FLIPS IN GRAPHS

TOM BOHMAN∗†, ANDRZEJ DUDEK∗, ALAN FRIEZE∗‡, AND OLEG PIKHURKO∗§

Abstract. We study a problem motivated by a question related to quantum-error-correcting codes. Combinatorially, it involves the following graph parameter:

\[ f(G) = \min \{|A| + \{|x \in V \setminus A : d_A(x) \text{ is odd}\} : A \neq \emptyset\}, \]

where \( V \) is the vertex set of \( G \) and \( d_A(x) \) is the number of neighbors of \( x \) in \( A \). We give asymptotically tight estimates of \( f \) for the random graph \( G_{n,p} \) when \( p \) is constant. Also, if \( f(n) = \max\{f(G) : |V(G)| = n\} \)

then we show that \( f(n) \leq (0.382 + o(1))n \).

Key words. quantum-error-correcting codes, random graphs

AMS subject classifications. 94B25, 05C80

1. Introduction. In this paper we consider a problem which is motivated by a question from quantum-error-correcting codes.

Given a graph \( G \) with \( \pm 1 \) signs on vertices, each vertex can perform at most one of the following three operations: \( O_1 \) (flip all neighbors, i.e., change their signs), \( O_2 \) (flip oneself), and \( O_3 \) (flip oneself and all neighbors). We want to start with all +1’s, execute some non-zero number of operations and return to all +1’s. The diagonal distance \( f(G) \) is the minimum number of operations needed (with each vertex doing at most one operation).

Trivially,

\[ f(G) \leq \delta(G) + 1 \quad (1.1) \]

holds, where \( \delta(G) \) denotes the minimum degree. Indeed, a vertex with the minimum degree applies \( O_1 \) and then its neighbors fix themselves applying \( O_2 \). Let

\[ f(n) = \max f(G), \]

where the maximum is taken over all non-empty graphs of order \( n \).

Given a graph \( G \), one can ultimately construct a quantum error correcting code, see [3, 5, 6]. A common metric to measure the code robustness against noise is the quantity called “code distance” which is bounded from above by \( f(G) \). Although it is more important to find explicit graphs \( G \) with large \( f(G) \) (see the case \( k = 0 \) of Section “QECC” in [2] for known constructions), theoretical upper and lower bounds on \( f(n) \) are also of interest.

In this paper we asymptotically determine the diagonal distance of the random graph \( G_{n,p} \) for any \( p \in (0, 1) \).

We denote the symmetric difference of two sets \( A \) and \( B \) by \( A \triangle B \) and the logarithmic function with base \( e \) as \( \log \).

THEOREM 1.1. There are absolute constants \( \lambda_0 \approx 0.189 \) and \( p_0 \approx 0.894 \), see (2.4) and (3.3), such that for \( G = G_{n,p} \) asymptotically almost surely:

∗Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, {tbohman,adudek,af1p,pikhurko}@andrew.cmu.edu
†Research partially supported by NSF grant DMS-0401147
‡Research partially supported by NSF grant DMS-0753472
§Research partially supported by NSF grant DMS-0758057
Hence, if \( l \) is binomial with parameters \( b \) and \( f \), then asymptotically almost surely the diagonal distance of \( p \) when \( \lambda \) is at least \( l \).

Let \( \lambda = l/n \) and \( a = a/n \). We may assume that \( \lambda < \frac{1}{4} \). Consequently, \( \lambda - a < \frac{1}{2} (1 - a) \), and hence, we can approximate the summand in (2.1) by

\[
g_n (H(\alpha) + (1 - \alpha) (H(\frac{\lambda - a}{1 - \alpha} - 1) + O(\log n/n)),
\]

where \( H \) is the binary entropy function defined as \( H(p) = -p \log_2 p - (1 - p) \log_2 (1 - p) \). For more information about the entropy function and its properties see, e.g., [1]. Let

\[
g_\lambda (\alpha) = H(\alpha) + (1 - \alpha) \left( H \left( \frac{\lambda - a}{1 - \alpha} - 1 \right) \right).
\]

The maximum of \( g_\lambda (\alpha) \) is attained exactly for \( \alpha = 2 \lambda / 3 \), since

\[
g_\lambda'(\alpha) = \log_2 \frac{2(\lambda - \alpha)}{\alpha}.
\]
Now the function
\[ h(\lambda) = g_\lambda(2\lambda/3) \]

is concave on \( \lambda \in [0,1] \) since
\[ h''(\lambda) = \frac{1}{(\lambda-1)\lambda \log 2} < 0. \]

Moreover, observe that \( h(0) = -1 \) and \( h(1) = H(2/3) - 1/3 > 0. \) Thus the equation \( h(\lambda) = 0 \) has a unique solution \( \lambda_0 \) and one can compute that
\[ \lambda_0 = 0.1892896249152306 \ldots \]

Therefore, if \( \lambda = \lambda_0 - K \log n/n \) for large enough \( K > 0, \) then the left hand side of (2.1) goes to zero and similarly for \( \lambda = \lambda_0 + K \log n/n \) it goes to infinity. In particular, \( f(G) > (\lambda_0 - o(1))n \) asymptotically almost surely.

Let us show that this constant \( \lambda_0 \) is best possible, i.e., asymptotically almost surely \( f(G) \leq (\lambda_0 + K \log n/n)n. \) Let \( \lambda = \lambda_0 + K \log n/n, \) \( n \) be large, and \( l = \lambda n. \) Let \( \alpha = 2\lambda/3 \) and \( a = [\alpha n]. \) We pick a random \( a \)-set \( A \subset V \) and compute \( b. \) Let \( X_A \) be an indicator random variable so that \( X_A = 1 \) if and only if \( b = b(A) \leq l - a. \) Let \( X = \sum_{|A|=a} X_A. \) We succeed if \( X > 0. \)

The expectation \( E(X) = \binom{n}{a} \text{Pr}(\text{Bin}(n-a, 1/2) \leq l-a) \) tends to infinity, by our choice of \( \lambda. \) We now show that \( X > 0 \) asymptotically almost surely by using the Chebyshev inequality. First note that for \( A \cap C \neq \emptyset \) we have
\[ \text{Cov}(X_A, X_C) = \text{Pr}(X_A = X_C = 1) - \text{Pr}(X_A = 1) \text{Pr}(X_C = 1) = 0. \]

Indeed, if \( x \in V \setminus (A \cup C) \), then \( \text{Pr}(x \in B(A)|X_C = 1) = 1/2, \) since \( A \setminus C \neq \emptyset \) and no adjacency between \( x \) and all vertices in \( A \setminus C \) is exposed by the event \( X_C = 1. \)

Similarly, if \( x \in C \setminus A, \) then \( A \cap C \neq \emptyset \) and an adjacency between \( x \) and \( A \cap C \) is independent of the occurrence of \( X_C = 1. \) This implies that \( \text{Pr}(x \in B(A)|X_C = 1) = 1/2 \) as well. Thus \( \text{Pr}(X_A = 1|X_C = 1) = \text{Pr}(\text{Bin}(n-a, 1/2) \leq l-a) = \text{Pr}(X_A = 1), \) and consequently, \( \text{Cov}(X_A, X_C) = 0. \)

Now consider the case when \( A \cap C = \emptyset. \) Let \( s \) be a vertex in \( A. \) Define a new indicator random variable \( Y \) which takes the value \( 1 \) if and only if \( |B(C) \setminus \{s\}| \leq l-a. \) Observe that
\[ \text{Pr}(Y = 1) = \text{Pr}(\text{Bin}(n-a-1, 1/2) \leq l-a) \]
\[ \leq 2 \text{Pr}(\text{Bin}(n-a, 1/2) \leq l-a) = 2 \text{Pr}(X_A = 1). \]

Moreover,
\[ \text{Pr}(X_A = 1|Y = 1) = \text{Pr}(\text{Bin}(n-a, 1/2) \leq l-a) = \text{Pr}(X_A = 1), \]

since for every \( x \in V \setminus A \) the adjacency between \( x \) and \( s \) is not influenced by \( Y = 1. \) Finally note that \( X_C \leq Y. \) Thus,
\[ \text{Cov}(X_A, X_C) \leq \text{Pr}(X_A = X_C = 1) \]
\[ \leq \text{Pr}(X_A = Y = 1) = \text{Pr}(Y = 1) \text{Pr}(X_A = 1|Y = 1) \leq 2(\text{Pr}(X_A = 1))^2. \]
Consequently,

\[
\text{Var}(X) = E(X) + \sum_{A \cap C \neq \emptyset, A \neq C} \text{Cov}(X_A, X_C) + \sum_{A \cap C = \emptyset} \text{Cov}(X_A, X_C)
\]

\[
\leq E(X) + 2 \sum_{A \cap C = \emptyset} (\Pr(X_A = 1))^2
\]

\[
= E(X) + 2 \binom{n}{a} \left(\frac{n-a}{a}\right) (\Pr(X_A = 1))^2 = o(E(X)^2),
\]

as \(E(X) = \binom{n}{a} \Pr(X_A = 1)\) tends to infinity and \(\binom{n-a}{a} = o\left(\binom{n}{a}\right)\). Hence, Chebyshev’s inequality yields that \(X > 0\) asymptotically almost surely.

**Remark 2.1.** A version of the well-known Gilbert-Varshamov bound (see, e.g., [4]) states that if

\[
2^{-n} \sum_{i=1}^{\frac{n}{2}} \binom{n}{i} 3^i < 1,
\]

then \(f(n) \geq 1\). Observe that this is consistent with bound (2.1). Let \(\lambda = \frac{1}{n}\). We can approximate the left hand side of (2.5) by

\[
2^{n(H(\lambda) + \lambda \log_2 3 - 1 + o(1))}.
\]

One can check after some computation that

\[
H(\lambda) + \lambda \log_2 3 - 1 = g_\lambda(2\lambda/3).
\]

Therefore, (2.1) and (2.5) give asymptotically the same lower bound on \(f(n)\).

### 3. Random Graphs for Arbitrary \(p\)

Let \(G = G_{n,p}\) be a random graph with \(p \in (0, 1)\).

Observe that for a fixed set \(A \subset V\), \(|A| = a\), the probability that a vertex from \(V \setminus A\) belongs to \(B(A)\) is

\[
p(a) = \sum_{0 \leq i < \frac{n}{2}} \left(\frac{a}{2i+1}\right) p^{2i+1} (1-p)^a = \frac{1 - (1-2p)^a}{2}.
\]

(If this is unfamiliar, write \(1 - (1-2p)^a = ((1-p) + p)^a - ((1-p) - p)^a\) and expand.)

#### 3.1. \(0 < p < \lambda_0\)

For \(p < \lambda_0\) we begin with the upper bound \(f(G) \leq \delta(G) + 1\), see (1.1). For the lower bound it is enough to show that

\[
\sum_{2 \leq a \leq pn} \binom{n}{a} \Pr(\text{Bin}(n-a, p(a)) \leq pn - a) = o(1),
\]

since \(\delta(G) + 1 \leq np\) asymptotically almost surely. (We may assume that \(p = \Omega\left(\frac{\log n}{n}\right)\); for otherwise \(\delta(G) = 0\) with high probability and the theorem is trivially true.) This implies that with high probability if \(|A| + |B| \leq pn\), then \(|A| = 1\).
3.1.1. **p Constant.** We split this sum into two sums for \(2 \leq a \leq \sqrt{n}\) and \(\sqrt{n} < a \leq pn\), respectively. Let \(X = Bin(n, a, p(a))\) and
\[
\varepsilon = 1 - \frac{pm - a}{(n - a)p(a)} \geq 1 - \frac{p}{p(2)} = 1 - \frac{1}{2 - 2p} > 0.
\]
We will use the following version of Chernoff’s bound,
\[
\Pr(Bin(N, \rho) \leq (1 - \theta)N \rho) \leq e^{-\theta^2N\rho/2}.
\]
Hence, we see that
\[
\Pr(Bin(n - a, p(a)) \leq pn - a) = \Pr(X \leq (1 - \varepsilon)E(X)) \leq \exp\{-\varepsilon^2E(X)/2\} = \exp\{-\Theta(n)\},
\]
and consequently,
\[
\sum_{2 \leq a < \sqrt{n}} \binom{n}{a} \Pr(Bin(n - a, p(a)) \leq pn - a)
\leq \sqrt{n} \binom{n}{\sqrt{n}} \exp\{-\Theta(n)\} \leq \exp\{O(\sqrt{n} \log n)\} \exp\{-\Theta(n)\} = o(1).
\]
Now we bound the second sum corresponding to \(\sqrt{n} < a \leq pn\). Note that
\[
\sum_{\sqrt{n} \leq a \leq pn} \binom{n}{a} \Pr(Bin(n - a, p(a)) \leq pn - a)
= \sum_{\sqrt{n} \leq a \leq pn} \binom{n}{a} \Pr\left(Bin\left(n - a, \frac{1}{2} + e^{-\Omega(n^{1/2})}\right) \leq pn - a\right)
\leq n2^{h(p(\rho)) + o(1)} = o(1).
\]
Here \(h\) is defined in (2.3) and the right hand limit is zero since \(p < \lambda_0\).

3.1.2. **p = o(1).** We follow basically the same strategy as above and show that (3.1) holds for large \(a\) and something similar when \(a\) is small. Suppose then that \(p = 1/\omega\) where \(\omega = \omega(n) \to \infty\). First consider those \(a\) for which \(ap \geq 1/\omega^{1/2}\). In this case \(p(a) \geq (1 - e^{-2ap})/2\). Thus,
\[
\sum_{ap \geq 1/\omega^{1/2}} \binom{n}{a} \Pr(Bin(n - a, p(a)) \leq pn - a)
= \sum_{ap \geq 1/\omega^{1/2}} \exp\{O(n \log \omega/\omega)\} e^{-\Omega(n/\omega^{1/2})} = o(1).
\]
If \(ap \leq 1/\omega^{1/2}\) then \(p(a) = ap(1 + O(ap))\). Then
\[
\sum_{ap \leq 1/\omega^{1/2}} \binom{n}{a} \Pr(Bin(n - a, p(a)) \leq pm - a) \leq \sum_{ap \leq 1/\omega^{1/2}} \frac{ne^{-np/10}}{a} a = o(1)
\]
provided \( np \geq 11 \log n \).

If \( np \leq \log n - \log \log n \) then \( G = G_{n,p} \) has isolated vertices asymptotically almost surely and then \( f(G) = 1 \). So we are left with the case where \( \log n - \log \log n \leq np \leq 11 \log n \).

We next observe that if there is a set \( A \) for which \( 2 \leq |A| \) and \( |A| + |B(A)| \leq np \) then there is a minimal size such set. Let \( H_A = (A, E_A) \) be a graph with vertex set \( A \) and an edge \((v, w) \in E_A\) if and only if \( v, w \) have a common neighbor in \( G \). \( H_A \) must be connected, else \( A \) is not minimal. So we can find \( t \leq a - 1 \) vertices \( T \) such that \( A \cup T \) spans at least \( t + a - 1 \) edges between \( A \) and \( T \). Thus we can replace the estimate (3.2) by

\[
\sum_{a p \leq 1 / \omega_{1/2}} \sum_{2 \leq a \leq np} \binom{n}{a} \binom{n}{t} \binom{a}{t + a - 1} p^{t + a - 1} \Pr (\text{Bin}(n - a - t, p(a)) \leq pm - a)
\]

\[
\leq \sum_{a p \leq 1 / \omega_{1/2}} \sum_{2 \leq a \leq np} \binom{ne}{a} \binom{ne}{t} \binom{a}{t + a - 1} e^{-anp/10}
\]

\[
\leq \frac{1}{2^{e^2np}} \sum_{2 \leq a \leq np} a \left( (e^2np)^2 e^{-np/10} \right)^a = o(1).
\]

### 3.2. \( p_0 < p < 1 \)

First let us define the constant \( p_0 \). Let

\[ p_0 \approx 0.8941512242051071 \ldots \] (3.3)

be a root of \( 2p - 2p^2 = \lambda_0 \). For the upper bound let \( A = \{ x, y \} \), where \( x \) and \( y \) satisfy \( |N(x) \triangle N(y)| \leq |N(x') \triangle N(y')| \) for any \( x', y' \in V(G) \). Then \( B = B(A) = N(x) \triangle N(y) \), and thus, asymptotically almost surely \(|B| \leq (2p-2p^2)n\) plus a negligible error term \( o(n) \). (We may assume that \( 1 - p = \Omega \left( \frac{\log n}{n} \right) \); for otherwise we have two vertices of degree \( n - 1 \) with high probability, and hence, \( f(G)=2 \).)

To show the lower bound it is enough to prove that

\[
\sum_{3 \leq a \leq (2p-2p^2)n} \binom{n}{a} \Pr (\text{Bin}(n - a, p(a)) \leq (2p - 2p^2)n - a) = o(1).
\]

Indeed, this implies that if \(|A| + |B| \leq (2p - 2p^2)n\), then \(|A| = 1 \) or \( 2 \). But if \(|A| = 1\), then in a typical graph \(|B| = (p + o(1))n > (2p - 2p^2)n\) since \( p > 1/2 \).

### 3.2.1. \( p \) Constant.

As in the previous section we split the sum into two sums for \( 3 \leq a \leq \sqrt{n} \) and \( \sqrt{n} < a \leq pn \), respectively. Let

\[ \varepsilon = 1 - \frac{(2p - 2p^2)n - a}{(n - a)p(a)} \geq 1 - \frac{2p - 2p^2}{p(a)} > 0. \]

To confirm the second inequality we have to consider two cases. The first one is for \( a \) odd and at least \( 3 \). Here,

\[ 1 - \frac{2p - 2p^2}{p(a)} > 1 - \frac{2p - 2p^2}{1/2} = (2p - 1)^2 > 0. \]
The second case, for \( a \) even and at least 4, gives
\[
1 - \frac{2p - 2p^2}{p(a)} > 1 - \frac{2p - 2p^2}{p(2)} = 0.
\]
Now one can apply Chernoff bounds with the given \( \varepsilon \) to show that
\[
\sum_{3 \leq a \leq \sqrt{n}} \binom{n}{a} \Pr \left( \text{Bin}(n - a, p(a)) \leq (2p - 2p^2)n - a \right) = o(1).
\]
Now we bound the second sum corresponding to \( \sqrt{n} < a \leq (2p - 2p^2)n \). Note that
\[
\sum_{\sqrt{n} \leq a \leq (2p - 2p^2)n} \binom{n}{a} \Pr \left( \text{Bin}(n - a, p(a)) \leq (2p - 2p^2)n - a \right)
= \sum_{\sqrt{n} \leq a \leq (2p - 2p^2)n} \binom{n}{a} \Pr \left( \text{Bin} \left( n - a, \frac{1}{2} + O(\varepsilon^{-(n/2)}) \right) \leq (2p - 2p^2)n - a \right)
\leq n2^{nh(2p-2p^2)+o(1)} = o(1)
\]
since \( p > p_0 \) implies that \( 2p - 2p^2 < \lambda_0 \).

3.2.2. \( p = 1 - o(1) \). One can check it by following the same strategy as above and in Section 3.1.2.

3.3. \( \lambda_0 \leq p \leq p_0 \). Let \( a = 2\lambda_0/3 \), \( a = \lfloor an \rfloor \). Fix an \( a \)-set \( A \subset V \) and generate our random graph and determine \( B = B(A) \) with \( b = |B| \). Let \( \varepsilon = (\log n)^4/\sqrt{n} \) and let \( X_A \) be the indicator random variable for \( a + b \leq (\lambda_0 + \varepsilon)n \) and \( X = \sum A X_A \). Then
\[
p(a) = \frac{1}{2} + e^{-O(n)}
\]
and with \( g_\lambda(\alpha) \) as defined in (2.2),
\[
E(X) = \exp \{ (g_{\lambda_0+\varepsilon}(2\lambda_0/3) + o(1))n \log 2 \}. \tag{3.4}
\]
Now
\[
g_{\lambda+\varepsilon}(\alpha) = g_\lambda(\alpha) + (1 - \alpha) \left( H \left( \frac{\lambda + \varepsilon - \alpha}{1 - \alpha} \right) - H \left( \frac{\lambda - \alpha}{1 - \alpha} \right) \right)
= g_\lambda(\alpha) + \varepsilon \log_2 \left( \frac{1 - \lambda}{\lambda - \alpha} \right) + O(\varepsilon^2).
\]
Plugging this into (3.4) with \( \lambda = \lambda_0 \) and \( \alpha = 2\lambda_0/3 \) we see that
\[
E(X) = \exp \left\{ \left( \varepsilon \log_2 \left( \frac{1 - \lambda_0}{\lambda_0/3} \right) + O(\varepsilon^2) \right) n \log 2 \right\} = e^{\Omega((\log n)^4n^{1/2})}. \tag{3.5}
\]
Next, we estimate the variance of \( X \). We will argue that for \( A, C \in \binom{V}{a} \) either \( |A \triangle C| \) is small (but the number of such pairs is small) or \( |A \triangle C| \) is large (but then the covariance \( \text{Cov}(X_A, X_C) \) is very small since if we fix the adjacency of some vertex \( x \) to \( C \), then the parity of \( |N(x) \cap (A \setminus C)| \) is almost a fair coin flip). Formally,
\[
\text{Var}(X) = E(X) + \sum_{A \neq C} \text{Cov}(X_A, X_C)
\leq E(X) + \sum_{|A \triangle C| < 2\sqrt{n}} \Pr(X_A = X_C = 1)
+ \sum_{|A \triangle C| > 2\sqrt{n}, |A \cap C| > \sqrt{n}} \text{Cov}(X_A, X_C)
+ \sum_{|A \cap C| < \sqrt{n}} \Pr(X_A = X_C = 1).
\]
Since $E(X)$ goes to infinity, clearly $E(X) = o(E(X)^2)$. We show in Claims 3.1, 3.2 and 3.3 that the remaining part is also bounded by $o(E(X)^2)$. Then Chebyshev’s inequality will imply that $X > 0$ asymptotically almost surely.

**Claim 3.1.** \[ \sum_{|A\Delta C|<2\sqrt{n}} \Pr(X_A = X_C = 1) = \Theta(E(X)^2) \]

**Proof.** We estimate trivially \( \Pr(X_A = X_C = 1) \leq \Pr(X_A = 1) \). Then,

\[
\sum_{|A\Delta C|<2\sqrt{n}} \Pr(X_A = 1) = \binom{n}{a} \sum_{0 \leq i < \sqrt{n}} \binom{n-a}{i} \binom{a}{a-i} \Pr(X_A = 1)
= E(X) \sum_{0 \leq i < \sqrt{n}} \binom{n-a}{i} \left( \frac{a}{a-i} \right) \leq E(X) \sqrt{\frac{n}{2}} \binom{n}{a} \leq 2^{\Theta(\sqrt{n}\log n)}.
\]

Thus, (3.5) yields that \( \sum_{|A\Delta C|<2\sqrt{n}} \Pr(X_A = X_C = 1) = o(E(X)^2) \).

**Claim 3.2.** \[ \sum_{|A\Delta C|\geq 2\sqrt{n}, |A\cap C|\geq \sqrt{n}} \operatorname{Cov}(X_A, X_C) = o(E(X)^2) \]

**Proof.** If \( x \in V \setminus (A \cup C) \), then \( \Pr(z \in B(A)|X_C = 1) = 2^{-1+o(1/n)} \), since we can always find at least \( \sqrt{n} \) vertices in \( A \setminus C \) with no adjacency with \( x \) determined by the event \( X_C = 1 \). Similarly, if \( x \in C \setminus A \), then there are at least \( \sqrt{n} - 1 \) vertices in \( A \cap C \) such that their adjacency with \( x \) is independent of the occurrence of \( X_C = 1 \). This implies that

\[
\Pr(X_A = 1|X_C = 1) = \sum_{0 \leq i < \sqrt{n}} \binom{n-a}{i} \left( \frac{a}{a-i} \right)^{2^{-(n-a)+o(1)}} = 2^{o(1)} \Pr(X_A = 1),
\]

and consequently, \( \operatorname{Cov}(X_A, X_C) = o(\Pr(X_A = 1)^2) \). Hence,

\[
\sum_{|A\Delta C|\geq 2\sqrt{n}, |A\cap C|\geq \sqrt{n}} \operatorname{Cov}(X_A, X_C) \leq \binom{n}{a}^2 o(\Pr(X_A = 1)^2) = o(E(X)^2). \]

**Claim 3.3.** \[ \sum_{|A\cap C|<\sqrt{n}} \Pr(X_A = X_C = 1) = o(E(X)^2) \]

**Proof.** First let us estimate the number of ordered pairs \((A, C)\) for which \( |A\cap C| < \sqrt{n} \). Note,

\[
\sum_{|A\cap C|<\sqrt{n}} 1 = \binom{n}{a} \sum_{0 \leq i < \sqrt{n}} \binom{n-a}{i} \left( \frac{a}{a-i} \right) \leq \sqrt{n} \binom{n}{a} \left( \frac{a}{\sqrt{n}} \right)^{2^{-(n-a)+o(1)}} = 2^{n(H(\alpha)+H(\frac{1}{2\alpha})(1-\alpha)+o(1))}.
\]

(3.6)

Now we will bound \( \Pr(X_A = X_C = 1) \) for fixed \( a \)-sets \( A \) and \( C \). Let \( S \subseteq A \setminus C \) be a set of size \( s = |S| = \lfloor \sqrt{n} \rfloor \). Define a new indicator random variable \( Y \) which takes the value 1 if and only if \( |B(C) \setminus S| \leq (\lambda_0 + \varepsilon)n - a \). Clearly, \( X_C \leq Y \) and

\[
\Pr(Y = 1) = \Pr(Bin(n-a-s, p(a)) \leq (\lambda_0 + \varepsilon)n - a)
\leq 2^{s+o(1)} \sum_{0 \leq i \leq (\lambda_0 + \varepsilon)n - a} \binom{n-a}{i} 2^{-(n-a)} = 2^{s+o(1)} \Pr(X_A = 1).
\]

Now if we condition on the existence or otherwise of all edges \( F' \) between \( C \) and \( V \setminus S \) then if \( x \in V \setminus A \)

\[
\Pr(x \in B(A) \mid F' \text{ and } F'') = \left[ \frac{1 - (1 - 2p)^s}{2}, \frac{1 + (1 - 2p)^s}{2} \right],
\]

and consequently, \( \operatorname{Cov}(X_A, X_C) = o(E(X)^2) \). Hence,

\[
\sum_{|A\Delta C|\geq 2\sqrt{n}, |A\cap C|\geq \sqrt{n}} \operatorname{Cov}(X_A, X_C) \leq \binom{n}{a}^2 o(\Pr(X_A = 1)^2) = o(E(X)^2). \]
where $F''$ is the set of edges between $x$ and $A \setminus S$. This implies that

$$\Pr(X_A = 1|Y = 1) = \sum_{0 \leq i \leq (\lambda_0 + \epsilon)n-a} \binom{n-a}{i} 2^{-(n-a)+O(\sqrt{n})} = 2^{O(\sqrt{n})} \Pr(X_A = 1).$$

Consequently,

$$\Pr(X_A = X_C = 1) \leq \Pr(X_A = Y = 1) \leq 2^{O(\sqrt{n})} \Pr(X_A = 1)^2.$$

Hence, (3.6) implies

$$\sum_{|A \cap C| < \sqrt{n}} \Pr(X_A = X_C = 1) \leq 2^{n(\alpha(H(\alpha) + H(n^{-\alpha}(1-\alpha))+(1-\alpha)))} \Pr(X_A = 1)^2.$$

To complete the proof it is enough to note that

$$E(X)^2 = 2^{n(2H(\alpha) + o(1))} \Pr(X_A = 1)^2$$

and

$$2H(\alpha) > H(\alpha) + H\left(\frac{\alpha}{1-\alpha}\right) (1-\alpha).$$

Indeed, the last inequality follows from the strict concavity of the entropy function, since then $(1-\alpha)H\left(\frac{\alpha}{1-\alpha}\right) + \alpha H(0) \leq H(\alpha)$ with the equality for $\alpha = 0$ only. \qed

Now we show that $f(G_{n,p}) \geq (\lambda_0 - \epsilon)n$. We show that

$$\sum_{1 \leq a \leq (\lambda_0 - \epsilon)n} \binom{n}{a} \Pr(Bin(n-a, p(a)) \leq (\lambda_0 - \epsilon)n-a) = o(1).$$

As in previous sections we split this sum into two sums but this time we make the break into $1 \leq a \leq (\log n)^2$ and $\log n)^2 < a \leq (\lambda_0 - \epsilon)n$, respectively. In order to estimate the first sum we use the Chernoff bounds with deviation $1-\theta$ from the mean where

$$\theta = 1 - \frac{(\lambda_0 - \epsilon)n-a}{(n-a)p(a)} \geq 1 - \frac{\lambda_0 - \epsilon}{p(a)} \geq 1 - \frac{\lambda_0 - \epsilon}{\lambda_0} = \frac{\epsilon}{\lambda_0}.$$

Consequently,

$$\sum_{2 \leq a < (\log n)^2} \binom{n}{a} \Pr(Bin(n-a, p(a)) \leq (\lambda_0 - \epsilon)n-a) \leq (\log n)^2 \binom{n}{(\log n)^2} \exp\{-\Omega((\log n)^4)\} \leq \exp\{-\Omega((\log n)^4)\} = o(1).$$

Now we bound the second sum corresponding to $(\log n)^2 < a \leq (\lambda_0 - \epsilon)n$

$$\sum_{(\log n)^2 \leq a \leq (\lambda_0 - \epsilon)n} \binom{n}{a} \Pr(Bin(n-a, p(a)) \leq (\lambda_0 - \epsilon)n-a) = 2^{n(h(\lambda_0 - \epsilon)+O(\log n/n))} = o(1).$$
4. General Graphs. Here we present the proof of Theorem 1.2. First, we prove a weaker result $f(n) \leq (0.440 \ldots + o(1)) n$.

Suppose we aim at showing that $f(n) \leq \lambda n$. We fix some $\alpha$ and $\rho$ and let $a = \alpha n$ and $r = \rho n$. For each $a$-set $A$ let $R(A)$ consist of all sets that have Hamming distance at most $r$ from $B(A)$. If

$$\binom{n}{a} \sum_{i=0}^{r} \binom{n}{i} = 2^{n(H(\alpha)+H(\rho)+o(1))} > 2^n,$$  \hspace{1cm} (4.1)

then there are $A, A'$ such that $R(A) \cap R(A') \ni C$ is non-empty. This means that $C$ is within Hamming distance $r$ from both $B = B(A)$ and $B' = B(A')$. Thus $|B \Delta B'| \leq 2r$.

Let all vertices in $A'' = A \Delta A'$ flip their neighbors, i.e., execute operation $O_1$. The only vertices outside of $A''$ that can have an odd number of neighbors in $A''$ are restricted to $(B \Delta B') \cup (A \Delta A')$. Thus

$$f(G) \leq |A \Delta A'| + |(B \Delta B') \cup (A \cap A')| \leq 2a + 2r = 2n(\alpha + \rho).$$ \hspace{1cm} (4.2)

Consequently, we try to minimize $\alpha + \rho$ subject to $H(\alpha) + H(\rho) > 1$. Since the entropy function is strictly concave, the optimum satisfies $\alpha = \rho$, otherwise replacing each of $\alpha, \rho$ by $(\alpha + \rho)/2$ we strictly increase $H(\alpha) + H(\rho)$ without changing the sum. Hence, the optimum choice is $\alpha = \rho \approx 0.1100278644383959 \ldots$

the smaller root of $H(x) = 1/2$, proving that $f(n) \leq (0.440 \ldots + o(1)) n$.

In order to obtain a better constant we modify the approach taken in (4.1). Let us take $\delta = 0.275$, $\alpha = 0.0535$, $a = \lfloor \alpha n \rfloor$, $d = \lfloor \delta n \rfloor$. Look at the collection of sets $B(A), A \in \binom{n}{a}$. This gives $\binom{n}{a} = 2^{nH(\alpha)+o(1)}$ binary $n$-vectors.

We claim that some two of these vectors are at distance at most $d$. If not, then inequality (5.4.1) in [4] says that

$$H(\alpha) + o(1) \leq \min\{1 + g(u^2) - g(u^2 + 2\delta u + 2\delta ) : 0 \leq u \leq 1 - 2\delta\},$$

where $g(x) = H((1 - \sqrt{1 - x})/2)$. In particular, if we take $u = 1 - 2\delta = 0.45$, we get 0.30108 + $o(1) \leq 0.30103$, a contradiction.

Thus, we can find two different $a$-sets $A$ and $A'$ such that $|B(A) \Delta B(A')| \leq d$. As in (4.2), we can conclude that $f(G) \leq 2a + d \leq (0.382 + o(1)) n$.

5. Acknowledgment. The authors would like to thank Shiang Yong Looi for suggesting this problem.

REFERENCES