{2k}{a}$ and therefore (because their number is bounded by a polynomial in n) their value also decreases to 0 at the average. An analogous result follows in the same way, if $p(n)$ tends at least polynomially to one. \square

6. Analysis of kernelization with Reduction Rule Set 1

As the kernel size after Reduction Rule Set 2 is an upper bound for the kernel size after Reduction Rule Set 1, all upper bounds of the previous section for Reduction Rule Set 2 immediately translate to upper bounds on the kernel size for Reduction Rule Set 1 by Gramm et al. [17]. However, in this section we show that Reduction Rule Set 1 reduces the kernel even further:

Theorem 6.1 (Main Theorem). *Let n be an integer with $n \rightarrow \infty$ and $k(n)$ a function $\mathbb{N} \rightarrow \mathbb{N}$. Let $p(n)$ be a function $\mathbb{N} \rightarrow (0, 1)$ with $p < 1 - r^k$ for a real constant $0 < r < 1$ and $CC_k(n, p)$ the probability space defined in Definition 3.1. Then there exists a real constant $c(r) > 0$, such that the Reduction Rule Set 1 by Gramm et al. [17] results w.h.p. in a kernel of size zero for every $k(n) < c(r) \cdot \log n$.*

Proof. First note that because of $p < 1 - r^k$ and the definition of the model, we also have $q < \sqrt{1 - r} =: \hat{r}$ with $0 < \hat{r} < 1$. This will be used in the following argument. We split the rest of the proof depending on the size of q .

The main tool of the proof are vectors of weight one. They characterize a unique clique. Thus, if every such vector is present in the graph, we can identify every clique (see Case 2). Otherwise (in Case 1), we show that also no other vector with higher weight was drawn. Therefore the graph can be reduced completely using only the Reduction Rule Set 2. Note that by Theorem 4.1 we can apply both reduction rule sets.

Case 1: $q = \mathcal{O}\left(\frac{\log k}{n}\right)$. Then w.h.p. no vector has weight greater than one. To prove this, we note that the probability that every vector drawn has weight ≤ 1 is $P := ((1 - q)^k + k \cdot q \cdot (1 - q)^{k-1})^n$. Furthermore, it is $P = \left((1 - q)^k \cdot \left(1 + k \cdot \frac{q}{1 - q}\right)\right)^n$. Since $(1 - q)^k > 1 - qk$ and $1 + k \cdot \frac{q}{1 - q} > 1 + kq$ we have $P > ((1 - qk) \cdot (1 + kq))^n = (1 - q^2k^2)^n$. And because of $k = \mathcal{O}(\log n)$, this is $\left(1 - \mathcal{O}\left(\frac{\log^2 n \cdot \log \log n}{n^2}\right)\right)^n = \left(1 - \frac{\mathcal{O}(1)}{n}\right)^n$ and hence P tends to one and w.h.p. all drawn vectors have weight at most one.

After deleting all duplicated vectors with rule (2.2) in every component, there is at most one nonzero entry and hence only isolated nodes are left, which will be cleared up by rule (1.1). Thus the graph reduces completely, which finishes the first case.

Case 2: $q = \omega\left(\frac{\log k}{n}\right)$.

In the remaining cases, we will show that w.h.p. every vector v_t with an entry one only in the t -th component (and all others zero) for $1 \leq t \leq k$ is present. After deleting duplicates, there is only exactly one for every such t . Let c_t be the node corresponding to the unique vector v_t and let C_t be the set of nodes whose vectors have an entry one in the t -th component. By definition, they form a clique and $c_t \in C_t$. Because of the absence of other nonzero entries in v_t , there are no other neighbors of c_t other than the ones part of C_t . Hence C_t is a maximal clique and for every edge incident to c_t , this is the only one which it is part of. Therefore reduction rule (1.2) will reduce this clique by covering all of its edges, which equally removes all entries one out of the t -th components of the vectors. In fact, if an edge between two nodes in C_t is also part of another clique C_s with $s \neq t$, then this edge will not be deleted (because both vectors corresponding to the nodes incident to this edge have a nonzero entry in the s -th component) by this procedure and therefore can contribute to the C_s . This procedure can be repeated for all components $1 \leq t \leq k$. This leaves only zero vectors and isolated vertices. These can be deleted by reduction rule (1.1) and the graph is completely reduced to an empty kernel.

Thus we only have to show that the probability P of the absence of at least one vector with weight one tends to zero as n tends to infinity. Let $Q_k := q \cdot (1 - q)^{k-1}$ be the probability that a specific vector with weight one occurs by drawing one vector as defined in Definition 3.1. Then the probability that this vector is not part of the n drawn ones is $(1 - Q_k)^n$ and therefore the probability P that at least one of the k vectors of weight one is not present, is at most k times this value and hence $P \leq k(1 - Q_k)^n$. Because of the inequality $\log(1 - x) < -x \forall x$, we further get $P \leq \exp(\log k - n \cdot Q_k)$. Therefore it suffices that $Q_k = \omega\left(\frac{\log k}{n}\right)$. Because Q_k as a function of q is concave with maximum at $q = \frac{1}{k}$, we split at this point in cases 2a and 2b.

Case 2a: $q = \omega\left(\frac{\log k}{n}\right)$ and $q \leq \frac{1}{k}$. In this case, the second factor $(1 - q)^{k-1}$ of Q_k is bounded below by a constant $c > 0$. Therefore $Q_k = q \cdot (1 - q)^{k-1} > cq = \omega\left(\frac{\log k}{n}\right)$. Hence the graph w.h.p. gets completely reduced.

Case 2b: $q = \omega\left(\frac{\log k}{n}\right)$ and $\frac{1}{k} < q < \hat{r}$. We note that Q_k as a function of q is monotone decreasing in this case. Thus $Q_k > \hat{r} \cdot (1 - \hat{r})^k$.

Define $c(r) := -\frac{1}{2} \cdot \frac{1}{\log(1 - \hat{r})}$. Because of $0 < \hat{r} < 1$, this constant is positive and $\log(1 - \hat{r}) \cdot (c(r) \cdot \log n) = -\frac{1}{2} \log n$. Therefore for all $k < c(r) \cdot \log n$ we have

$$\begin{aligned}
Q_k &> \hat{r} \cdot (1 - \hat{r})^k > \hat{r} \cdot \exp(1 - \hat{r})^{c(r) \cdot \log n} \\
&= \hat{r} \cdot \exp(\log(1 - \hat{r}) \cdot (c(r) \cdot \log n)) \\
&= \hat{r} \cdot \exp\left(-\frac{1}{2} \log n\right) = \frac{\hat{r} \cdot \sqrt{n}}{n} = \omega\left(\frac{\log k}{n}\right),
\end{aligned}$$

and again the graph w.h.p. will be reduced completely.

Summing up, we have in every case a complete reduction, which proves the statement. \square

For edge probabilities p tending to one, [Theorem 6.1](#) only allows non-constant parameters k . However, also for constant k we can show an equivalent result if p only tends “slow enough” to one:

Corollary 6.2. *Under the same conditions as in [Theorem 6.1](#), but with constant k , graphs drawn from $CC_k(n, p)$ reduce under the Reduction Rule Set 1 by [Gramm et al. \[17\]](#) w.h.p. completely, if $p < 1 - n^{-2/k}$.*

Proof. As in [Theorem 5.1](#) and with the following remark we know that for constant k and with this restriction on p our set of reduction rules reduces w.h.p. to a kernel of size $2^k - 1$. There all vectors with positive weight are present exactly ones, especially the vectors with weight one. But then the graph can be reduced completely by reduction rule (1.2) as in the second part of the proof of [Theorem 6.1](#). \square

7. Conclusion

It is known that the parameterized CLIQUE problem is $W[1]$ -complete in the worst-case, but solvable on average in FPT-time for Erdős–Rényi random graphs [\[10\]](#) and Chung–Lu random graphs [\[11,12\]](#).

We now study the average-case behavior of the problem CLIQUE COVER, which lies in FPT, but is unlikely to have a subexponential kernel in the worst-case. For studying kernel sizes, we have to focus our attention to yes-instances. Previous experiments have used the Erdős–Rényi model, but this model is not appropriate for an asymptotic analysis as it gives w.h.p. only no-instances (cf. [Lemma 2.2](#)). We instead study random intersection graphs, which naturally arise from the property that every covering of a graph with k cliques inherits an assignment of binary vectors to nodes. We prove that the set of reduction rules given by [Gramm et al. \[17\]](#) reduces the graph w.h.p. completely for small enough k and most edge probabilities p (cf. [Theorem 6.1](#)).

Acknowledgements

The first author thanks Matthias Mnich for discussions on average kernelization at the Max-Planck Institute for Informatics that inspired this work. The research leading to these results has received funding from the German Research Foundation (DFG grant FR 2988) and the German–Israeli Foundation for Scientific Research and Development (GIF grant G-2315-1061.6/2012).

References

- [1] P.K. Agarwal, N. Alon, B. Aronov, S. Suri, Can visibility graphs be represented compactly?, *Discrete Comput. Geom.* 12 (1994) 347–365.
- [2] M. Behrisch, A. Taraz, Efficiently covering complex networks with cliques of similar vertices, *Theor. Comput. Sci.* 355 (1) (2006) 37–47.
- [3] M. Blanchette, E. Kim, A. Vetta, Clique cover on sparse networks, in: *Proceedings of the 14th Meeting on Algorithm Engineering & Experiments (ALENEX)*, SIAM, 2012, pp. 93–102.
- [4] H.L. Bodlaender, R.G. Downey, M.R. Fellows, D. Hermelin, On problems without polynomial kernels, *J. Comput. Syst. Sci.* 75 (8) (2009) 423–434.
- [5] B. Bollobás, P. Erdős, J. Spencer, D.B. West, Clique coverings of the edges of a random graph, *Combinatorica* 13 (1) (1993) 1–5.
- [6] M.-S. Chang, H. Müller, On the tree-degree of graphs, in: *Graph-Theoretic Concepts in Computer Science*, Boltenhagen, 2001, 2001, pp. 44–54.
- [7] M. Cygan, M. Pilipczuk, M. Pilipczuk, Known algorithms for edge clique cover are probably optimal, in: *Proceedings of the 24th Annual ACM–SIAM Symposium on Discrete Algorithms (SODA)*, SIAM, 2013, pp. 1044–1053.
- [8] M. Cygan, S. Kratsch, M. Pilipczuk, M. Pilipczuk, M. Wahlström, Clique cover and graph separation: new incompressibility results, *ACM Trans. Comput. Theory* 6 (2) (2014) 6.
- [9] P. Erdős, A. Rényi, On the evolution of random graphs, *Magy. Tud. Akad. Mat. Kut. Intéz. Közl.* 5 (1960) 17–61.
- [10] N. Fountoulakis, T. Friedrich, D. Hermelin, On the average-case complexity of parameterized clique, *Theor. Comput. Sci.* 576 (2015) 18–29.
- [11] T. Friedrich, A. Krohmer, Parameterized clique on scale-free networks, in: *Proceedings of the 23rd International Symposium on Algorithms and Computation (ISAAC)*, 2012, pp. 659–668.
- [12] T. Friedrich, A. Krohmer, Parameterized clique on inhomogeneous random graphs, *Discrete Appl. Math.* 184 (2015) 130–138.
- [13] A. Frieze, B. Reed, Covering the edges of a random graph by cliques, *Combinatorica* 15 (4) (1995) 489–497.
- [14] M.R. Garey, D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W.H. Freeman, 1979.
- [15] F. Gavril, Algorithms for minimum coloring, maximum clique, minimum covering by cliques, and maximum independent set of a chordal graph, *SIAM J. Comput.* 1 (2) (1972) 180–187.
- [16] J. Gramm, J. Guo, F. Hüffner, R. Niedermeier, H. Piepho, R. Schmid, Algorithms for compact letter displays: comparison and evaluation, *Comput. Stat. Data Anal.* 52 (2) (2007) 725–736.
- [17] J. Gramm, J. Guo, F. Hüffner, R. Niedermeier, Data reduction and exact algorithms for clique cover, *ACM J. Exp. Algorithmics* 13 (2009), article ID number 2.2.

- [18] J.-L. Guillaume, M. Latapy, Bipartite structure of all complex networks, *Inf. Process. Lett.* 90 (5) (2004) 215–221.
- [19] A. Gyárfás, A simple lower bound on edge coverings by cliques, *Discrete Math.* 85 (1) (1990) 103–104.
- [20] D. Hermelin, R. Rizzi, S. Vialette, Algorithmic aspects of the intersection and overlap numbers of a graph, in: *Proceedings of the 23rd International Symposium on Algorithms and Computation (ISAAC)*, 2012, pp. 465–474.
- [21] D. Hoover, Complexity of graph covering problems for graphs of low degree, *J. Comb. Math. Comb. Comput.* 11 (1992) 187–208.
- [22] L.T. Kou, L.J. Stockmeyer, C.K. Wong, Covering edges by cliques with regard to keyword conflicts and intersection graphs, *Commun. ACM* 21 (2) (1978) 135–139.
- [23] C. Lund, M. Yannakakis, On the hardness of approximating minimization problems, *J. ACM* 41 (5) (1994) 960–981.
- [24] S. Ma, W.D. Wallis, J. Wu, Clique covering of chordal graphs, *Util. Math.* 36 (1989) 151–152.
- [25] K. Michal, E.R. Scheinerman, K.B. Singer-Cohen, On random intersection graphs: the subgraph problem, *Comb. Probab. Comput.* 8 (1999) 131–159.
- [26] J. Orlin, Contentment in graph theory: covering graphs with cliques, *Indag. Math. (Proceedings)* 80 (5) (1977) 406–424.
- [27] H.-P. Piepho, An algorithm for a letter-based representation of all-pairwise comparisons, *J. Comput. Graph. Stat.* 13 (2) (2004) 456–466.
- [28] M. Pilipczuk, Tournaments and optimality: new results in parameterized complexity, PhD thesis, University of Bergen, Norway, 2013.
- [29] S. Rajagopalan, M. Vachharajani, S. Malik, Handling irregular ILP within conventional VLIW schedulers using artificial resource constraints, in: *Proceedings of the International Conference on Compilers, Architecture, and Synthesis for Embedded Systems (CASES)*, 2000, pp. 157–164.
- [30] J. Zhao, O. Yagan, V. Gligor, Random intersection graphs and their applications in security, wireless communication, and social networks, in: *Proceedings of the Information Theory and Applications Workshop (ITA)*, 2015, arXiv:1504.03161.