The Lovász Number of Random Graphs

Amin Coja-Oghlan*

Humboldt-Universität zu Berlin, Institut für Informatik,
Unter den Linden 6, 10099 Berlin, Germany
coja@informatik.hu-berlin.de

Abstract. We study the Lovász number ϑ along with two further SDP relaxations $\vartheta_{1/2}$, ϑ_2 of the independence number and the corresponding relaxations $\bar{\vartheta}$, $\bar{\vartheta}_{1/2}$, $\bar{\vartheta}_2$ of the chromatic number on random graphs $G_{n,p}$. We prove that $\bar{\vartheta}, \bar{\vartheta}_{1/2}, \bar{\vartheta}_2(G_{n,p})$ in the case $p < n^{-1/2-\varepsilon}$ are concentrated in intervals of constant length. Moreover, we estimate the probable value of $\vartheta, \bar{\vartheta}(G_{n,p})$ etc. for essentially the entire range of edge probabilities p. As applications, we give improved algorithms for approximating $\alpha(G_{n,p})$ and for deciding k-colorability in polynomial expected time.

1 Introduction and Results

Given a graph $G = (V, E)$, let $\alpha(G)$ be the independence number, let $\omega(G)$ be the clique number, and let $\chi(G)$ be the chromatic number of G. Further, let $\bar{G}$ signify the complement of G. Since it is NP-hard to compute any of $\alpha(G)$, $\omega(G)$ or $\chi(G)$, it is remarkable that there exists an efficiently computable function $\vartheta(G)$ that is "sandwiched" between $\alpha(G)$ and $\chi(\bar{G})$, i.e. $\alpha(G) \leq \vartheta(G) \leq \chi(\bar{G})$. Passing to complements, and letting $\bar{\vartheta}(G) = \vartheta(\bar{G})$, we have $\omega(G) \leq \bar{\vartheta}(G) \leq \chi(G)$. The function ϑ was introduced by Lovász, and is called the *Lovász number* of G (cf. [16,21]).

Though $\vartheta(G)$ is sandwiched between $\alpha(G)$ and $\chi(\bar{G})$, Feige [7] proved that the gap between $\alpha(G)$ and $\vartheta(G)$ or between $\chi(\bar{G})$ and $\vartheta(G)$ can be as large as $n^{1-\varepsilon}$, $\varepsilon > 0$. Indeed, unless NP=coRP, none of $\alpha(G)$, $\omega(G)$, $\chi(G)$ can be approximated within a factor of $n^{1-\varepsilon}$, $\varepsilon > 0$, in polynomial time [17,9]. However, though there exist graphs G such that $\vartheta(G)$ is not a good approximation of $\alpha(G)$ (or $\bar{\vartheta}(G)$ of $\chi(G)$), it might be the case that the Lovász number performs well on "average" instances. In fact, several algorithms for random and semirandom graph problems are based on computing ϑ [4,5,8]. Therefore, the aim of this paper is to study the Lovász number of random graphs more thoroughly.

The standard model of a random graph is the binomial model $G_{n,p}$, pioneered by Erdős and Renyi. We let $0 < p = p(n) < 1$ be a number that may depend on n. Let $V = \{1, \ldots, n\}$. Then the random graph $G_{n,p}$ is obtained by including each of the $\binom{n}{2}$ possible edges $\{v, w\}$, $v, w \in V$, with probability p independently. Though $G_{n,p}$ may fail to model some types of input instances appropriately, both the combinatorial structure and the algorithmic theory of

* supported by the Deutsche Forschungsgemeinschaft (grant DFG FOR 413/1-1).

S. Arora et al. (Eds.): APPROX 2003+RANDOM 2003, LNCS 2764, pp. 228–239, 2003.
© Springer-Verlag Berlin Heidelberg 2003

$G_{n,p}$ are of fundamental interest [18,12]. We say that $G_{n,p}$ has some property A *with high probability* (whp.), if $\lim_{n\to\infty} \mathrm{P}(G_{n,p}$ has property $A) = 1$.

We also address two further SDP relaxations $\vartheta_{1/2}$, ϑ_2 of α (cf. [27]) on random graphs. These relaxations satisfy $\alpha(G) \le \vartheta_{1/2}(G) \le \vartheta(G) \le \vartheta_2(G) \le \chi(\bar{G})$, for all G. Passing to complements, and setting $\bar{\vartheta}_i(G) = \vartheta_i(\bar{G})$ $(i = 1/2, 2)$, one gets $\omega(G) \le \bar{\vartheta}_{1/2}(G) \le \bar{\vartheta}(G) \le \bar{\vartheta}_2(G) \le \chi(G)$. The relaxation $\bar{\vartheta}_{1/2}(G)$ coincides with the well-known *vector chromatic number* $\boldsymbol{\chi}(G)$ of Karger, Motwani, and Sudan [20].

The Concentration of $\bar{\vartheta}$, $\bar{\vartheta}_{1/2}$, $\bar{\vartheta}_2$. A remarkable fact concerning the chromatic number of sparse random graphs $G_{n,p}$, $p \le n^{-\varepsilon-1/2}$, is that $\chi(G_{n,p})$ is concentrated in an interval of constant length. Indeed, Shamir and Spencer [26] proved that there is a function $u = u(n,p)$ such that in the case $p = n^{-\beta}$, $1/2 < \beta < 1$, we have $\mathrm{P}(u \le \chi(G_{n,p}) \le u + \lceil (2\beta+1)/(2\beta-1)\rceil) = 1 - o(1)$. Furthermore, Łuczak [25] showed that in the case $5/6 < \beta < 1$, the chromatic number is concentrated in width one. In fact, Alon and Krivelevich [2] could prove that two point concentration holds for the entire range $p = n^{-\beta}$, $1/2 < \beta < 1$. The two following theorems state similar results as given by Shamir and Spencer and by Łuczak for the relaxations $\bar{\vartheta}_{1/2}(G_{n,p})$, $\bar{\vartheta}(G_{n,p})$, and $\bar{\vartheta}_2(G_{n,p})$ of the chromatic number.

Theorem 1. *Suppose that $c_0/n \le p \le n^{-\beta}$ for some large constant $c_0 > 0$ and some number $1/2 < \beta < 1$. Then $\bar{\vartheta}_{1/2}(G_{n,p})$, $\bar{\vartheta}(G_{n,p})$, $\bar{\vartheta}_2(G_{n,p})$ are concentrated in width $s = \frac{2}{2\beta-1} + o(1)$, i.e. there exist numbers u, u', u'' depending on n and p such that whp. $u \le \bar{\vartheta}_{1/2}(G_{n,p}) \le u + s$, $u' \le \bar{\vartheta}(G_{n,p}) \le u' + s$, and $u'' \le \bar{\vartheta}_{1/2}(G_{n,p}) \le u'' + s$.*

Theorem 2. *Suppose that $c_0/n < p \le n^{-5/6-\delta}$ for some large constant c_0 and some $\delta > 0$. Then $\bar{\vartheta}_{1/2}(G_{n,p})$, $\bar{\vartheta}(G_{n,p})$, and $\bar{\vartheta}_2(G_{n,p})$ are concentrated in width 1.*

In contrast to the chromatic number, $\bar{\vartheta}_{1/2}$, $\bar{\vartheta}$, and $\bar{\vartheta}_2$ need not be integral. Therefore, the above results do *not* imply that $\bar{\vartheta}_{1/2}(G_{n,p})$, $\bar{\vartheta}(G_{n,p})$, $\bar{\vartheta}_2(G_{n,p})$ are concentrated on a constant number of points.

The Probable Value of $\vartheta(G_{n,p})$, $\bar{\vartheta}(G_{n,p})$, etc. Concerning the probable value of $\vartheta(G_{n,p})$ and $\bar{\vartheta}(G_{n,p})$, Juhász [19] gave the following partial answer: If $\ln(n)^6/n \ll p \le 1/2$, then with high probability we have $\vartheta(G_{n,p}) = \Theta(\sqrt{n/p})$ and $\bar{\vartheta}(G_{n,p}) = \Theta(\sqrt{np})$. However, we shall indicate in Sec. 4 that Juhász's proof fails in the case of sparse random graphs (e.g. $np = O(1)$). Making use of concentration results on ϑ, $\bar{\vartheta}$ etc., we can compute the probable value not only of $\vartheta(G_{n,p})$ and $\bar{\vartheta}(G_{n,p})$, but also of $\vartheta_i(G_{n,p})$ and $\bar{\vartheta}_i(G_{n,p})$, $i = 1/2, 2$, for essentially the entire range of edge probabilities p.

Theorem 3. *Suppose that $c_0/n \le p \le 1/2$ for some large constant $c_0 > 0$. Then there exist constants $c_1, c_2, c_3, c_4 > 0$ such that*

$$c_1\sqrt{n/p} \le \vartheta_{1/2}(G_{n,p}) \le \vartheta(G_{n,p}) \le \vartheta_2(G_{n,p}) \le c_2\sqrt{n/p} \qquad (1)$$
$$\text{and} \quad c_3\sqrt{np} \le \bar\vartheta_{1/2}(G_{n,p}) \le \bar\vartheta(G_{n,p}) \le \bar\vartheta_2(G_{n,p}) \le c_4\sqrt{np}$$

with high probability. More precisely,

$$\mathrm{P}(c_3\sqrt{np} \le \bar\vartheta_{1/2}(G_{n,p}) \le \bar\vartheta(G_{n,p}) \le \bar\vartheta_2(G_{n,p})) \ge 1 - \exp(-n). \qquad (2)$$

Assume that $c_0/n \le p = o(1)$. Then $\alpha(G_{n,p}) \sim 2\ln(np)/p$ and $\chi(G_{n,p}) \sim np/(2\ln(np))$ whp. (cf. [18]). Hence, Thm. 3 shows that $\vartheta_2(G_{n,p})$ ($\bar\vartheta_{1/2}(G_{n,p})$) approximates $\alpha(G_{n,p})$ ($\chi(G_{n,p})$) within a factor of $O(\sqrt{np})$. In fact, if $np = O(1)$, then we get a constant factor approximation. Our estimate on the vector chromatic number $\bar\vartheta_{1/2}(G_{n,p})$ answers a question of Krivelevich [22].

Finally, consider the *random regular graph* $G_{n,r}$. The proof of the following theorem is somewhat technical, and is omitted.

Theorem 4. *Let c_0 be a sufficiently large constant, and let $c_0 \le r = o(n^{1/4})$. There are constants $c_1, c_2 > 0$ such that whp. the random regular graph $G_{n,r}$ satisfies $c_1 n/\sqrt{r} \le \vartheta_{1/2}(G_{n,r}) \le \vartheta(G_{n,r}) \le \vartheta_2(G_{n,r}) \le c_2 n/\sqrt{r}$. Moreover, there is a constant $c_3 > 0$ such that in the case $c_0 \le r = o(n^{1/2})$ we have* $\mathrm{P}(c_3\sqrt{r} \le \bar\vartheta_{1/2}(G_{n,r}) \le \bar\vartheta(G_{n,r}) \le \bar\vartheta_2(G_{n,r})) \ge 1 - \exp(-n)$.

Algorithmic Applications. There are two types of algorithms for NP-hard random graph problems. First, there are *heuristics* that *always* run in polynomial time, and *almost always* output a good solution. On the other hand, there are algorithms that guarantee some approximation ratio on *any* input instance, and which have a polynomial *expected* running time when applied to $G_{n,p}$. In this paper, we deal with algorithms with a polynomial expected running time.

First, we consider the maximum independent set problem in random graphs. Krivelevich and Vu [23] gave an algorithm that in the case $p \gg n^{-1/2}$ approximates the independence number of $G_{n,p}$ in polynomial expected time within a factor of $O(\sqrt{np}/\ln(np))$. Moreover, they ask whether a similar algorithm exists for smaller values of p. As a first answer, Coja-Oghlan and Taraz [4], gave an $O(\sqrt{np}/\ln(np))$-approximative algorithm for the case $p \gg \ln(n)^6/n$.

Theorem 5. *Suppose that $c_0/n \le p \le 1/2$. There is an algorithm* `ApproxMIS` *that for any input graph G outputs an independent set of size at least $\frac{\alpha(G)\ln(np)}{c_1\sqrt{np}}$, and which applied to $G_{n,p}$ runs in polynomial expected time. Here $c_0, c_1 > 0$ denote constants.*

As a second application, we give an algorithm for deciding within polynomial expected time whether the input graph is k-colorable. Instead of $G_{n,p}$, we shall even consider the *semirandom model* $G^+_{n,p}$ that allows for an adversary to add edges to the random graph. We say that *the expected running time of an algorithm $\mathcal{A}$ is polynomial over $G^+_{n,p}$*, if there is some constant l such that the expected running time of $\mathcal{A}$ is $O(n^l)$ regardless of the behavior of the adversary.

Theorem 6. *Suppose that $k = o(\sqrt{n})$, and that $p \geq c_0 k^2/n$, for some constant $c_0 > 0$. There exists an algorithm* $\texttt{Decide}_k$ *that for any input graph G decides whether G is k-colorable, and that applied to $G^+_{n,p}$ has a polynomial expected running time.*

The algorithm $\texttt{Decide}_k$ is essentially identical with Krivelevich's algorithm for deciding k-colorability in polynomial expected time [22]. However, the analysis given in [22] requires that $np \geq \exp(\Omega(k))$. The improvement results from the fact that the analysis given in this paper relies on the asymptotics for $\bar{\vartheta}_{1/2}(G_{n,p})$ derived in Thm. 3 (instead of the concept of semi-colorings). Finally, we mention that our algorithm $\texttt{Decide}_k$ also applies to random regular graphs $G_{n,r}$.

Theorem 7. *Suppose that $c_0 k^2 \leq r = o(n^{1/2})$ for some constant $c_0 > 0$. Then, applied to $G_{n,r}$, the algorithm* $\texttt{Decide}_k$ *has polynomial expected running time.*

Notation. Throughout we let $V = \{1, \ldots, n\}$. If $G = (V, E)$ is a graph, then $A(G)$ is the adjacency matrix of G. By $\mathbf{1}$ we denote the vector with all entries $= 1$, and J denotes a square matrix with all entries $= 1$. If M is a real symmetric $n \times n$-matrix, then $\lambda_1(M) \geq \cdots \lambda_n(M)$ signify the eigenvalues of M.

2 Preliminaries

Let $G = (V, E)$ be a graph, let $(v_1, \ldots, v_n)$ be an n-tuple of unit vectors in $\mathbf{R}^n$, and let $k > 1$. Then $(v_1, \ldots, v_n)$ is a *vector k-coloring* of G if $\langle v_i, v_j \rangle \leq -1/(k-1)$ for all edges $\{i, j\} \in E$. Furthermore, $(v_1, \ldots, v_n)$ is a *strict* vector k-coloring if $\langle v_i, v_j \rangle = -1/(k-1)$ for all $\{i, j\} \in E$. Finally, we say that $(v_1, \ldots, v_n)$ is a *rigid* vector k-coloring if $\langle v_i, v_j \rangle = -1/(k-1)$ for all $\{i, j\} \in E$ and $\langle v_i, v_j \rangle \geq -1/(k-1)$ for all $\{i, j\} \notin E$. Following [20,14,3], we define

$$\begin{aligned} \bar{\vartheta}_{1/2}(G) &= \inf\{k > 1 \mid G \text{ admits a vector } k\text{-coloring}\}, \\ \bar{\vartheta}(G) = \bar{\vartheta}_1(G) &= \inf\{k > 1 \mid G \text{ admits a strict vector } k\text{-coloring}\}, \\ \bar{\vartheta}_2(G) &= \inf\{k > 1 \mid G \text{ admits a rigid vector } k\text{-coloring}\}. \end{aligned} \tag{3}$$

Observe that $\bar{\vartheta}_{1/2}(G)$ is precisely the *vector chromatic number* introduced by Karger, Motwani, and Sudan [20]; $\bar{\vartheta}_2$ occurs in [14,27]. Further, we let $\vartheta_{1/2}(G) = \bar{\vartheta}_{1/2}(\bar{G})$, $\vartheta(G) = \vartheta_1(G) = \bar{\vartheta}(\bar{G})$, and $\vartheta_2(G) = \bar{\vartheta}_2(\bar{G})$. It is shown in [20] that the above definition of ϑ is equivalent with Lovász's original definition (cf. [16]).

Proposition 8. *Let $G = (V, E)$ be a graph of order n, and let $S \subset V$. Let $G[S]$ denote the subgraph of G induced on S. Then $\vartheta_i(G) \leq \vartheta_i(G[S]) + \vartheta_i(G[V \setminus S])$.*

It is obvious from the definitions that for any weak subgraph H of G we have $\bar{\vartheta}_i(H) \leq \bar{\vartheta}_i(G)$, $i \in \{1/2, 1, 2\}$. In addition to ϑ, $\vartheta_{1/2}$, and ϑ_2, we consider the semidefinite relaxation of MAX CUT invented by Goemans and Williamson [15]: $\mathrm{SMC}(G) = \max \sum_{i<j} \frac{a_{ij}}{2} (1 - \langle v_i, v_j \rangle)$ s.t. $\|v_i\| = 1$, where the max is taken over $v_1, \ldots, v_n \in \mathbf{R}^n$.

Finally, we need the following concentration result on $\vartheta_{1/2}, \vartheta, \vartheta_2$. For ϑ, the proof can be found in [5]. Using suitable characterizations of $\vartheta_{1/2}$, ϑ_2, the argument given in [5] can be adapted to cover these cases as well.

Theorem 9. *Suppose that $p \leq 0.99$, and that $n \geq n_0$ for a certain constant $n_0 > 0$. Let m be a median of $\vartheta(G_{n,p})$.*

i. Let $\xi \geq \max\{10, m^{1/2}\}$. Then $\mathrm{P}(\vartheta(G_{n,p}) \geq m+\xi) \leq 30\exp(-\xi^2/(5m+10\xi))$.
ii. Let $\xi > 10$. Then $\mathrm{P}(\vartheta(G_{n,p}) \leq m-\xi) \leq 3\exp(-\xi^2/10m)$.

The same holds with ϑ replaced by $\vartheta_{1/2}$ or by ϑ_2.

3 The Concentration Results

Proof of Thm. 1. Let p and β be as in Thm. 1. The proof is based on the following large deviation result, which is a consequence of Azuma's inequality.

Lemma 10. *Suppose that $X : G_{n,p} \to \mathbf{R}$ is a random variable that satisfies the following conditions for all graphs $G = (V, E)$.*

- *For all $v \in V$ the following holds. Let $G^* = G + \{\{v, w\}|\ w \in V,\ w < v\}$, and let $G_* = G - \{\{v, w\}|\ w \in V,\ w < v\}$. Then $|X(G^*) - X(G_*)| \leq 1$.*
- *If H is a weak subgraph of G, then $X(H) \leq X(G)$.*

Then $\mathrm{P}(|X - \mathrm{E}(X)| > t\sqrt{n}) \leq 2\exp(-t^2/2)$.

Let $\omega = \omega(n)$ be a sequence tending to infinity slowly, e.g. $\omega(n) = \ln\ln(n)$. Furthermore, let $k = k(n, p) = \inf\{x > 0|\ \mathrm{P}(\bar{\vartheta}_2(G_{n,p}) \leq x) \geq \omega^{-1}\}$. For any graph $G = (V, E)$ let $Y(G) = \min\{\#U|\ U \subset V,\ \bar{\vartheta}_2(G-U) \leq k\}$. Then $\bar{\vartheta}_2(G) \leq k$ if and only if $Y(G) = 0$. Hence, $\mathrm{P}(Y = 0) \geq \omega^{-1}$. Moreover, by Prop. 8, the random variable Y satisfies the assumptions of L. 10. Let $\mu = \mathrm{E}(Y)$. Then $\mu \leq \sqrt{n}\omega$. Thus, by L. 10, $Y \leq 2\sqrt{n}\omega$ with high probability. The following lemma is implicit in [26] (cf. the proof of L. 8 in [26]).

Lemma 11. *Let $\delta > 0$. Whp. the random graph $G = G_{n,p}$ enjoys the following property. If $U \subset V$, $\#U \leq 2\sqrt{n}\omega$ then $\chi(G[U]) \leq s$, where $s > \frac{2}{2\beta-1} + \delta$.*

To conclude the proof of Thm. 1, let $G = G_{n,p}$, and suppose that there is some $U \subset V$, $\#U \leq 2\sqrt{n}\omega$, such that $\bar{\vartheta}_2(G - U) \leq k \leq \bar{\vartheta}_2(G)$. Since by L. 11 $\bar{\vartheta}_2(G[U]) \leq \chi(G[U]) \leq s$ whp., Prop. 8 entails that $k \leq \bar{\vartheta}_2(G) \leq k + s$ whp.

Proof of Thm. 2. Let ω be a sequence tending to infinity slowly. The random graph $G = G_{n,p}$ admits no $U \subset V$, $\#U \leq \omega^3\sqrt{n}$, spanning more than $3(\#U - \varepsilon)/2$ edges whp., where $\varepsilon > 0$ is a small constant. Let k be as in the proof of Thm. 1. Then whp. there is a set $U \subset V$, $\#U \leq \omega\sqrt{n}$, such that $\bar{\vartheta}_2(G - U) \leq k$. Following Łuczak [25], we let $U = U_0$, and construct a sequence $U_0, \ldots, U_m$ as follows. If there is no edge $\{v, w\} \in E$ with $v, w \in N(U_i) \setminus U_i$, then we let $m = i$ and finish. Otherwise, we let $U_{i+1} = U_i \cup \{v, w\}$ and continue. Then $m \leq$

$m_0 = \omega^2\sqrt{n}$, because otherwise $\#U_{m_0} = (2 + o(1))\omega^2\sqrt{n}$ and $\#E(G[U_{m_0}]) \geq 3(1 - o(1))\#U_{m_0}/2$. Let $R = U_m$.

By L. 11, $\bar{\vartheta}_2(G[R]) \leq \chi(G[R]) \leq 3$. Furthermore, $I = N(R) \setminus R$ is an independent set. Let $G_1 = G[R \cup I]$, $S = V \setminus (R \cup I)$, and $G_2 = G[S \cup I]$. Then $\bar{\vartheta}_2(G_2) \leq k$, and $\bar{\vartheta}_2(G_1) \leq 4$. In order to prove that $\bar{\vartheta}_2(G) \leq k+1$, we shall first construct a rigid vector $k+1$-coloring of G_2 that assigns the same vector to all vertices in I. Thus, let $(x_v)_{v \in S \cup I}$ be a rigid vector k-coloring of G_2. Let x be a unit vector perpendicular to x_v for all $v \in S$. Moreover, let $\alpha = (k^2-1)^{-1/2}$, and set $y_v = (\alpha^2+1)^{-1/2}(x_v - \alpha x)$ for $v \in S$, and $y_v = x$ for $v \in I$. Then $(y_v)_{v \in S \cup I}$ is a rigid vector $(k+1)$-coloring of G_2. In a similar manner, we can construct a rigid vector 4-coloring $(y'_v)_{v \in R \cup I}$ of G_1 that assigns the same vector x' to all vertices in I.

Applying a suitable orthogonal transformation if necessary, we may assume that $x = x'$. Let $l = \max\{4, k+1\}$. Since $N(R) \subset R \cup I$, we obtain a rigid vector l-coloring $(z_v)_{v \in V}$ of G, where $z_v = y_v$ if $v \in S \cup I$, and $z_v = y'_v$ if $v \in R$. By the lower bound on $\bar{\vartheta}_2(G_{n,p})$ in Thm. 3 (which does not rely on Thm. 2 of course), choosing c_0 large enough we may assume that $k \geq 4$, whence $k \leq \bar{\vartheta}_2(G) \leq k+1$.

4 The Probable Value of $\vartheta(G_{n,p})$, $\bar{\vartheta}(G_{n,p})$, etc.

4.1 The Lower Bound on $\bar{\vartheta}_{1/2}(G_{n,p})$

To bound $\bar{\vartheta}_{1/2}(G_{n,p})$ from below, we make use of an estimate on the probable value of the SDP relaxation SMC of MAX CUT (cf. Sec. 2). Suppose that $c_0/n \leq p \leq 1 - c_0/n$ for some large constant $c_0 > 0$. Combining Thms. 4 and 5 of [6] instantly yields that there is a constant $\lambda > 0$ such that

$$\mathrm{P}\left(\mathrm{SMC}(G_{n,p}) > \frac{1}{2}\binom{n}{2}p + \lambda n^{3/2}p^{1/2}(1-p)^{1/2}\right) \leq \exp(-2n). \tag{4}$$

Let $G = (V, E)$ be a graph with adjacency matrix $A = (a_{ij})_{i,j=1,\ldots,n}$. Let $v_1, \ldots, v_n$ be a vector k-coloring of G, where $k = \bar{\vartheta}_{1/2}(G) \geq 2$. Then $\|v_i\| = 1$ for all i, and $\langle v_i, v_j \rangle \leq -1/(k-1)$ whenever $\{i,j\} \in E$. Therefore,

$$\mathrm{SMC}(G) \geq \sum_{i<j} \frac{a_{ij}}{2}(1 - \langle v_i, v_j \rangle) \geq \#E\left(\frac{1}{2} + \frac{1}{k-1}\right). \tag{5}$$

Let $c_0/n \leq p \leq 1 - c_0/n$ for some large constant $c_0 > 0$. By Chernoff bounds (cf. [18, p. 26]),

$$\mathrm{P}\left(\#E(G_{n,p}) < \binom{n}{2}p - 8n^{3/2}p^{1/2}(1-p)^{1/2}\right) \leq \exp(-2n). \tag{6}$$

Combining (4), (5), and (6), we conclude that

$$\bar{\vartheta}_{1/2}(G_{n,p}) \geq \bar{\vartheta}_{1/2}(G_{n,p}) - 1 \geq \frac{\binom{n}{2}p - 8n^{3/2}p^{1/2}(1-p)^{1/2}}{(\lambda+4)n^{3/2}p^{1/2}(1-p)^{1/2}} \geq \frac{1}{2(\lambda+4)}\sqrt{\frac{np}{1-p}}$$

holds with probability at least $1 - \exp(-n)$. As $\bar{G}_{n,p} = G_{n,1-p}$, this proves (2) and the lower bounds in Thm. 3.

4.2 Spectral Considerations

Let us briefly recall Juhász's proof that $\vartheta(G_{n,p}) \leq (2+o(1))\sqrt{n(1-p)/p}$ for constant values of p, say. Given a graph $G=(V,E)$, we consider the matrix $M = M(G) = (m_{ij})_{i,j=1,\ldots,n}$, where

$$m_{ij} = \begin{cases} 1 & \text{if } \{i,j\} \notin E \\ (p-1)/p & \text{otherwise,} \end{cases} \qquad (i \neq j), \tag{7}$$

and $m_{ii} = 1$ for all i. Then $\lambda_1(M) \geq \vartheta(G)$. Moreover, as p is constant, the result of Füredi and Komlos [13] on the eigenvalues of random matrices applies and yields that $\vartheta(G_{n,p}) \leq \lambda_1(M) \leq (2+o(1))\sqrt{n(1-p)/p}$ whp. This argument carries over to the case $\ln(n)^7/n \leq p \leq 1/2$ (cf. [4]):

Lemma 12. *Let $\ln(n)^7/n \leq p \leq 1/2$. Then $\|M(G_{n,p})\| \leq 3\sqrt{n/p}$ whp.*

However, it is easily seen that in the sparse case, e.g. if $np = O(1)$, we have $\lambda_1(M) \gg n$ whp. The reason is that in the case $np \geq \ln(n)^7$ the random graph $G_{n,p}$ is "almost regular", which is not true if $np = O(1)$. We will get around this problem by chopping off all vertices of degree considerably larger than np, as first proposed in [1]. Thus, let $\varepsilon > 0$ be a small constant, and consider the graph $G' = (V', E')$ obtained from $G = G_{n,p}$ by deleting all vertices of degree greater than $(1+\varepsilon)np$.

Lemma 13. *Suppose that $c_0/n \leq p \leq \ln(n)^7/n$ for some large constant c_0. Let $G = G_{n,p}$, and let $M' = M(G')$. Then $\mathrm{P}(\|M'\| \leq c_1\sqrt{n/p}) \geq 9/10$, where $c_1 > 0$ denotes some constant.*

To prove L. 13, we make use of the following lemma, which is implicit in [10, Sections 2 and 3]; the proof is based on the method of Kahn and Szemeredi [11].

Lemma 14. *Let $G = G_{n,p}$ be a random graph, where $c_0/n \leq p \leq \ln(n)^7/n$ for some large constant $c_0 > 0$. Let $n' = \#V(G')$, $e = n'^{-1/2}\mathbf{1} \in \mathbf{R}^{n'}$, and $A' = A(G')$. For each $\delta > 0$ there is a constant $C(\delta) > 0$ such that in the case $np \geq C(\delta)$ with probability $\geq 1-\delta$ we have*

$$\max\{|\langle A'v, e\rangle|, |\langle A'v, w\rangle|\} \leq c_1\sqrt{np} \text{ for all } v, w \perp \mathbf{1},\ \|v\| = \|w\| = 1. \tag{8}$$

Here $c_1 > 0$ denotes a certain constant.

In addition, the proof of Lemma 13 needs the following observation.

Lemma 15. *Let c_1 be a large constant. The probability that in $G = G_{n,p}$ there exists a set $U \subset V$, $\#U \geq n/2$, such that $|\#E(G[U]) - \#U^2p/2| \geq c_1(\#U)^{3/2}p^{1/2}$ is less than $\exp(-n)$.*

Proof. There are at most 2^n sets U. By Chernoff bounds (cf. [18, p. 26]), for a fixed U the probability that $|\#E(G[U]) - \#U^2p/2| \geq c_1(\#U)^{3/2}p^{1/2}$ is at most $\exp(-2n)$, provided that c_0, c_1 are large enough. $\square$

Proof of Lemma 13. Let $G = G_{n,p}$, let $n' = \#V(G')$, and let A', e be as in L. 14. Without loss of generality, we may assume that $V' = V(G') = \{1, \ldots, n'\}$. Let $c_1 > 0$ be a sufficiently large constant. Let J signify the $n' \times n'$ matrix with all entries equal to 1. Letting $\delta > 0$ be sufficiently small and $c_0 \geq C(\delta)$, we assume in the sequel that (8) holds, and that G has the property stated in L. 15. Let $z \in \mathbf{R}^{n'}$, $\|z\| = 1$. Then we have a decomposition $z = \alpha e + \beta v$, $\|v\| = 1$, $v \perp \mathbf{1}$, $\alpha^2 + \beta^2 = 1$. Since $\|M'z\| \leq \|M'e\| + \|M'v\|$, if suffices bound $\max_{v \perp e, \|v\|=1} \|M'v\|$ and $\|M'e\|$.

Let $\rho : \mathbf{R}^{n'} \to \mathbf{R}^{n'}$ be the projection on the space $\mathbf{1}^\perp$. Then $A'v = \rho A'v + \langle A'v, e\rangle e$, whence $\|A'v\| \leq \|\rho A'v\| + c_1\sqrt{np}$, for all unit vectors $v \perp \mathbf{1}$. In order to bound $\|\rho A'v\|$, we estimate $\|\rho A'\rho\|$ via (8):

$$\|\rho A'\rho\| = \sup_{\|y\|=1} |\langle \rho A'\rho y, y\rangle| = \sup_{\|y\|=1} |\langle A'\rho y, \rho y\rangle| = \sup_{\|y\|=1,\ \mathbf{1}\perp y} |\langle A'y, y\rangle| \leq c_1\sqrt{np}.$$

Consequently, $\|M'v\| = \|(J - \frac{1}{p}A')v\| = \frac{1}{p}\|A'v\| \leq 2c_1\sqrt{n/p}$ ($v \perp \mathbf{1}$, $\|v\| = 1$).

To bound $\|M'e\|$, note that $-pM' = A' - pJ$. Let $\bar{d} = 2\#E(G')/n'$, and $x = A'e - (\bar{d}/n')Je$. Then $x \perp \mathbf{1}$, and by (8) we have $\|x\|^2 = \langle A'e, x\rangle - \langle(\bar{d}/n')Je, x\rangle = \langle A'e, x\rangle \leq c_1\sqrt{np}\|x\|$, whence $\|x\| \leq c_1\sqrt{np}$. By L. 15, $|\bar{d} - n'p| \leq c_1\sqrt{np}$. As a consequence, $\|(\bar{d}/n')Je - pJe\| \leq c_1\sqrt{np}$. Therefore, $\|pM'e\| \leq \|x\| + \|(\bar{d}/n')Je - pJe\| \leq 2c_1\sqrt{np}$, i.e. $\|M'e\| \leq 2c_1\sqrt{n/p}$. □

4.3 Bounding $\vartheta_2(G_{n,p})$ from Above

Let $c_0/n \leq p \leq 1/2$ for some large constant $c_0 > 0$. The following lemma is a consequence of the characterization of $\bar{\vartheta}_2$ as an eigenvalue minimization problem given in [27].

Lemma 16. *Let G be any graph. Let $M = M(G)$. Then $\lambda_1(M) \geq \bar{\vartheta}_2(G)$.*

In the case $\ln(n)^7/n \leq p \leq 1/2$, combining L. 12 and L. 16 yields that $\vartheta_2(G_{n,p}) \leq c_2\sqrt{n/p}$ whp. for some constant $c_2 > 0$, as desired. Thus, let us assume that $c_0/n \leq p \leq \ln(n)^7/n$ in the sequel. Let $\varepsilon > 0$ be a small constant.

Lemma 17. *With probability at least $9/10$ the random $G_{n,p}$ has at most $1/p$ vertices of degree greater than $(1 + \varepsilon)np$.*

Proof. For each vertex v of $G_{n,p}$, the degree $d(v)$ is binomially distributed with mean $(n - 1)p$. By Chernoff bounds (cf. [18, p. 26]), the probability that $d(v) > (1 + \varepsilon)np$ is at most $\exp(-\varepsilon^2 np/100)$. Hence, the expected number of vertices v such that $d(v) > (1 + \varepsilon)np$ is at most $n\exp(-\varepsilon^2 np/100) < 1/(10p)$, provided $np \geq c_0$ for some large constant $c_0 > 0$. Therefore, the assertion follows from Markov's inequality. □

Let $G = G_{n,p}$, and let $G' = (V', E')$ be the graph obtained from G by deleting all vertices of degree greater than $(1 + \varepsilon)np$. Let $V'' = V \setminus V'$, and $G'' = G[V'']$. Combining L. 17 and L. 13, we obtain that

$$\mathrm{P}\left(\vartheta_2(G') \leq c_2\sqrt{n/p} \text{ and } \vartheta_2(G'') \leq \#V(G'') \leq 1/p \leq \sqrt{n/p}\right) > 1/2,$$

where c_2 denotes a suitable constant. Consequently, Prop. 8 yields that

$$\mathrm{P}(\vartheta_2(G_{n,p}) \leq (c_2+1)\sqrt{n/p}) > 1/2.$$

Let $\mu = (c_2+1)\sqrt{n/p}$, $t = \ln(n)\sqrt{n}$, and note that $t = o(\sqrt{n/p})$. Then, by Thm. 9, $\mathrm{P}(\vartheta_2(G_{n,p}) > \mu + t) \leq 30\exp\left(-\Omega(\ln(n)^2)\right) = o(1)$. Since $t < \sqrt{n/p}$, we get that $\vartheta_2(G_{n,p}) \leq (c_2+2)\sqrt{n/p}$ with high probability.

4.4 Bounding $\bar{\vartheta}_2(G_{n,p})$ from Above

Let us first assume that $\ln(n)^7/n \leq p \leq 1/2$. Let $G = (V,E) = G_{n,p}$ be a random graph, and consider the matrix $\bar{M} = \frac{1}{1-p}E_n - \frac{p}{1-p}M(G)$, where E_n is the $n \times n$-unit matrix, and $M(G)$ is the matrix defined in (7). Combining L. 12 and L. 16, we have $\bar{\vartheta}_2(G) \leq \lambda_1(\bar{M}) \leq \|\frac{1}{1-p}E - \frac{p}{1-p}M'\| \leq \frac{p}{1-p}\|M\| + 2 \leq c_4\sqrt{np}$ whp., where $c_4 > 0$ is a certain constant.

Now let $c_0/n \leq p \leq \ln(n)^7/n$ for some large constant $c_0 > 0$. In this case, the proof of our upper bound on $\bar{\vartheta}_2(G_{n,p})$ relies on the concentration result Thm. 2.

Lemma 18. *Whp. the random graph $G = G_{n,p}$ admits no set $U \subset V$, $\#U \leq 1/p$, such that $\chi(G[U]) > \sqrt{np}$.*

Proof. We shall prove that for all $U \subset V$, $\#U = \nu \leq 1/p$, we have $\#E(G[U]) < \nu\sqrt{np}/2$. Then each subgraph $G[U]$ has a vertex of degree $< \sqrt{np}$, a fact which immediately implies our assertion. Thus, let $\nu \leq 1/p$. The probability that there exists some $U \subset V$, $\#U = \nu$, $\#E(G[U]) \geq \nu\sqrt{np}/2$, is at most

$$\binom{n}{\nu}\binom{\binom{\nu}{2}}{\nu\sqrt{np}/2}p^{\nu\sqrt{np}/2} \leq \left(\frac{\mathrm{e}n}{\nu}\left(\frac{\mathrm{e}\nu\sqrt{p}}{\sqrt{n}}\right)^{\sqrt{np}/2}\right)^{\nu}$$

Let $b_\nu = (\mathrm{e}n/\nu)(\mathrm{e}\nu\sqrt{p}/\sqrt{n})^{\sqrt{np}/2}$. Observe that the sequence $(b_\nu)_{\nu=1,\ldots,n}$ is monotone increasing, and that $b_{1/p} = \mathrm{e}np(\mathrm{e}/\sqrt{np})^{\sqrt{np}/2} \leq \exp(-2)$. Therefore, $\sum_{\nu=\ln(n)}^{1/p} b_\nu^\nu \leq b_{1/p}^{\ln(n)}/p \leq n^{-2}p^{-1} = o(1)$. Moreover, if $\nu \leq \ln(n)$, then $b_\nu \leq \mathrm{e}n\nu^{-1}(\mathrm{e}\nu\sqrt{p}/\sqrt{n})^{\sqrt{np}/2} \leq 1/n$, whence $\sum_{\nu=1}^{\ln n} b_\nu^\nu = o(1)$. Thus, $\sum_{\nu=1}^{1/p} b_\nu^\nu = o(1)$, thereby proving the lemma. □

Let $G = (V,E) = G_{n,p}$ be a random graph, and let $G' = (V',E')$ be the graph obtained from G by removing all vertices of degree greater than $(1+\varepsilon)np$, where $\varepsilon > 0$ is small but constant. Let $V'' = V \setminus V'$, and let $G'' = G[V'']$. By L. 17, with probability at least $9/10$ we have $\#V'' \leq 1/p$. Therefore, by L. 18, $\mathrm{P}(\bar{\vartheta}_2(G'') \leq \sqrt{np}) \geq \mathrm{P}(\chi(G'') \leq \sqrt{np}) \geq 9/11$. To bound $\bar{\vartheta}_2(G')$, we consider the matrix $\bar{M} = \frac{1}{1-p}E_{n'} - \frac{p}{1-p}M(G')$. By L. 16, $\bar{\vartheta}_2(G') \leq \lambda_1(\bar{M})$. Moreover, by L. 13, with probability $\geq 9/10$ we have $\bar{\vartheta}_2(G') \leq \lambda_1(\bar{M}) \leq \frac{p}{1-p}\|M'\| + 2 \leq c_4\sqrt{np}$, for some constant $c_4 > 0$. Prop. 8 implies that $\bar{\vartheta}_2(G) \leq \bar{\vartheta}_2(G') + \bar{\vartheta}_2(G'')$, whence we conclude that $\mathrm{P}(\bar{\vartheta}_2(G_{n,p}) \leq (c_4+1)\sqrt{np}) > 1/2$. Since Thm. 2 shows that $\bar{\vartheta}_2(G_{n,p})$ is concentrated in width one, we have

$$\mathrm{P}\left(\bar{\vartheta}_{1/2}(G_{n,p}) \leq \bar{\vartheta}(G_{n,p}) \leq \bar{\vartheta}_2(G_{n,p}) \leq (c_4+1)\sqrt{np} + 1\right) = 1 - o(1),$$

thereby completing the proof of Thm. 3.

Remark 19. One could prove slightly weaker results on the probable value of $\vartheta(G_{n,p})$ and $\bar{\vartheta}(G_{n,p})$ than provided by Thm. 3 without applying any concentration results, or bounds on the SDP relaxation SMC of MAX CUT. Indeed, using only L. 17, 18, 13 (thus implicitly [10]) and the estimates proposed in [19], one could show that for each $\delta > 0$ there is $C(\delta) > 0$ such that $\mathrm{P}(c_1\sqrt{n/p} \leq \vartheta(G_{n,p}) \leq c_2\sqrt{n/p}) \geq 1-\delta$ and $\mathrm{P}(c_3\sqrt{np} \leq \bar{\vartheta}(G_{n,p}) \leq c_4\sqrt{np}) \geq 1-\delta$, provided $np \geq C(\delta)$. Such an approach is mentioned without proof independently in the latest version of [10].

5 Approximating the Independence Number and Deciding *k*-Colorability

Approximating the Independence Number. The algorithm `ApproxMIS` for approximating the independence number consists of two parts. First, we employ a certain greedy procedure that on input $G = G_{n,p}$ finds a large independent set whp. Secondly, we compute $\vartheta(G)$ to bound $\alpha(G)$ from above. Following [23], to find a large independent set of $G = G_{n,p}$, we run the greedy algorithm for graph coloring and pick the largest color class it produces.

Lemma 20. *The probability that the largest color class produced by the greedy coloring algorithm contains* $< \ln(np)/(2p)$ *vertices is at most* $\exp(-n)$.

Proof. The proof given in [23] for the case that $p \geq n^{\varepsilon-1/2}$ carries over. □

The following algorithm is essentially identical with the one given in [4].

Algorithm 21. `ApproxMIS`(G)
Input: A graph $G = (V, E)$. *Output:* An independent set of G.

1. Run the greedy algorithm for graph coloring on input G. Let I be the largest resulting color class. If $\#I < \ln(np)/(2p)$, then go to 5.
2. Compute $\vartheta(G)$. If $\vartheta(G) \leq C\sqrt{n/p}$, then output I and terminate. Here C denotes some sufficiently large constant (cf. the analysis below).
3. Check whether there exists a subset S of V, $\#S = 25\ln(np)/p$, such that $\#V \setminus (S \cup N(S)) > 12(n/p)^{1/2}$. If no such set exists, then output I and terminate.
4. Check whether in G there is an independent set of size $12(n/p)^{1/2}$. If this is not the case, then output I and terminate.
5. Enumerate all subsets of V and output a maximum independent set.

Lemma 22. *The expected running time of* `ApproxMIS`$(G_{n,p})$ *is polynomial.*

Proof. The first two steps can be implemented in polynomial time. By Thm. 3, the median μ of $\vartheta(G_{n,p})$ is at most $c\sqrt{n/p}$, for some constant c. Therefore, Thm. 9 entails that the probability that `ApproxMIS` runs step 3 is less than $\exp(-(n/p)^{1/2})$, provided C is large enough. Furthermore, up to polynomial factors, step 3 consumes time $\leq \exp(25\ln(np)^2/p) < \exp(\sqrt{n/p})$. Hence, the expected time spent executing step 3 is polynomial. Taking into account L. 20, the expected running time of the remaining steps can be estimated as in the proof of Thm. 4 in [4]. □

Finally, it is not hard to show that `ApproxMIS` guarantees the desired approximation ratio.

Deciding *k*-Colorability. Following [22], we decide k-colorability by computing the vector chromatic number of the input graph. Let $k = k(n)$ be a sequence of positive integers such that $k(n) = o(\sqrt{n})$. Since the vector chromatic number is always a lower bound on the chromatic number, the answer of the following algorithm is correct for all input graphs G.

Algorithm 23. $\texttt{Decide}_k(G)$
Input: A graph $G = (V, E)$. *Output:* Either "$\chi(G) \leq k$" or "$\chi(G) > k$".

1. If $\bar{\vartheta}_{1/2}(G) > k$ then terminate with output "$\chi(G) > k$".
2. Otherwise, compute $\chi(G)$ in time $o(\exp(n))$ using Lawler's algorithm [24], and answer correctly.

Lemma 24. *Suppose that $p \geq Ck^2/n$ for some large constant C. Then the expected running time of $\texttt{Decide}_k(G^+_{n,p})$ is polynomial.*

Proof. In [20] it is shown that $\bar{\vartheta}_{1/2}$ can be computed in polynomial time. Since the second step consumes time $o(\exp(n))$, (2) shows that the expected running time of $\texttt{Decide}_k$ on input $G^+_{n,p}$ is polynomial. □

The analysis of $\texttt{Decide}_k$ on input $G_{n,r}$, $r \geq Ck^2$, is based on Thm. 4 and yields the proof of Thm. 7.

Acknowledgment. I am grateful to M. Krivelevich and C. Helmberg for helpful discussions, and to U. Feige and E. Ofek for providing me with their technical report [10].

References

1. Alon, N., Kahale, N.: A spectral technique for coloring random 3-colorable graphs. SIAM J. Comput. **26** (1997) 1733–1748
2. Alon, N., Krivelevich, M.: The concentration of the chromatic number of random graphs. Combinatorica **17** 303–313
3. Charikar, M.: On semidefinite programming relaxations for graph coloring and vertex cover. Proc. 13th SODA (2002) 616–620
4. Coja-Oghlan, A., Taraz, A.: Exact and approximative algorithms for colouring $G(n, p)$. preprint (available from http://www.informatik.hu-berlin.de/∼coja/). A preliminary version has appeard in Proc. 20th STACS (2003) 487–498

5. Coja-Oghlan, A.: Finding large independent sets in polynomial expected time. Proc. 20th STACS (2003) 511–522
6. Coja-Oghlan, A., Moore, C., Sanwalani, V.: MAX k-CUT and approximating the chromatic number of random graphs. available from http://www.informatik.hu-berlin.de/∼coja/. An extended abstract version has appeared in Proc. ICALP 2003.
7. Feige, U.: Randomized graph products, chromatic numbers, and the Lovász theta function. Combinatorica **17**(1) 79–90
8. Feige, U., Kilian, J.: Heuristics for semirandom graph problems. J. Comput. and System Sci. **63** (2001) 639–671
9. Feige, U. and Kilian, J.: Zero knowledge and the chromatic number. Proc. 11th IEEE Conf. Comput. Complexity (1996) 278–287.
10. Feige, U., Ofek, E.: Spectral techniques applied to sparse random graphs, report MCS03-01, Weizmann Institute (2003) (available from http://www.wisdom.weizmann.ac.il/math/research.shtml)
11. Friedman, J., Kahn, J., Szemeredi, E.: On the second eigenvalue in random regular graphs. Proc. 21st STOC (1989) 587–598.
12. Frieze, A., McDiarmid, C.: Algorithmic theory of random graphs. Random Structures & Algorithms **10** (1997) 5–42
13. Füredi, Z., Komloś, J.: The eigenvalues of random symmetric matrices, Combinatorica **1** (1981) 233–241
14. Goemans, M.X., Kleinberg, J.: The Lovasz theta function and a semidefinite programming relaxation of vertex cover. SIAM J. on Discrete Math. **11** (1998) 1–48
15. Goemans, M.X., Williamson, D.P.: Improved approximation algorithms for maximum cut and satisfyability problems using semidefinite programming. J. ACM **42** (1995) 1115–1145.
16. Grötschel, M., Lovász, L., Schrijver, A.: Geometric algorithms and combinatorial optimization. Springer (1988)
17. Håstad, J.: Clique is hard to approximate within $n^{1-\varepsilon}$. Proc. 37th FOCS (1996) 627–636
18. Janson, S., Łuczak, T., Ruciński, A.: Random Graphs. Wiley (2000)
19. Juhász, F.: The asymptotic behaviour of Lovász ϑ function for random graphs. Combinatorica **2** (1982) 269–270
20. Karger, D., Motwani, R., Sudan, M.: Approximate graph coloring by semidefinite programming. J. ACM **45** (1998) 246–265
21. Knuth, D.: The sandwich theorem, Electron. J. Combin. **1** (1994)
22. Krivelevich, M.: Deciding k-colorability in expected polynomial time, Information Processing Letters **81** (2002), 1–6
23. Krivelevich, M., Vu, V.H.: Approximating the independence number and the chromatic number in expected polynomial time. J. of Combinatorial Optimization **6** (2002) 143–155
24. Lawler, E.L.: A note on the complexity of the chromatic number problem, Information Processing Letters **5** (1976) 66–67
25. Łuczak, T.: A note on the sharp concentration of the chromatic number of random graphs. Combinatorica **11** (1991) 45–54
26. Shamir, E., Spencer, J.: Sharp concentration of the chromatic number of random graphs $G_{n,p}$. Combinatorica **7** (1987) 121–129
27. Szegedy, M.: A note on the θ number of Lovász and the generalized Delsarte bound. Proc. 35th FOCS (1994) 36–39