The strong matching number of a random graph

Lane Clark
Department of Mathematics
Southern Illinois University at Carbondale
Carbondale, IL 62901–4408, U.S.A.
lclark@math.siu.edu

Abstract
The strong matching number sm(G) of a graph G is the maximum number of edges in G that induces a matching in the graph. For fixed 0 < p < 1, El Maftouhi and Marquez Gordones [Australasian Journal of Combinatorics 10 (1994), 97–104] showed that sm(G_{n,p}) is one of only a finite number of values for a.e. G_{n,p} ∈ G(n,p). We show that, in fact, sm(G_{n,p}) is one of only two possible values for a.e. G_{n,p} ∈ G(n,p); determine the probability of attaining each value; and find the limiting distribution of the number of maximum strong matchings in G_{n,p} ∈ G(n,p).

1. Introduction

The vertex set (edge set) of a finite simple undirected graph G is denoted by V(G) (E(G)). The order (size) of G is |V(G)| (|E(G)|). For \( \phi \neq S \subseteq V(G) \), the subgraph G[S] of G induced by S has vertex set S and edge set those edges of G both ends of which are in S. A set M ⊆ E(G) is a matching of G provided no two edges in M have a common end-vertex. A matching M of G is a strong matching if and only if M = E(G[S]) where S = S(M) is the set of all end-vertices of edges in M (i.e., G[S] is a 1-regular induced subgraph of G). Equivalently, a strong matching of G is a set \{e_1, \ldots, e_m\} of pair-wise vertex-disjoint edges of G such that no edge of G connects an end-vertex of e_i with an end-vertex of e_j for 1 ≤ i ≠ j ≤ m. Observe that G has a strong matching of size k whenever it has a strong matching of size \ell for 1 ≤ k ≤ \ell and that any edge of G is itself a strong matching. The strong matching number sm(G) of G is the maximum number of edges in a strong matching of G (here sm(G) = 0 for the empty graph). Though not expressed in terms of the above parameter, [4, 5, 7] contain related results. The concept and the notation sm(G), though not the terminology, appear in [6].

The probability space G(n,p) consists of all graphs with vertex set [n] := \{1, \ldots, n\} in which edges are chosen independently with probability p = p(n). For a random graph G_{n,p} ∈ G(n,p), Pr(G_{n,p}) = p^m q^{N-m} when G_{n,p} has size m where q = 1 - p and N = n \choose 2. A class of graphs which is closed under isomorphism
is called a property of graphs. We say almost every (a.e.) \( G_{n,p} \in \mathcal{G}(n, p) \) has a property \( Q \) provided \( \Pr \left( G_{n,p} \in \mathcal{G}(n, p) \text{ has } Q \right) \to 1 \) as \( n \to \infty \). As usual, \( E(Y) \) and \( \text{Var}(Y) \) denote the expectation and variance of \( Y \). A random variable having Poisson distribution with mean \( \lambda > 0 \) is denoted by \( \text{Po}(\lambda) \) and one having normal distribution with mean 0 and variance 1 by \( N(0, 1) \). We write \( Y_n \xrightarrow{d} Y \) when the sequence \( Y_n \) converges in distribution to \( Y \).

Recently, El Maftouhi and Marquez Gordones [4] showed that for fixed \( 0 < p < 1 \), \( \text{sm}(G_{n,p}) \) is concentrated for a.e. \( G_{n,p} \in \mathcal{G}(n, p) \). Throughout, \( d = 1/(1-p) = 1/q \).

**Theorem** (El Maftouhi and Marquez Gordones [4]). For fixed \( 0 < p < 1 \), there exist positive constants \( c_1 \) and \( c_2 \) depending only on \( p \) such that:

1. a.e. \( G_{n,p} \in \mathcal{G}(n, p) \) contains a strong matching of size \( m \) for each \( m \) satisfying
   \[ m \leq \log_d n - \frac{1}{2} \log_d \log_d n - c_1. \]
2. a.e. \( G_{n,p} \in \mathcal{G}(n, p) \) does not contain a strong matching of size \( m \) for each \( m \) satisfying
   \[ m \geq \log_d n - \frac{1}{2} \log_d \log_d n + c_2. \]

We show that, in fact, \( \text{sm}(G_{n,p}) \) is one of only two possible values for a.e. \( G_{n,p} \in \mathcal{G}(n, p) \); determine the probability of attaining each value; and find the limiting distribution of the number of maximum strong matchings in \( G_{n,p} \in \mathcal{G}(n, p) \). More precisely we prove the following results.

**Theorem.** Fix \( 0 < p < 1 < 2c < 2 \) and let

\[ m = \left\lfloor \log_d n - \frac{1}{2} \log_d \log_d n + c + \frac{1}{2} \log_d \left( \frac{p}{2} \right) \right\rfloor. \]

For all constant \( 0 < \delta < 2 - 2c \),

\[ \Pr \left( m - 1 \leq \text{sm}(G_{n,p}) \leq m \right) = 1 - o(n^{-\delta}). \]

In fact, for all constant \( 0 < \delta' < 2c - 1 \),

\[ \Pr \left( \text{sm}(G_{n,p}) = m - 1 \right) = e^{-\lambda_m} - o(n^{-\delta}) + o(n^{-\delta'}) \]
\[ \Pr \left( \text{sm}(G_{n,p}) = m \right) = 1 - e^{-\lambda_m} - o(n^{-\delta'}). \]

Here \( \lambda_m \) is the expected number of strong matchings of size \( m \) in \( G_{n,p} \in \mathcal{G}(n, p) \) and is given in (1) in the next section. In addition, if \( \lim_{n \to \infty} \lambda_m = \lambda \in (0, \infty) \), then

\[ Y_m \xrightarrow{d} \text{Po}(\lambda) \]

while, if \( \lim_{n \to \infty} \lambda_m = \infty \), then

\[ \frac{Y_m - \lambda_m}{\sqrt{\lambda_m}} \xrightarrow{d} N(0, 1). \]

Here \( Y_m(G_{n,p}) \) is the number of strong matchings of size \( m \) in \( G_{n,p} \in \mathcal{G}(n, p) \) and is defined in the next section.

We write \( a \preceq b \) to indicate that the inequality \( a \leq b \) holds for all sufficiently large integers \( n \). All other inequalities hold absolutely for the range of parameters being considered. We denote the nonnegative integers by \( \mathbb{N} \), the positive integers
by \( \mathbb{Z}^+ \) and the real numbers by \( \mathbb{R} \). Recall that \( f(n) = o(g(n)) \) means that \( \lim_{n \to \infty} f(n)/g(n) = 0 \), \( f(n) \gg g(n) \) that \( g(n) = o(f(n)) \) and \( f(n) \sim g(n) \) that \( \lim_{n \to \infty} f(n)/g(n) = 1 \). For \( x \in \mathbb{R} \), \( (x)_0 = 1 \) and \( (x)_k = (x) \cdots (x - k + 1) \) for \( k \in \mathbb{Z}^+ \). Our notation and terminology generally follows Bollobás [3].

2. Results

For \( n \geq 4m + 1 \geq 5 \), let \( M_1, \ldots, M_t \) be the distinct \( m \)-matchings (i.e., having precisely \( m \) edges) in \([n]\) and \( S_i = S(M_i) \) be the set of all \( 2m \) end-vertices of edges in \( M_i \) (\( 1 \leq i \leq t \)). Here \( t = (n)_{2m}/m!2^m \sim n^{2m}/m!2^m \) as \( n \to \infty \) when \( m = o(n^{1/2}) \). For \( G_{n,p} \in \mathcal{G}(n,p) \), let

\[
X_i(G_{n,p}) := \begin{cases} 
1 & \text{, } M_i \text{ is a strong matching in } G_{n,p} \\
0 & \text{, otherwise,}
\end{cases}
\]

hence,

\[
E(X_i) = p^m q^M ; M := \binom{2m}{2} - m = 2m^2 - 2m,
\]

since the edge set of \( G_{n,p}[S_i] \) is precisely \( M_i \). Let

\[
Y_m = Y_{m,n} := \sum_{i=1}^t X_i,
\]

hence, for \( m = o(n^{1/2}) \),

\[
\lambda_m = \lambda_{m,n} := E(Y_m) = \frac{p^m q^M(n)_{2m}}{m!2^m} \sim \frac{p^m q^M n^{2m}}{m!2^m} := \bar{\lambda}_{m,n} = \bar{\lambda}_m \tag{1}
\]

(in fact, \( \lambda_m \leq \bar{\lambda}_m \)). If \( E(X_iX_j) \neq 0 \), then \( ab \in M_i \) if and only if \( ab \in M_j \) whenever \( a, b \in S_i \cap S_j \). Hence, \( M_j \) must consist of \( k \) edges of \( M_i \); \( \ell \) other edges each adjacent to precisely one edge of \( M_i \); and \( m - k - \ell \) other edges each adjacent to no edge of \( M_i \). Necessarily, \( 0 \leq k + \ell \leq m \) and \( 0 \leq k \leq m - 1 \) for \( i \neq j \). Hence, for each \( 1 \leq i \leq t \),

\[
\sum_{|S_i \cap S_j| \geq 2} E(X_iX_j) = \sum_{|S_i \cap S_j| \geq 2} \left\{ \binom{m}{k} \binom{m - k}{\ell} 2^\ell \left( \frac{n - 2m}{2m - 2k - \ell} \right) (2m - 2k - \ell)^\ell \right. \\
\left. \times \frac{\binom{2m - 2k - 2\ell}{2, \ldots, 2}}{(m - k - \ell)!} \frac{p^{2m-k}q^{2M-(\binom{2k+\ell}{2})+k}}{n^2} \right\} \\
\leq \sum_{k, \ell, m - k - \ell} \binom{m}{k, \ell, m - k - \ell} \frac{p^{2m-k}q^{2M-(\binom{2k+\ell}{2})+k}}{n^2} \frac{2m - 2k - 2\ell}{(m - k - \ell)!} \tag{2}
\]

49
Here, we first choose the $k$ common edges of $M_i$ and $M_j$; then choose the ends of $\ell$ other edges of $M_i$; next choose the remaining $2m - 2k - \ell$ vertices of $S_j - S_i$; then match these $\ell$ (ordered) vertices in $S_i$ with $\ell$ vertices of $S_j - S_i$; and finally match the remaining $2m - 2k - 2\ell$ vertices of $S_j - S_i$. Note that each $G_{n,p}[S_i] \cup G_{n,p}[S_j]$ has the same set of $2m - k$ edges and the same set of $2M - \binom{2k + \ell}{2} + k$ non-edges.

We note that (by independence),

$$\sum_{i=1}^{t} \sum_{\substack{|S_i \cap S_j| \leq 1 \atop i \neq j}} E(X_iX_j) \leq E(Y_m)^2,$$

so that,

$$\text{Var}(Y_m) \leq \lambda_m + \sum_{i=1}^{t} \sum_{\substack{|S_i \cap S_j| \geq 2 \atop i \neq j}} E(X_iX_j). \quad (3)$$

In what follows $0 < p < 1$ is constant, $q = 1 - p$, $d = 1/q$ and $m = \log_d n - \frac{1}{2} \log_d \log_d n + c(n) \in \mathbb{Z}^+$ where $c(n)$ is a bounded function. We will estimate (2) generally for these parameters and apply these estimates to specific such $m$ in Theorems 1, 2 and Corollary 3.

Let $T := \log_d(3d^2) > 2$. Now,

$$S_1 := \sum_{\substack{2 \leq 2k + \ell \leq 2T \atop 0 \leq k + \ell \leq m}} \binom{m}{k, \ell, m - k - \ell} \frac{p^{2m-k}q^{2M-(\binom{2k+\ell}{2})+k}n^{2m-2k-\ell}}{2^{m-k-2\ell}(m-k-\ell)!}$$

$$\leq \frac{p^{2m}q^{2M}n^{2m}}{m! \cdot 2^m} \sum_{\eta} \frac{2^{k+2\ell}d^{(\binom{2k+\ell}{2})}m^{2k+2\ell}}{p^k n^{2k+\ell}}$$

$$\leq \hat{c} \frac{p^{2m}q^{2M}m^{4T}n^{2m-2}}{m! \cdot 2^m} = \frac{\hat{c}\lambda m^p q^m q^M m^{4T}}{n^2}, \quad (4)$$

where $\hat{c} = 4T^2(16p^{-1})Td^{2T^2}$. We next need to carefully estimate the terms in (2).

For $k, \ell \in \mathbb{N}$ with $0 \leq k + \ell \leq m$, let

$$f(k, \ell) := \frac{2^{k+2\ell}d^{(\binom{2k+\ell}{2})}q^k}{p^k n^{2k+\ell}(m-k-\ell)!} \leq \frac{n^{(2k+\ell)}\left\{ \frac{2k+\ell}{2} \log_d n - 1 + \log_d(4p^{-1}) \right\}}{n^{(m-k-\ell)}(m-k-\ell)!}.$$

If $m/2 \leq 2k + \ell \leq 2 \log_d n - 4 \log_d m$,

$$1 - \frac{2k + \ell}{2 \log_d n} - \frac{\log_d(4p^{-1})}{\log_d n} - \frac{T}{2k + \ell} \geq \frac{7 \log_d m}{4 \log_d n} > 0$$

so that,

$$f(k, \ell) \leq \frac{1}{m! n^T}. \quad (5)$$
Next, if $2T \leq 2k + \ell \leq m/2$,

$$1 - \frac{2k + \ell}{2 \log_d n} - \frac{\log_d(4p^{-1})}{\log_d n} - \frac{T}{2k + \ell} \geq \frac{1}{5}$$

and, again,

$$f(k, \ell) \leq \frac{1}{m! n^T}.$$  \hspace{1cm} (6)

Hence, (5), (6) and the Multinomial Theorem imply

$$S_2 := \sum_{2T \leq 2k + \ell \leq 2 \log_d n - 4 \log_d m} \left( \sum_{0 \leq k + \ell \leq m} \sum_{0 \leq k' \leq \ell} \binom{m}{k, \ell, m - k - \ell} p^{2m-k} q^{2M-(2k+\ell)} n^{2m-2k-\ell} \right) \leq \frac{p^{2m} q^{2M} n^{2m}}{m! 2^m n^T} \left( \sum_{m} \binom{m}{k, \ell, m - k - \ell} \right)$$

$$\leq \frac{3^m p^{2m} q^{2M} n^{2m}}{m! 2^m n^T} \leq \frac{\lambda_m p^m q^M}{n^2}.$$ \hspace{1cm} (7)

For $k, \ell \in \mathbb{N}$ with $0 \leq k + \ell \leq m - 1$,

$$f(k, \ell + 1) = \frac{4(m - k - \ell)d^{2k+\ell}}{n} f(k, \ell).$$

If, in addition, $(5 \log_d n)/4 \leq 2k + \ell$,

$$\frac{4(m - k - \ell)d^{2k+\ell}}{n} \geq 4n^{1/4} \geq 1$$

so that $(2k + \ell \leq 2k + m - k$ here),

$$f(k, \ell) \leq f(k, m - k)$$

$$\leq n^{(m+k)} \left\{ \frac{m+k}{2 \log_d n} - 1 \right\} + \frac{(m+k) \log_d(4p^{-1})}{\log_d n}.$$ 

If $2 \log_d n - 4 \log_d m \leq 2k + \ell$,

$$2k + \ell \geq \frac{5}{4} \log_d n; \quad k \geq \frac{\log_d n}{4}; \quad m + k > \log_d n$$

and (by considering the derivative with respect to real $k$),

$$(m + k) - \frac{(m + k)^2}{2 \log_d n}$$

decreases as $k$ increases for all sufficiently large $n$. 

51
If, further, \( k \leq m - \log_d m \),

\[
(m + k) - \frac{(m + k)^2}{2 \log_d n} \geq (2m - \log_d m) - \frac{(2m - \log_d m)^2}{2 \log_d n} \geq \frac{5}{4} \log_d \log_d n
\]

and \((m + k)/\log_d n\) is bounded

\[
(m + k) - \frac{(m + k)^2}{2 \log_d n} - \frac{(m + k) \log_d (4p^{-1})}{\log_d n} - \frac{m \log_d m}{\log_d n} - T \geq \frac{\log_d \log_d n}{5}
\]

so that,

\[
f(k, \ell) \leq \frac{1}{m! n^T}.
\]  

(8)

Hence, (8) and the Multinomial Theorem imply

\[
S_3 := \sum_{\substack{2 \log_d n - 4 \log_d m \leq 2k + \ell \leq 2m \\
0 \leq k \leq m - \log_d m \\
0 \leq k + \ell \leq m}}^* \binom{m}{k, \ell, m - k - \ell} \frac{p^{2m - k} q M^{-(2k + \ell) + k} n^{2m - 2k - \ell}}{2^{m - k - 2\ell} (m - k - \ell)!} \leq \frac{\lambda_m q^m q M}{n^2}.
\]  

(9)

For \( k, \ell \in \mathbb{N} \) with \( 0 \leq k + \ell \leq m \), let

\[
g(k, \ell) := \frac{p^{m - k} q M^{-(2k + \ell) + k} n^{2m - 2k - \ell}}{2^{m - k - 2\ell} (m - k - \ell)!}
\]

hence, for \( 0 \leq k + \ell \leq m - 1 \),

\[
g(k + 1, \ell) = \frac{2(m - k - \ell) d^{4k + 2\ell}}{pn^2} g(k, \ell).
\]

If, in addition, \( 2 \log_d n - 4 \log_d m \leq 2k + \ell \),

\[
\frac{2(m - k - \ell) d^{4k + 2\ell}}{pn^2} \geq n^{3/2} \geq 1
\]

so that \( (2k + \ell \leq 2(m - \ell) + \ell \) here),

\[
g(k, \ell) \leq g(m - \ell, \ell) = (2pn)^\ell q^{2m \ell - \binom{\ell + 1}{2} - \ell} \quad = \left( \frac{2p \log_d n}{n} \right)^\ell \frac{(\ell + 1) + \ell - 2c(n) \ell}{\log_d n}.
\]
If, further, $m - \log_d m \leq k$, then $\ell \leq \log_d m$ and

$$g(k, \ell) \leq \left( \frac{2p \log_d n}{n} \right)^\ell n \frac{(2|c(n)|+2)\log^2_m n}{\log_d n}$$

hence, for all $\ell \geq 1$,

$$g(k, \ell) \leq 2n \left\{ \frac{\tilde{c}(\log_d \log_d n)^2}{\log_d n} - 1 \right\}$$

(10)

where $\tilde{c} \geq 2|c(n)| + 3$. Also,

$$2 \log_d n - 4 \log_d m \leq 2k$$

so that (recall $k \leq m - 1$),

$$g(k, 0) \leq g(m - 1, 0) = \frac{pq^{4m-4}n^2}{2} \leq \frac{d^{2\tilde{c}} \log^2_d n}{n^2}.$$  

(11)

Hence, (10) and (11) imply

$$S_4 := \sum_{\substack{2 \log_d n - 4 \log_d m \leq 2k + \ell \leq 2m \\ m - \log_d m \leq k \leq m - 1 \\ 0 \leq k + \ell \leq m \\ 0 \leq k, \ell}} p^{2m-k}q^{2M-(\ell+1)/2}n^{2m-2k-\ell} \frac{2m-k-\ell}{(m-k-\ell)!} \binom{m}{k, \ell, m-k-\ell}$$

\[ \leq 2p^m q^M n \left\{ \frac{\tilde{c}(\log_d \log_d n)^2}{\log_d n} - 1 \right\} \sum_n \binom{m}{k, \ell, m-k-\ell} \]

\[ \leq 2p^m q^M m^{2+\log_d m} n \left\{ \frac{\tilde{c}(\log_d \log_d n)^2}{\log_d n} - 1 \right\} \]

\[ \leq 2p^m q^M n \left\{ \frac{2\tilde{c}(\log_d \log_d n)^2}{\log_d n} - 1 \right\}. \]

(12)

Consequently, for each $1 \leq i \leq t$, (2), (4), (7), (9) and (12) imply

$$\sum_{|S_i \cap S_j| \geq 2 \atop i \neq j} E(X_i X_j) \leq \left\{ 2 + (\tilde{c} + 2)\tilde{\lambda}_m \right\} p^m q^M n \left\{ \frac{2\tilde{c}(\log_d \log_d n)^2}{\log_d n} - 1 \right\}$$

hence,

$$\sum_{i=1}^t \sum_{|S_i \cap S_j| \geq 2 \atop i \neq j} E(X_i X_j) \leq \left\{ 2 + (\tilde{c} + 2)\tilde{\lambda}_m \right\} \tilde{\lambda}_m n \left\{ \frac{2\tilde{c}(\log_d \log_d n)^2}{\log_d n} - 1 \right\}$$

(13)
and, (3) and (13) imply
\[ \text{Var}(Y_m) \leq 2 + (c + 2) \tilde{\lambda}_m \tilde{\lambda}_m n \left\{ \frac{2c (\log_d \log_d n)^2}{\log_d n} - 1 \right\} \] (14)

For \( m = \log_d n - \frac{1}{2} \log_d \log_d n + c(n) \in \mathbb{Z}^+ \) where \( c(n) \) is a bounded function, standard estimates give,
\[ \log_d \left( \frac{\log_d n}{m} \right) = \frac{\log_d \log_d n}{2 \ln d \log_d n} + o \left( \frac{\log_d \log_d n}{\log_d n} \right) \]
hence, (1) and Stirling's formula imply,
\[ \tilde{\lambda}_m = n^{2 - 2c(n) + \log_d (ep/2) + \{c(n) + 1/2 \ln d - 0.5 \log_d (ep/2) \} \} \frac{\log_d \log_d n}{\log_d n} + o \left( \frac{\log_d \log_d n}{\log_d n} \right) \] (15)

We are now ready to prove that \( \text{sm}(G_{n,p}) \) is one of only two possible values for a.e. \( G_{n,p} \in \mathcal{G}(n,p) \).

**Theorem 1.** Fix \( 0 < p < 1 < 2c < 2 \) and let \( m = \left[ \log_d n - \frac{1}{2} \log_d \log_d n + c + \frac{1}{2} \log_d \left( \frac{ep}{2} \right) \right] \). For all constant \( 0 < \delta < 2 - 2c \),
\[ \Pr \left( m - 1 \leq \text{sm}(G_{n,p}) \leq m \right) = 1 - o(n^{-\delta}). \]

**Proof.** From (15) (and \( \lambda_{m+1} \sim \tilde{\lambda}_{m+1} \)), we have
\[ \lambda_{m+1} = o(n^{-2c+\epsilon}) \quad (0 < \epsilon < 1) \]
hence, Markov's inequality implies
\[ \Pr (Y_{m+1} \geq 1) = o(n^{-2c+\epsilon}) \quad (0 < \epsilon < 1). \] (16)

It is readily seen that \( m - 1 = \left\lfloor \log_d n - \frac{1}{2} \log_d \log_d n + c + \frac{1}{2} \log_d \left( \frac{ep}{2} \right) \right\rfloor \) for \( n \in \mathbb{Z}^+ \)
with density 1; otherwise, \( m - 1 = \left\lfloor \log_d n - \frac{1}{2} \log_d \log_d n + c - 1 + \frac{1}{2} \log_d \left( \frac{ep}{2} \right) \right\rfloor \).
From (15) (and \( \lambda_{m-1} \sim \tilde{\lambda}_{m-1} \)), we have in either case
\[ \lambda_{m-1} \gg n^{2 - 2c - \epsilon} \quad (\epsilon > 0) \] (17)
hence, (14) (applied to \( Y_{m-1} \) with \( \tilde{c} = 2c + |\log_d (ep/2)| + 3 \), (17) and Chebyshev's inequality imply
\[ \Pr (Y_{m-1} = 0) = o(n^{-2 + 2c + \epsilon}) \quad (0 < \epsilon < 2 - 2c). \] (18)

Hence, for all constant \( 0 < \epsilon < 2 - 2c \), (16) and (18) imply
\[ \Pr (m - 1 \leq \text{sm}(G_{n,p}) \leq m) = \Pr (Y_{m-1} \geq 1) - \Pr (Y_{m+1} \geq 1) = 1 - o(n^{-2 + 2c + \epsilon}) \]
since the event \((Y_k \geq 1)\) contains the event \((Y_\ell \geq 1)\) for all \(1 \leq k \leq \ell\). Our result follows upon letting \(\delta = 2 - 2c - \epsilon\). ■

**Remark.** It is readily seen that the theorem remains true if \(\epsilon = \epsilon(n) \to 0\) slowly enough.

We now discuss the Stein-Chen method of approximating the distribution of a random variable with a Poisson distribution (see [1–3]). For \(A \subseteq \mathbb{N}\) and \(\lambda > 0\), let \(x = x_{\lambda,A} : \mathbb{N} \to \mathbb{R}\) by \(x(0) = 0\) and

\[
x(m + 1) := \lambda^{-m-1}e^{-\lambda}m! \{\text{Po}(\lambda, A \cap C_m) - \text{Po}(\lambda, A) \text{Po}(\lambda, C_m)\}, \quad m \in \mathbb{N}
\]

where \(C_m := \{0, \ldots, m\}\) and \(\text{Po}(\lambda, B) := e^{-\lambda} \sum_{k \in B} \lambda^k / k!\) for \(B \subseteq \mathbb{N}\). Then

(1) \(\Delta x := \sup_{m \in \mathbb{N}} |x(m + 1) - x(m)| \leq 2 \min\{1, \lambda^{\lambda}\}\)

and

(2) for any probability space \((\Omega, \mathcal{F}, \text{Pr})\) and any \(\mathcal{F}\)-measurable non-negative integer valued random variable \(Y : \Omega \to \mathbb{N},\)

\[
\Pr(Y \in A) - \text{Po}(\lambda, A) = E\{\lambda x(Y + 1) - Y x(Y)\}. \tag{19}
\]

Define the total variation distance \(d_{TV}(Y, \text{Po}(\lambda))\) between \(Y\) and \(\text{Po}(\lambda)\) by

\[
d_{TV}(Y, \text{Po}(\lambda)) := \sup_{A \subseteq \mathbb{N}} |\Pr(Y \in A) - \text{Po}(\lambda, A)|.
\]

For a sequence \((\Omega_n, \mathcal{F}_n, \text{Pr}_n)\) of probability spaces and a sequence \(Y_n\) of \(\mathcal{F}_n\)-measurable non-negative integer valued random variables with expectation \(\lambda_n\), if

\[
d_{TV}(Y_n, \text{Po}(\lambda_n)) = o(1) \text{ as } n \to \infty,
\]

we say \(Y_n\) is **Poisson convergent**. Necessarily, \(Y_n \overset{d}{\to} \text{Po}(\lambda)\) when \(\lim_{n \to \infty} \lambda_n = \lambda \in (0, \infty)\) while \((Y_n - \lambda_n) / \sqrt{\lambda_n} \overset{d}{\to} \mathcal{N}(0, 1)\) when \(\lim_{n \to \infty} \lambda_n = \infty\). Here, \((\Omega_n, \mathcal{F}_n, \text{Pr}_n) = \mathcal{G}(n, p)\).

Again, \(m = \log_d n - \frac{1}{2} \log_d n + c(n) \in \mathbb{Z}^+\) where \(c(n)\) is a bounded function. For \(1 \leq i \leq t\), let

\[
V_i := \sum_{|S_i \cap S_j| \geq 2} X_j \quad \text{and} \quad W_i := \sum_{|S_i \cap S_j| \leq 1} X_j
\]

so that \(X_i\) and \(W_i\) are independent in \(\mathcal{G}(n, p)\) and \(Y_m = V_i + W_i + X_i\) for each \(1 \leq i \leq t\). For any function \(x : \mathbb{N} \to \mathbb{R},\)

\[
\lambda_m x(Y_m + 1) - Y_m x(Y_m) = p^m q^M \sum_{i=1}^{t} \{x(Y_m + 1) - x(W_i + 1)\}
\]

\[
+ \sum_{i=1}^{t} (p^m q^M - X_i) x(W_i + 1)
\]

\[
+ \sum_{i=1}^{t} X_i \{x(W_i + 1) - x(Y_m)\}. \tag{20}
\]
First,
\[ |x(Y_m + 1) - x(W_i + 1)| \leq \Delta x(X_i + V_i) \]
while crude estimates give,
\[ E(X_i + V_i)^* \leq p^m q^{M \frac{10m^4(n)_{2m}}{m!2^m n^2}} = \frac{10m^4\lambda_m}{n^2} \]
hence,
\[ p^m q^{M \sum_{i=1}^{t} E|x(Y_m + 1) - x(W_i + 1)|} \leq \frac{20m^4\lambda_m}{n^2}. \quad (21) \]

Next,
\[ |X_i \{x(W_i + 1) - x(Y_m)\}| \leq \Delta xX_i V_i \]
hence, (13) implies
\[ \sum_{i=1}^{t} E|X_i \{x(W_i + 1) - x(Y_m)\}| \leq \left\{4 + (2\tilde{c} + 4)\tilde{\lambda}_m\right\} n^{\frac{2\bar{c}(\log_d \log_d n)^2}{\log_d n} - 1}. \quad (22) \]

Consequently, (19), (20), (21), (22) and the independence of \( X_i \) and \( W_i \) imply,
\[ d_{TV}(Y_m, \text{Po}(\lambda_m)) \leq \frac{20m^4\lambda_m}{n^2} + \left\{4 + (2\tilde{c} + 4)\tilde{\lambda}_m\right\} n^{\frac{2\bar{c}(\log_d \log_d n)^2}{\log_d n} - 1}, \quad (23) \]

since our estimates are independent of the set \( A \).

We are now ready to prove that \( Y_m \) is Poisson convergent for appropriate \( m \) and, hence, determine the probability that \( \text{sm}(G_{n,p}) = m - 1 \) or \( m \) for \( G_{n,p} \in \mathcal{G}(n,p) \).

**Theorem 2.** Fix \( 0 < p < 1 < 2c < 2 \) and let \( m = \lceil \log_d n - \frac{1}{3} \log_d \log_d n + c + \frac{1}{2} \log_d \left( \frac{c p}{2} \right) \rceil \). Then, for all constant \( 0 < \delta' < 2c - 1 \),
\[ d_{TV}(Y_m, \text{Po}(\lambda_m)) = o(n^{-\delta'}). \]

Hence, for all constant \( 0 < \delta < 2 - 2c \), \( 0 < \delta' < 2c - 1 \),
\[ \Pr(\text{sm}(G_{n,p}) = m - 1) = e^{-\lambda_m} - o(n^{-\delta}) + o(n^{-\delta'}) \]
\[ \Pr(\text{sm}(G_{n,p}) = m) = 1 - e^{-\lambda_m} - o(n^{-\delta'}). \]

**Proof.** From (15) (and \( \lambda_m \sim \tilde{\lambda}_m \)), we have
\[ \lambda_m = o(n^{2-2c+\epsilon}) \quad (\epsilon > 0) \quad (24) \]

hence, for all constant \( 0 < \epsilon' < 2c - 1 \), (23) (with \( \bar{c} = 2c + |\log_d(ep/2)| + 5 \)) and (24) imply
\[ d_{TV}(Y_m, \text{Po}(\lambda_m)) = o(n^{1-2c+\epsilon'}). \quad (25) \]

56
Hence, for all constant $0 < \epsilon < 2 - 2c$, $0 < \epsilon' < 2c - 1$, (16), (18) and (25) imply

$$
\Pr(\text{sm}(G_{n,p}) = m - 1) = \Pr(Y_{m-1} \geq 1) - \Pr(Y_m \geq 1)
= e^{-\lambda_m} - o(n^{-2+2c+\epsilon}) + o(n^{1-2c+\epsilon'})
$$

$$
\Pr(\text{sm}(G_{n,p}) = m) = \Pr(Y_m \geq 1) - \Pr(Y_{m+1} \geq 1)
= 1 - e^{-\lambda_m} - o(n^{1-2c+\epsilon'})
$$

since the event $(Y_k \geq 1)$ contains the event $(Y_{k+1} \geq 1)$. Our result follows upon letting $\delta = 2 - 2c - \epsilon$ and $\delta' = 2c - 1 - \epsilon'$. ■

Finally, we find the limiting distribution of the number of maximum strong matchings in $G_{n,p} \in \mathcal{G}(n, p)$.

**Corollary 3.** Fix $0 < p < 1 < 2c < 2$ and let $m = \left\lfloor \log_d n - \frac{1}{2} \log_d \log_d n + c + \frac{1}{2} \log_d \left( \frac{ep}{2} \right) \right\rfloor$. If $\lim_{n \to \infty} \lambda_n = \lambda \in (0, \infty)$, then

$$
Y_m \xrightarrow{d} \text{Po}(\lambda),
$$

while, if $\lim_{n \to \infty} \lambda_n = \infty$, then

$$
\frac{Y_m - \lambda_m}{\sqrt{\lambda_m}} \xrightarrow{d} N(0, 1).
$$

**Remark.** For all such $c$ and $m$, there exists a set $S$ of positive integers having positive density with $\lim_{n \to \infty} \lambda_n = \infty$ when $n \in S$.

**References**


(Received 15/3/2000)