The vertex degree distribution of random intersection graphs

Dudley Stark
School of Mathematical Sciences
Queen Mary, University of London
London E1 4NS U.K.

December 3, 2002

ROUGH DRAFT: NOT FOR CIRCULATION

Abstract

Random intersection graphs are a model of random graphs in which each vertex is assigned a subset of a set of objects independently and two vertices are adjacent if their assigned subsets are not disjoint. The number of vertices is denoted by \( n \) and the number of objects is supposed to be \([n^\alpha]\) for some \( \alpha > 0 \). We determine the distribution of the degree of a typical vertex and show that it changes sharply between \( \alpha < 1 \), \( \alpha = 1 \), and \( \alpha > 1 \).

1 Introduction

Random intersection graphs were introduced in [4]. Given a set \( V \) of \( n \) vertices and a set \( W \) of \( m \) objects, define a bipartite graph \( G^*(n, m, p) \) with independent vertex sets \( V \) and \( W \) and edges between \( v \in V \) and \( w \in W \) existing independently and with probability \( p \). The random intersection graph \( G(n, m, p) \) derived from \( G^*(n, m, p) \) is defined on the vertex set \( V \) with vertices \( v_1, v_2 \in V \) adjacent if and only if there exists some \( w \in W \) such that
both $v_1$ and $v_2$ are adjacent to $w$ in $G^*(n, m, p)$. We may interpret the vertices of $W_v \subset W$ adjacent to $v \in V$ as a random subset of $W$, in which case two vertices $v_1, v_2 \in V$ are adjacent iff $W_{v_1} \cap W_{v_2} \neq \emptyset$.

The properties of $G(n, m, p)$ were studied in [1, 4] and contrasted with the well known random graph model $G(n, p)$, in which vertices are made adjacent to each other independently and with probability $p$. In [1, 4] the number of objects $m$ is taken to be $m = \lfloor n^\alpha \rfloor$ for a fixed $\alpha > 0$. It is found in [4] that the thresholds for the existence of small subgraphs in $G(n, m, p)$ show different behaviors from what is seen in $G(n, p)$. When $\alpha > 6$ [1] showed that the total variation distance between the distributions of $G(n, m, p)$ and $G(n, p)$ converges to 0 when $\hat{p}$ is defined appropriately.

Intersection graphs can be viewed as relationship graphs. For example, if $V$ represents mathematicians and $W$ represents mathematics papers, and an edge is put between $v \in V$ and $w \in W$ iff mathematician $v$ was an author of paper $w$, then the resulting intersection graph is the collaboration graph on $V$, where two mathematicians are connected by an edge iff they have written a paper together. The roles of $V$ and $W$ can be interchanged, in which case two papers are adjacent iff they have an author in common. The random graph $G(n, m, p)$ is dual in this way to $G(m, n, p)$. If $m = \lfloor n^\alpha \rfloor$, then $n \in [m^{1/\alpha}, (m + 1)^{1/\alpha})$ and so the dual of $G(n, m, p)$ with $\alpha = \beta > 1$ is basically $G(n, m, p)$ with $\alpha = \beta^{-1} < 1$. Data sets for relationship graphs and models of random intersection graphs with fixed degree sequences are analyzed in [5].

Interest was expressed in [1] in further understanding the differences between $G(n, m, p)$ and $G(n, p)$. In [1, 4] thresholds for various quantities were looked at, but not much attention was paid to limiting distributions. A fundamental quantity that has not been studied for random intersection graphs is the distribution of the degree of a typical vertex. We give the precise distribution for $G(n, m, p)$ in the form of a probability generating function in Theorem 1. The corresponding distribution for $G(n, p)$ is, of course, Binomial$(n - 1, p)$.

Let $X = X(n, m, p)$ be the number of vertices $V - \{v\}$ adjacent in $G(n, m, p)$ to a vertex $v \in V$. The probability generating function of $X(n, m, p)$ is defined to be $E x^X = \sum_{k=0}^{\infty} \mathbb{P}(X = k)x^k$. 

2
Theorem 1 The probability generating function \( F(x) = \mathbb{E}x^X \) is given by
\[
F(x) = \sum_{j=0}^{n-1} \binom{n-1}{j} x^j (1-x)^{n-1-j} \left[ 1 - p + p(1-p)^{n-1-j} \right]^m.
\]

Theorem 1 is proved by using a generating function version of the sieve method.

The expectation of \( X \) is given by
\[
\mathbb{E}X = (n-1) \left[ 1 - (1-p^2)^m \right] \tag{1}
\]
because the expression in square brackets is the probability that two vertices \( v, v_1 \) in \( V \) are simultaneously adjacent to some vertex \( w \in W \) in \( \mathcal{G}^*(n, m, p) \).

The derivative \( F'(1) \) also gives (1). If we let
\[
p = \sqrt{cn^{-(1+\alpha)/2}},
\]
then
\[
\mathbb{E}X = (n-1) \left[ 1 - (1 - mp^2 + O(mp^4)) \right] = c + o(1).
\]

With respect to vertex degree, defining \( p = p(n) \) by (2) for \( \mathcal{G}(n, p) \) is therefore analogous to defining \( p = cn^{-1} \) for \( \mathcal{G}(n, p) \).

The vertex degree distribution for \( \mathcal{G}(n, p) \) converges to the Poisson distribution with parameter \( c \) as \( n \to \infty \) when \( p = cn^{-1} \). With \( p = \sqrt{cn^{-(1+\alpha)/2}} \) the vertex degree distribution of \( \mathcal{G}(n, m, p) \) converges to a Poisson distribution in the limit if and only if \( \alpha > 1 \). We say that \( X_n \) is asymptotically almost surely \( a_n \) if \( \mathbb{P}(|X_n/a_n - 1| > \epsilon) \to 0 \) as \( n \to \infty \) for each \( \epsilon > 0 \).

Theorem 2 Let \( \mathcal{G}(n, m, p) \) denote the random intersection graph with \( m = \lfloor n^{\alpha} \rfloor \) and \( p = \sqrt{cn^{-(1+\alpha)/2}} \).

(i) If \( \alpha < 1 \), then the number of non-isolated vertices is asymptotically almost surely \( \sqrt{cn^{(1+\alpha)/2}} = o(n) \). It follows that the degree of a fixed vertex in \( V \) has a distribution which converges to \( \delta_0 \), the probability distribution with all mass at 0.

(ii) If \( \alpha = 1 \), then the degree of a fixed vertex in \( V \) has a distribution which converges weakly to the compound Poisson distribution of the random variable \( Z_1 + Z_2 + \cdots + Z_N \), where \( N, Z_1, Z_2, \ldots \) are i.i.d. Poisson(\( \sqrt{c} \)) random variables.
(iii) If $\mathcal{G}(n, m, p)$ with $\alpha > 1$ and $p = \sqrt{c}n^{-(1+\alpha)/2}$, then the degree of a fixed vertex has distribution which converges weakly to a Poisson limiting distribution with parameter $c$.

Theorem 2 can roughly be explained in the following way. When $\alpha < 1$ the probability in $\mathcal{G}(n, m, p)$ that any one vertex $v \in V$ is connected to a vertex in $W$ goes to 0 and so the degree distribution in $\mathcal{G}(n, m, p)$ converges to $\delta_0$ and most of the vertices are isolated. When $\alpha = 1$ a vertex $v \in V$ will have approximately a Poisson($\sqrt{c}$) number of neighbors in $W$ and each of those neighbors have independently about Poisson($\sqrt{c}$) number of neighbors in $V$, not including $v$. When $\alpha$ is large enough, [1] shows that it becomes unlikely that a vertex $w \in W$ has more than two neighbors in $V$ and as a result the events that different edges in $\mathcal{G}(n, m, p)$ exist become independent.

Theorem 3 shows that if $\alpha > 1$ and $p$ grows faster than $n^{-(1+\alpha)/2}$, but not as fast as $\min\left(n^{-2/3-\alpha/3}, n^{-1/3-\alpha/2}\right)$, then $X$ converges to normal when it is rescaled.

**Theorem 3** Let $\mathcal{G}(n, m, p)$ denote the random intersection graph with $m = \lfloor n^\alpha \rfloor$, suppose that $\alpha > 1$, and suppose that $p$ satisfies $nm\mathbb{E}p^2 \to \infty$ and $p = o(n^{-2/3-\alpha/3})$ if $1 < \alpha \leq 2$, $p = o\left(n^{-1/3-\alpha/2}\right)$ if $\alpha > 2$. Under these assumptions,

$$\frac{X - \mathbb{E}X}{\sigma(X)} \Rightarrow N(0, 1),$$

where $\sigma(X)$ is the standard deviation of $X$ and $N(0, 1)$ is the standard normal distribution.

The formula in Theorem 1 is derived in Section 2. In Section 3 Theorem 2 is proven for $\alpha < 1$ by using Chebyshev’s inequality. Section 4 proves Theorem 2 for $\alpha \geq 1$ and Theorem 3 by analyzing the probability generating function found in Section 2.

## 2 The probability generating function

We will determine $F(x)$ by using Lemma 1, which is a probability generating function version of the sieve method. Lemma 1 is used, for example, by Takács in [7], though according to [3] it may also have been known to Jordan. For completeness we give a proof of Lemma 1. It is similar to an argument in
Section 4.2 of [8]. Let $P$ be a set of properties that a random object can take on. Let $p_k$ be the probability that the object takes on exactly $k$ properties in $P$. We are interested in the probability generating function

$$F(x) = \sum_k p_k x^k.$$ 

**Lemma 1** For $S \subseteq P$, we define $N_S$ to be the event that the random object possesses the properties $S$. Define $N_r$ to be

$$N_r = \sum_{|S|=r} \mathbb{P}(N_S)$$

and define $N(x)$ to be

$$N(x) = \sum_{r \geq 0} N_r x^r.$$ 

With the definitions above, we have

$$F(x) = N(x - 1).$$

**Proof** The proof is similar to the argument in Section 4.2 of [8], but replacing certain summations with expectations. Let $I_{N_S}$ be the indicator function of the event $N_S$. Let $Y$ be the number of properties that the random object possesses. We have

$$N_r = \sum_{|S|=r} \mathbb{P}(N_S) = \sum_{|S|=r} \mathbb{E}(I_{N_S}) = \mathbb{E} \left( \sum_{|S|=r} I_{N_S} \right) = \mathbb{E} \left( \binom{Y}{r} \right).$$

Therefore,

$$N(x) = \sum_{r \geq 0} N_r x^r = \sum_{r \geq 0} \mathbb{E} \left( \binom{Y}{r} x^r \right) = \mathbb{E} \left( (1 + x)^Y \right) = F(1 + x)$$

and $F(x) = N(x - 1)$. 

In our application to random intersection graphs, there are $n - 1$ properties consisting of the non-adjacency of the fixed vertex to the other $n - 1$
vertices. We use non-adjacency rather than adjacency in the initial analysis for ease of calculation; the generating function for the number of adjacent vertices follows immediately.

**Proof of Theorem 1**

Let $G(x)$ be the generating function of $p^k$, the probability that exactly $k$ vertices in $V - \{v\}$ are not adjacent to $v \in V$. The probability that the fixed vertex $v$ is adjacent to none of the vertices represented by $S \subset V - \{v\}$ is given by

$$P(N_S) = \sum_{k=0}^{m} \binom{m}{k} p^k (1-p)^{m-k} (1-p)^{k|S|},$$

where the index $k$ counts the number of elements of $W$ adjacent to $v$ in $G^*(n, m, p)$. Therefore,

$$N_r = \binom{n-1}{r} \sum_{k=0}^{m} \binom{m}{k} [p(1-p)^r]^k (1-p)^{m-k}$$

$$= \binom{n-1}{r} [1 - p + p(1-p)^r]^m.$$

Lemma 1 implies that $G(x) = N(x - 1)$, where

$$N(x) = \sum_{r=0}^{n-1} N_r x^r = \sum_{r=0}^{n-1} \binom{n-1}{r} x^r [1 - p + p(1-p)^r]^m.$$

Hence,

$$G(x) = \sum_{r=0}^{n-1} \binom{n-1}{r} (x-1)^r [1 - p + p(1-p)^r]^m$$

$$= \sum_{j=0}^{n-1} \binom{n-1}{j} (x-1)^{n-1-j} [1 - p + p(1-p)^{n-1-j}]^m.$$

Now use the identity $F(x) = x^{n-1} G(x^{-1})$. □
3 The number of isolated vertices

In this section we prove Theorem 2 for \( \alpha < 1 \).

**Lemma 2** Consider a random intersection graph \( G(n, m, p) \) with \( \alpha < 1 \) and \( p = o(n^{-\alpha}) \). Let \( Y \) be the number of non-isolated vertices. If \(bmp \to \infty \), then asymptotically almost surely \( Y \sim bmp \). In particular, if \( \alpha < 1 \) and \( p = \sqrt{c}n^{-(1+\alpha)/2} \), then \( Y \sim \sqrt{c}n^{(1+\alpha)/2} = o(n) \).

**Proof** Write \( W = \sum_{v \in V} I_v \), where \( I_v \) is the indicator that vertex \( v \in V \) is isolated, so that the number of non-isolated vertices is \( Y = n - W \). The probability that \( v \) is isolated is

\[
\mathbb{E}I_v = \sum_{k=0}^{m} \binom{m}{k} p^k (1-p)^{m-k} (1-p)^{(n-1)k} = \left[1 - p + p(1-p)^{n-1}\right]^m,
\]

where the index \( k \) represents the number of vertices in \( W \) which are adjacent to \( v \) in \( G^*(n, m, p) \). Hence

\[
\mathbb{E}W = n \left[1 - p + p(1-p)^{n-1}\right]^m,
\]

a formula computed in [6] by different means. When \( \alpha < 1 \) and \( p = o(n^{-\alpha}) \) the expectation of \( Y \) is

\[
\mathbb{E}Y = n - \mathbb{E}W = n - n \left[1 - p + p(1-p)^{n-1}\right]^m = n - n (1-p)^m \left(1 + O(mpe^{-(n-2)p})\right) = n - n (1-p)^m \left(1 + O(mp \exp(-n^{1-\alpha}))\right) = nmp + O(nm^2p^2).
\]

Next we calculate the variance of \( Y \). We have

\[
\mathbb{E}W(W - 1) = \sum_{v_2 \neq v_2} \mathbb{E}I_{v_1} I_{v_2} = n(n-1) \sum_{s=0}^{m} \binom{m}{s} (1-p)^{2s}[2p(1-p)]^{m-s}(1-p)^{(m-s)(n-2)} = n(n-1) \left[(1-p)^2 + 2p(1-p)^{n-1}\right]^m,
\]

7
where $s$ counts the number of vertices in $W$ adjacent to neither $v_1$ or $v_2$ in $G^*(n, m, p)$, leaving $m-s$ vertices in $W$ adjacent to exactly one of $v_1, v_2$. The factor $(1-p)^{(m-s)(n-2)}$ is the probability that none of the $m-s$ vertices are adjacent to other vertices in $V$. Now,

$$
\text{Var}(Y) = \text{Var}(W)
= n(n-1) \left[ (1-p)^2 + 2p(1-p)^{n-1} \right]^m + n \left[ 1 - p + p(1-p)^{n-1} \right]^m
= n(n-1)(1-p)^{2m} \left( 1 + O(\exp(-n^{1-a})) \right)
+ n(1-p)^m \left( 1 + O(\exp(-n^{1-a})) \right)
- n^2(1-p)^{2m} \left( 1 + O(\exp(-n^{1-a})) \right)
= n(1-p)^m - n(1-2p + p^2)^m + O\left(n^2 \exp(-n^{1-a})\right)
= O(nmp).
$$

An application of Chebyshev’s inequality completes the proof. ■

4 Limit laws for the vertex degree

In this section we prove Theorem 2 for $\alpha \geq 1$ and Theorem 3.

Lemma 3 Consider the random intersection graph $G(n, m, p)$ with $m = \lfloor n^\alpha \rfloor$ and $p = \sqrt{cn^{-1+\alpha}/2}$ with $\alpha = 1$. When $x \leq 1$, the probability generating function $\sum_k \mathbb{P}(X(n, m, p) = k)x^k$ satisfies

$$F(x) = \exp\left(-\sqrt{c} + \sqrt{ce^{-\sqrt{c}(1-x)}}\right) + O(n^{-1/4})$$

It follows that the probability that a fixed vertex of $V$ in $G(n, m, p)$ equals $k \geq 0$ asymptotically approaches the compound Poisson distribution given by $\mathbb{P}(Z_1 + Z_2 + \cdots + Z_N = k)$, where $N, Z_1, Z_2, \ldots$ are i.i.d. Poisson($\sqrt{c}$) distributed random variables.

Proof Write the formula for $F(x)$ given by Theorem 1 for fixed $x \leq 1$ as

$$F(x) = \sum_{|j-nx| \leq n^{3/4}} \left( \begin{array}{c} n-1 \\ j \end{array} \right) x^j (1-x)^{n-1-j} \left[ 1 - \sqrt{cn^{-1}} + \sqrt{cn^{-1}(1-\sqrt{cn^{-1}})^{n-1-j}} \right]^n$$

8
\[ + \sum_{|j-nx|>n^{3/4}} \binom{n-1}{j} x^j(1-x)^{n-1-j} \left[ 1-cn^{-1}+\sqrt{cn^{-1}}(1-\sqrt{cn^{-1}})^{n-1-j} \right]^n. \]

The second sum is bounded by \( \sum_{|j-nx|>n^{3/4}} \binom{n-1}{j} x^j(1-x)^{n-1-j} \), which is \( o(1) \) by large deviation bounds for the binomial; see Theorem 2.1 of [2], for example.

As for the first sum, we have
\[ (1-\sqrt{cn^{-1}})^{n-1-j} = e^{-\sqrt{\sqrt{1-x}}} + O(n^{-1/4}). \]

uniformly for all \( j \) such that \( |j-nx| \leq n^{3/4} \). Hence,
\[ \left[ 1-\sqrt{cn^{-1}}+\sqrt{cn^{-1}}(1-\sqrt{cn^{-1}})^{n-1-j} \right]^n = \exp \left( -\sqrt{c} + \sqrt{c} \exp(-\sqrt{1-x}) \right) + O(n^{-1/4}) \]

uniformly for all \( j \) such that \( |j-nx| \leq n^{3/4} \) and
\[ \sum_{|j-nx|\leq n^{3/4}} \binom{n-1}{j} x^j(1-x)^{n-1-j} \left[ 1-\sqrt{cn^{-1}}+\sqrt{cn^{-1}}(1-\sqrt{cn^{-1}})^{n-1-j} \right]^n \]
\[ = \exp \left( -\sqrt{c} + \sqrt{c} \exp(-\sqrt{1-x}) \right) + O(n^{-1/4}). \]

The Laplace transform \( F(e^{-t}) \) converges pointwise to \( \exp \left( -\sqrt{c} + \sqrt{c} \exp(-\sqrt{1-x}) \right) \)
which, as is easily checked, is the Laplace transform of \( Z_1 + Z_2 + \cdots + Z_N \).

**Lemma 4** Consider the random intersection graph \( G(n,m,p) \) with \( m = \lfloor n^\alpha \rfloor \) and \( p = \sqrt{cn^{-1(1+\alpha)/2}} \) with \( \alpha > 1 \). When \( x \leq 1 \), the probability generating function \( \sum_k \mathbb{P}(X(n,m,p) = k)x^k \) satisfies
\[ F(x) = e^{-c+c^x} + O(n^{-1}) + O(n^{1-\alpha)/2}). \]

It follows that the probability \( p_k \) that a fixed vertex of \( V \) in \( G(n,m,p) \) equals \( k \geq 0 \) is asymptotically Poisson: \( p_k \sim e^{-c^k} / k! \).

**Proof** Expand \( (1-p)^{n-1-j} \) for \( j \in [1,n] \) as
\[ (1-p)^{n-1-j} = 1 - \sqrt{c}(n-1-j)n^{(1+\alpha)/2} + O(n^{1-\alpha}). \]

9
It follows that
\[
[1 - p + p(1 - p)^{n-1-j}]^m = [1 - c(n - 1 - j)n^{-1-\alpha} + O\left(n^{(1-3\alpha)/2}\right)]^m
= \exp\left(-c(n - 1 - j)n^{-1} + O\left(n^{(1-\alpha)/2}\right)\right)
\]
uniformly for \( j \in [1, n] \). Now, for \( x \in [0, 1], \)
\[
F(x) = \sum_{j=0}^{n-1} \binom{n-1}{j} x^j (1-x)^{n-1-j} e^{-c(n-1-j)/n} + O\left(n^{(1-\alpha)/2}\right)
= e^{-c(n-1)/n} (1 - x + xe^{c/n})^{n-1} + O\left(n^{(1-\alpha)/2}\right)
= e^{-c(n-1)/n} \exp\left(xc(n-1)n^{-1} + O\left(n^{-1}\right)\right) + O\left(n^{(1-\alpha)/2}\right)
= e^{-c+cx} + O\left(n^{-1}\right) + O\left(n^{(1-\alpha)/2}\right).
\]
The Laplace transform \( F(e^t) \) converges pointwise to \( \exp(-c + ce^{-t}) \), the Laplace transform of the Poisson distribution with parameter \( c \). ■

**Proof of Theorem 3**

Suppose that \( x_n \) is a sequence of complex numbers such that \( |x_n| = O(1) \). By the assumptions on \( p \) we have \( n^2 mp^3 = o(1) \), \( nm^2 p^4 = o(1) \), \( n^2 m^2 p^5 = o(1) \), and \( n^2 m^3 p^6 = o(1) \). Therefore,
\[
(1 - p)^{n-1-j} = 1 - p(n - 1 - j) + O(n^2 p^2)
\]
and
\[
\left[1 - p + p(1 - p)^{n-1-j}\right]^m = \left[1 - p^2(n - 1 - j) + O(n^2 p^3)\right]^m
= \exp\left(-mp^2(n - 1 - j) + O(n^2 mp^3)\right)
= (1 + o(1))\exp\left(-mp^2(n - 1 - j)\right).
\]
Furthermore,
\[
F(x_n) = (1 + o(1))\sum_{j=0}^{n-1} \binom{n-1}{j} x_n^j (1-x_n)^{n-1-j} e^{-mp^2(n-1-j)}
= (1 + o(1))e^{-mp^2(n-1)} \sum_{j=0}^{n-1} \binom{n-1}{j} (x_ne^{mp^2})^j (1-x_n)^{n-1-j}
\]

10
\[
\begin{align*}
&= (1 + o(1))e^{-mp^2(n-1)}(1 + x_n(1 + x_n e^{mp^2})^{n-1} \\
&= (1 + o(1))e^{-mp^2(n-1)}(1 + x_n mp^2 + O(m^2 p^4))^{n-1} \\
&= (1 + o(1))e^{-\mu + \mu x_n},
\end{align*}
\]

with \( \mu = mp^2(n-1) \). The equality (1) shows that \( \mathbb{E}X = \mu + o(1) \). Suppose that \( \mu / \sigma(X)^2 \to 1 \) as \( n \to \infty \). Writing \( \sigma(X) \), the characteristic function of \( (X - \mathbb{E}X)/\sigma \) is

\[
\begin{align*}
e^{-it\mathbb{E}X/\sigma} F(e^{it/\sigma}) &= (1 + o(1))e^{-it\mu/\sigma} F(e^{it/\sigma}) \\
&= (1 + o(1))e^{-it\mu/\sigma} \exp(-\mu + \mu e^{it/\sigma}) \\
&= (1 + o(1))e^{-it\mu/\sigma} \exp(\mu it/\sigma - \mu t^2/(2\sigma^2) + O(\mu/\sigma^3)) \\
&= (1 + o(1))e^{-t^2/2},
\end{align*}
\]

which converges to the characteristic function of the standard normal distribution.

It remains to be shown that \( \mu / \sigma^2 \to 1 \). By Theorem 1, the second derivative of \( F(x) \) at 1 equals

\[
F''(1) = \mathbb{E}X(X - 1) = (n - 1)(n - 2)[1 - 2(1 - p^2)^m + (1 - 2p^2 + p^3)^m],
\]

from which

\[
\begin{align*}
\sigma^2 &= (n - 1)(n - 2)[1 - 2(1 - p^2)^m + (1 - 2p^2 + p^3)^m] \\
&\quad + (n - 1)[1 - (1 - p^2)^m] - (n - 1)^2[1 - (1 - p^2)^m]^2 \\
&= (n - 1)(1 - p^2)^m + (n - 1)(n - 2)(1 - 2p^2 + p^3)^m \\
&\quad - (n - 1)^2(1 - 2p^2 + p^4)^m \\
&= (n - 1)[1 - mp^2 + O(m^2 p^4)] \\
&\quad + (n - 1)(n - 2)[1 - 2mp^2 + 2m^2 p^4 + O(mp^3) + O(m^2 p^5) + O(m^3 p^6)] \\
&\quad - (n - 1)^2[1 - 2mp^2 + 2m^2 p^4 + O(mp^4) + O(m^3 p^6)] \\
&= \mu + o(1).
\end{align*}
\]

References


