On the probability that a random digraph is acyclic

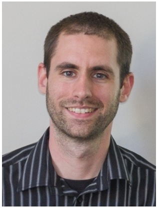
Dimbinaina Ralaivaosaona

University of Stellenbosch
naina@sun.ac.za

Joint work with
Vonjy Rasendrahasina and Stephan Wagner

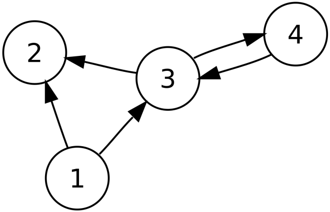
AofA2020
The 31st International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms

The team



Directed graphs (digraphs)

We consider directed graphs (digraphs) on the vertex set $\{1,\ 2,\ \cdots,n\}$ where loops and multiple edges (edges oriented in the same direction) are not allowed.



The subgraph induced by the vertices $\{3,4\}$ is called a *2-cycle*.

The following models will be mentioned :

Model $\mathcal{D}(n,p)$ **(no 2-cycles)**. Generate an undirected graph according the binomial model $\mathbb{G}(n,2p)$. Thereafter, a direction is chosen independently for each edge, with probability $\frac{1}{2}$ for each possible direction.

Model $\mathbb{D}(n,p)$ **(2-cycles can occur)**. Each of the $n(n-1)$ possible edges occurs independently with probability p.

In this work, we want to determine the probability that the random digraph $\mathcal{D}(n,p)$ is acyclic, i.e., no directed cycles. We are primarily interested in the sparse regime, where $p = \frac{\lambda}{n}$ and $\lambda = \mathcal{O}(1)$.

The following models will be mentioned :

Model $\mathcal{D}(n,p)$ (no 2-cycles). Generate an undirected graph according the binomial model $\mathbb{G}(n,2p)$. Thereafter, a direction is chosen independently for each edge, with probability $\frac{1}{2}$ for each possible direction.

Model $\mathbb{D}(n,p)$ (2-cycles can occur). Each of the $n(n-1)$ possible edges occurs independently with probability p.

In this work, we want to determine the probability that the random digraph $\mathcal{D}(n,p)$ is acyclic, i.e., no directed cycles. We are primarily interested in the sparse regime, where $p = \frac{\lambda}{n}$ and $\lambda = \mathcal{O}(1)$.

The following models will be mentioned :

Model $\mathcal{D}(n,p)$ (no 2-cycles). Generate an undirected graph according the binomial model $\mathbb{G}(n,2p)$. Thereafter, a direction is chosen independently for each edge, with probability $\frac{1}{2}$ for each possible direction.

Model $\mathbb{D}(n,p)$ (2-cycles can occur). Each of the $n(n-1)$ possible edges occurs independently with probability p.

In this work, we want to determine the probability that the random digraph $\mathcal{D}(n,p)$ is acyclic, i.e., no directed cycles. We are primarily interested in the sparse regime, where $p = \frac{\lambda}{n}$ and $\lambda = \mathcal{O}(1)$.

The model $\mathbb{D}(n,p)$ exhibits a phase transition that is somewhat similar to that of $\mathbb{G}(n,p)$ random graph model KARP (1990), ŁUCZAK (1990) :

Subcritical phase : $\lambda < 1$

- All strong components of $\mathbb{D}(n,p)$ are either cycles or single vertices.
- Every component of $\mathbb{D}(n,p)$ has at most $\omega(n)$ vertices, for any sequence $\omega(n)$ tending to infinity arbitrarily slowly.

Critical phase : $\lambda \sim 1$

- $\mathbb{D}(n,p)$ may have components of order $O(n^{1/3})$.

Supercritical phase : $\lambda > 1$

- There exists a component of linear size, while all the others contain at most $\omega(n)$ vertices.

Phase transition of random digraphs

The model $\mathbb{D}(n,p)$ exhibits a phase transition that is somewhat similar to that of $\mathbb{G}(n,p)$ random graph model Karp (1990), Łuczak (1990) :

Subcritical phase : $\lambda < 1$

- All strong components of $\mathbb{D}(n,p)$ are either cycles or single vertices.
- Every component of $\mathbb{D}(n,p)$ has at most $\omega(n)$ vertices, for any sequence $\omega(n)$ tending to infinity arbitrarily slowly.

Critical phase : $\lambda \sim 1$

- $\mathbb{D}(n,p)$ may have components of order $O(n^{1/3})$.

Supercritical phase : $\lambda > 1$

- There exists a component of linear size, while all the others contain at most $\omega(n)$ vertices.

Phase transition of random digraphs

Theorem (Karp (1990) and Łuczak (1990))

Let $p = \lambda/n$, where $\lambda \geqslant 0$ is a constant.

- When $\lambda < 1$, then w.h.p.
  - **(i)** all strong components of $\mathbb{D}(n,p)$ are either cycles or single vertices,
  - **(ii)** the number of vertices on cycles is at most ω, for any $\omega(n) \to \infty$
- when $\lambda > 1$, and let x be defined by $x < 1$ and $xe^{-x} = \lambda e^{-\lambda}$. Then w.h.p. $\mathbb{D}(n,p)$ contains a unique strong component of size $(1 - \frac{x}{\lambda})^2 n$. All other strong components are of logarithmic size

Theorem (Łuczak and Seierstad (2009))

Let $np = 1 + \varepsilon$, such that $\varepsilon = \varepsilon(n) \to 0$.

(i) If $\varepsilon^3 n \to -\infty$, then w.h.p. every component in $\mathbb{D}(n,p)$ is either a vertex or a cycle of length $O_p(1/|\varepsilon|)$.

(ii) If $\varepsilon^3 n \to \infty$, then w.h.p. $\mathbb{D}(n,p)$ contains a unique complex component, of order $(4 + o(1))\varepsilon^2 n$, whereas every other component is either a vertex or a cycle of length $O_p(1/\varepsilon)$.

Theorem (Karp (1990) and Łuczak (1990))

Let $p = \lambda/n$, where $\lambda \geqslant 0$ is a constant.

- When $\lambda < 1$, then w.h.p.
  - **(i)** all strong components of $\mathbb{D}(n,p)$ are either cycles or single vertices,
  - **(ii)** the number of vertices on cycles is at most ω, for any $\omega(n) \to \infty$
- when $\lambda > 1$, and let x be defined by $x < 1$ and $xe^{-x} = \lambda e^{-\lambda}$. Then w.h.p. $\mathbb{D}(n,p)$ contains a unique strong component of size $(1 - \frac{x}{\lambda})^2 n$. All other strong components are of logarithmic size

Theorem (Łuczak and Seierstad (2009))

Let $np = 1 + \varepsilon$, such that $\varepsilon = \varepsilon(n) \to 0$.

(i) If $\varepsilon^3 n \to -\infty$, then w.h.p. every component in $\mathbb{D}(n,p)$ is either a vertex or a cycle of length $O_p(1/|\varepsilon|)$.

(ii) If $\varepsilon^3 n \to \infty$, then w.h.p. $\mathbb{D}(n,p)$ contains a unique complex component, of order $(4 + o(1))\varepsilon^2 n$, whereas every other component is either a vertex or a cycle of length $O_p(1/\varepsilon)$.

The model $\mathcal{D}(n,p)$ of simple random digraphs was used by in Subramanian (2003), where the author studied induced acyclic subgraphs in random digraphs for fixed p.

Following this work, there are also some relatively recent results on the related question of the largest acyclic subgraph in random digraphs by Spencer and Subramanian (2008), and by Dutta and Subramanian (2011), (2014), and (2016).

The enumeration of acyclic digraphs originated in the 1970s by Liskovets (1969) Harary and Palmer (1973), Robinson (1973,1977) and Stanley (1973).

Let a_n denote the number of acyclic digraphs on n (labelled) vertices, then one has

$$a_n = \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} 2^{k(n-k)} a_{n-k} \quad \text{for} \quad n > 1$$

with initial value $a_0 = 1$.

The sequence $(a_n)_{n \geqslant 0}$ starts as follows (see OEIS A003024) :

$$1,\ 1,\ 3,\ 25,\ 543,\ 29281,\ 3781503,\ 1138779265,\ \ldots$$

Enumeration of DAGs

The enumeration of acyclic digraphs originated in the 1970s by Liskovets (1969) Harary and Palmer (1973), Robinson (1973,1977) and Stanley (1973).

Let a_n denote the number of acyclic digraphs on n (labelled) vertices, then one has

$$a_n = \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} 2^{k(n-k)} a_{n-k} \quad \text{for} \quad n > 1$$

with initial value $a_0 = 1$.

The sequence $(a_n)_{n \geqslant 0}$ starts as follows (see OEIS A003024) :

$$1,\ 1,\ 3,\ 25,\ 543,\ 29281,\ 3781503,\ 1138779265,\ \ldots$$

Generating functions

Introducing the so-called *graphic generating function*

$$A(x) = \sum_{n \geqslant 0} \frac{1}{n!} 2^{-\binom{n}{2}} \, a_n \, x^n, \quad \text{and let} \quad \phi(x) = \sum_{n \geqslant 0} \frac{(-1)^n}{n!} 2^{-\binom{n}{2}} x^n.$$

It follows from the recursive formula for $(a_n)_n$ that

$$A(x) = \frac{1}{\phi(x)}.$$

It can be shown that this function is meromorphic, and that the pole with minimum modulus occurs at $x \approx 1.48808$. From this, one can derive the asymptotic formula

$$\frac{a_n}{n!} 2^{-\binom{n}{2}} \sim \alpha \cdot \beta^n,$$

where $\alpha \approx 1.74106$ and $\beta \approx 0.672008$. This result appears in Liskovets (1973), Robinson (1973) and Stanley (1973).

It is not difficult to include the number of edges in the count : let $a_{n,m}$ denote the number of acyclic digraphs with n vertices and m edges, and set

$$A(x,y) = \sum_{n,m \geqslant 0} \frac{1}{n!} (1+y)^{-\binom{n}{2}} a_{n,m} x^n y^m.$$

This generating function is precisely the reciprocal of

$$\phi(x,y) = \sum_{k=0}^{\infty} \frac{(-x)^k}{k!\,(1+y)^{\binom{k}{2}}},$$

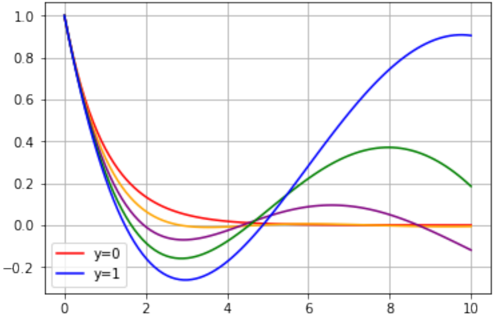
i.e.

$$A(x,y) = \frac{1}{\phi(x,y)},$$

By means of this identity : Bender, Richmond, Robinson and Wormald 1988 obtained an asymptotic formula for $a_{n,m}$ where $m \approx n^2$. This was done by studying the zeros of $\phi(x,y)$ when $y > 0$ is bounded away from zero.

It is known that, for $y > 0$, all zeros of $\phi(x, y)$ (i.e., solutions of $\phi(x, y) = 0$) are real, positive and distinct. This is discussed in Bender, Richmond, Robinson and Wormald 1988

The following figure shows the graphs of the function $\phi(x, y)$ for different values of y. Noting that when $y = 0$, we obtain $\phi(x, 0) = e^{-x}$.



Our results

The following theorem provides asymptotic estimates of the first few zeros of $\phi(x,y)$ as $y \to 0^+$.

Theorem

For a given y, let $\varrho_j(y)$ be the solution to the equation $\phi(x,y) = 0$ that is the j-th closest to zero. If $j \in \mathbb{N}$ is fixed, then we have

$$\varrho_j(y) = \frac{1}{e} y^{-1} - \frac{a_j}{2^{1/3} e} y^{-1/3} - \frac{1}{6e} + O(y^{1/3}), \quad \text{as } y \to 0^+,$$

where a_j is the zero of the Airy function $\mathrm{Ai}(z)$ that is j-th closest to 0. Furthermore, we have the following estimate for the partial derivative of $\phi(x,y)$ at $\varrho_j(y)$:

$$\phi_x(\varrho_j(y), y) \sim -\kappa_j \, y^{1/6} \exp\left(-\tfrac{1}{2} y^{-1} + 2^{-1/3} a_j \, y^{-1/3}\right), \quad \text{as } y \to 0^+,$$

where $\kappa_j = \pi^{1/2} 2^{7/6} e^{11/12} \mathrm{Ai}'(a_j)$.

Our results

The following theorem provides asymptotic estimates of the first few zeros of $\phi(x,y)$ as $y \to 0^+$.

Theorem

For a given y, let $\varrho_j(y)$ be the solution to the equation $\phi(x,y)=0$ that is the j-th closest to zero. If $j \in \mathbb{N}$ is fixed, then we have

$$\varrho_j(y) = \frac{1}{e} y^{-1} - \frac{a_j}{2^{1/3} e} y^{-1/3} - \frac{1}{6e} + O(y^{1/3}), \quad \text{as } y \to 0^+,$$

where a_j is the zero of the Airy function $\mathrm{Ai}(z)$ that is j-th closest to 0. Furthermore, we have the following estimate for the partial derivative of $\phi(x,y)$ at $\varrho_j(y)$:

$$\phi_x(\varrho_j(y), y) \sim -\kappa_j \, y^{1/6} \exp\left(-\tfrac{1}{2} y^{-1} + 2^{-1/3} a_j \, y^{-1/3}\right), \quad \text{as } y \to 0^+,$$

where $\kappa_j = \pi^{1/2} 2^{7/6} e^{11/12} \mathrm{Ai}'(a_j)$.

Our results

Theorem

Let $p = \lambda/n$ with $\lambda \geqslant 0$ fixed. Then, the probability $\mathbb{P}(n,p)$ that a random digraph $\mathcal{D}(n,p)$ is acyclic satisfies the following asymptotic formulas as $n \to \infty$:

$$\mathbb{P}(n,p) \sim \begin{cases} (1-\lambda)e^{\lambda+\lambda^2/2} & \text{if } 0 \leqslant \lambda < 1, \\ \gamma_1 n^{-1/3} & \text{if } \lambda = 1, \\ \gamma_2 n^{-1/3} e^{-c_1 n - c_2 n^{1/3}} & \text{if } \lambda > 1, \end{cases}$$

with

$$\gamma_1 = \frac{2^{-1/3}e^{3/2}}{2\pi} \int_{-\infty}^{\infty} \frac{1}{\mathrm{Ai}(-i2^{1/3}t)} \mathrm{d}t \approx 2.19037,$$

$$\gamma_2 = \frac{2^{-2/3}}{\mathrm{Ai}'(a_1)} \lambda^{5/6} e^{-\lambda^2/4+8\lambda/3-11/12},$$

$$c_1 = \frac{\lambda^2-1}{2\lambda} - \log\lambda,$$

$$c_2 = 2^{-1/3} a_1 \lambda^{-1/3}(1-\lambda),$$

and a_1 is the zero of the Airy function $\mathrm{Ai}(z)$ with the smallest modulus.

In the critical window, we have the following result :

Theorem

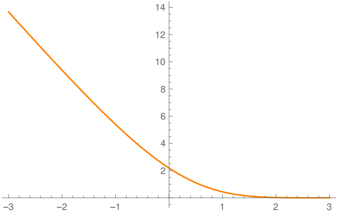
If $np = 1 + \mu n^{-1/3}$ such that $\mu = \mathcal{O}(1)$, then

$$\mathbb{P}(n,p) = (\varphi(\mu) + o(1))n^{-1/3}, \quad \text{as } n \to \infty,$$

where

$$\varphi(\mu) = 2^{-1/3} e^{3/2-\mu^3/6} \times \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \frac{e^{-\mu s}}{\mathrm{Ai}(-2^{1/3}s)} \mathrm{d}s.$$

Here is a numerical plot of $\varphi(\mu)$:



Main ideas

Lemma

The probability $\mathbb{P}(n,p)$ is given by

$$\mathbb{P}(n,p) = n!(1-p)^{\binom{n}{2}}[x^n]A\left(x, \tfrac{p}{1-2p}\right).$$

Idea of proof

$$\mathbb{P}(n,p) = \sum_{m=0}^{\binom{n}{2}} a_{n,m}\ (2p)^m(1-2p)^{\binom{n}{2}-m}\ 2^{-m}$$

$$= (1-2p)^{\binom{n}{2}} \sum_{m=0}^{\binom{n}{2}} a_{n,m}\left(\frac{p}{1-2p}\right)^m.$$

This can be written in terms of

$$[x^n]A(x,y) = \frac{1}{n!}(1+y)^{-\binom{n}{2}} \sum_{m=0}^{\binom{n}{2}} a_{n,m}y^m.$$

Main ideas

- To obtain the estimates of the zeros of $\phi(x, y)$, we first need to find an asymptotic estimate of $\phi(x, y)$ as $y \to 0^+$, where x is a function of y.
- Using Mahler's transformation Mahler (1940), $\phi(x, y)$ can be expressed in integral form as follows :

$$\phi(x, y) = \sqrt{\frac{\log(1+y)}{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\log(1+y)z^2 - x(1+y)^{1/2-iz}\right) dz,$$

After a change of variables, we have

$$\phi(x, y) = \frac{1}{\sqrt{2\pi\alpha}} \int_{-\infty}^{\infty} e^{f(z)} dz,$$

where

$$f(z) := -\frac{1}{2\alpha} z^2 - x\beta e^{-iz}.$$

and

$$\alpha := \log(1+y) \quad \text{and} \quad \beta := \sqrt{1+y}.$$

Main ideas

We have

$$f'(z) = -\frac{1}{\alpha} z + i x \beta e^{-iz}.$$

We can see that $f'(z) = 0$ if and only if

$$i z e^{iz} = -x\alpha\beta.$$

Hence the solutions are given by the branches of the Lambert W function. The fact that the Lambert function $W_0(z)$ has a singularity at $z = -1/e$ suggests that we should choose x and y in such a way that $x\alpha\beta$ is close to $1/e$. Motivated by this, let us define x_0 and δ such that

$$x_0 = \frac{1}{e\alpha\beta} \quad \text{and} \quad x = (1+\delta)x_0.$$

When **δ is zero**, we have a double saddle point at $z = i$, i.e.,

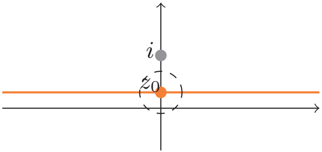
$$f'(i) = f''(i) = 0.$$

Recall that

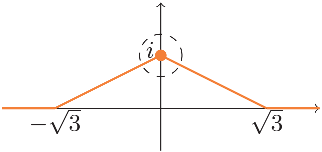
$$\phi(x,y) = \frac{1}{\sqrt{2\pi\alpha}} \int_{-\infty}^{\infty} e^{f(z)} dz,$$

and $z_0 = -iw$ (the solution of the saddle point equation) where $w = W_0(-x\alpha\beta)$.

Choose a path of integration according to the distance between z_0 and i.



For $|z_0 - i| \gg \alpha^{2/3}$



For $|z_0 - i| = \mathcal{O}(\alpha^{2/3})$

We are now able to give asymptotic estimates of $\phi(x,y)$ when $y \to 0^+$ (or $\alpha \to 0^+$), $x = (1+\delta)x_0$, with several ranges of δ.

Main ideas

Theorem

Let $\alpha = \log(1+y)$ and $x = (1+\delta)x_0$, and $w = W_0\big(-(1+\delta)/e\big)$. Then $\phi(x,y)$ satisfies the following asymptotic formulas as $\alpha \to 0$:

(I) If $\delta \geqslant -1$ and satisfies $\delta = -1 + o(1)$, then

$$\phi(x,y) \sim e^{(w+w^2/2)/\alpha}.$$

(II) If $\delta < 0$ and satisfies $\alpha^{2/3} \ll |\delta| \leqslant 1-\varepsilon$ for some constant $\varepsilon > 0$, then

$$\phi(x,y) \sim 2^{5/6}\pi^{1/2}\alpha^{-1/6}\mathrm{Ai}(R)e^{-\frac{1}{2}\alpha^{-1}+\theta\alpha^{-1/3}},$$

where $R = 2^{-2/3}(1+w)^2 w^{-4/3}\alpha^{-2/3}$ and $Ai(z)$ is the Airy function.

(III) If $\delta \sim \theta\alpha^{2/3}$ for a fixed constant $\theta \geqslant 0$, then

$$\phi(x,y) \sim 2^{-1/2}\pi^{-1/2}\alpha^{-1/6}\left(K_1(\theta) + K_2(\theta)\,\alpha^{1/3}\right)\, e^{-\frac{1}{2}\alpha^{-1}-\theta\alpha^{-1/3}},$$

where

$K_1(\theta) = \pi 2^{4/3}\mathrm{Ai}(-2^{1/3}\theta)$, $K_2(\theta) = \frac{5}{3}\pi\, 2^{1/3}\theta^2\mathrm{Ai}(-2^{1/3}\theta) - \frac{1}{3}\pi 2^{2/3}\mathrm{Ai}'(-2^{1/3}\theta)$.

The estimates in the previous theorem can be extended to complex values of x. In particular, if $x = x_0 e^{it\alpha^{2/3}}$, then

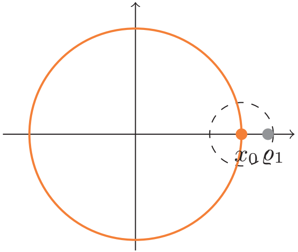
$$\phi(x,y) \sim \pi^{1/2} 2^{5/6} \alpha^{-1/6} \mathrm{Ai}(-i2^{1/3}t)\, e^{-\frac{1}{2}\alpha^{-1} - it\alpha^{-1/3}},$$

For $np = 1 + \mu n^{-1/3}$ and $\mu = \mathcal{O}(1)$, (with $y = p/(1-2p)$ and $\alpha = \log(1+y)$), we get $n = \alpha^{-1} + \mu\alpha^{-2/3}$. So

$$[x^n]A(x,y) = \frac{\alpha^{2/3}\mathrm{Ai}(0)}{2\pi\phi(\rho,y)\rho^n}\left(\int_{-\infty}^{\infty} \frac{e^{-i\mu t}}{\mathrm{Ai}(-i2^{1/3}t)}\mathrm{d}t + o(1)\right).$$

To get this, we use Cauchy integral formula, then apply the saddle point method.

$$[x^n]A(x,y) = \frac{1}{2\pi i}\oint_{|x|=x_0} \frac{1}{\phi(x,y)x^{n+1}}\mathrm{d}x$$

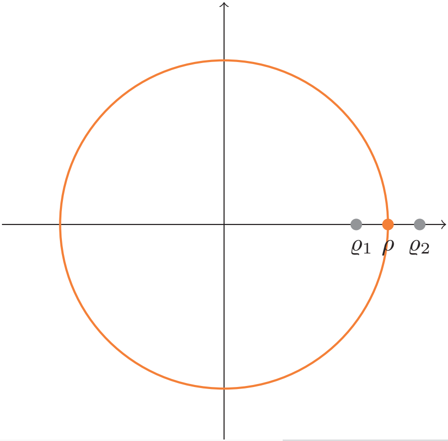


Supercritical

For $np = \lambda$ where $\lambda > 1$. By the residue theorem, we have

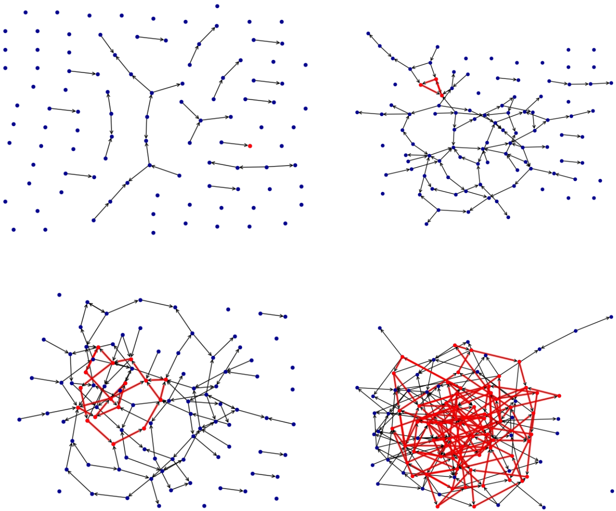
$$[x^n]A(x,y) = -\frac{1}{\varrho_1(y)^{n+1}\phi_x(\varrho_1(y),y)} + \frac{1}{2\pi i}\oint_{|x|=\rho} \frac{1}{\phi(x,y)x^{n+1}}\mathrm{d}x.$$

The main term comes from the 1st term on the right-hand side.



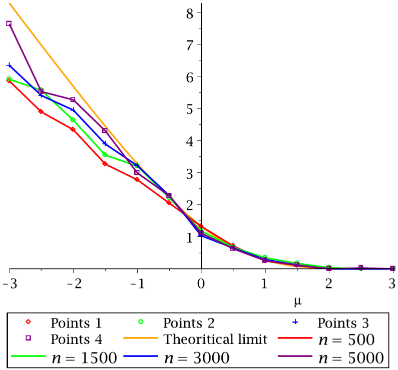
Simulations

For $n = 100$ and $p = 0.006,\ 0.023,\ 0.027,\ 0.033$



In the critical window, $np = 1 + \mu n^{-1/3}$, we have

$$\mathbb{P}(n,p) = (\varphi(\mu) + o(1))n^{-1/3}, \quad \text{as } n \to \infty,$$



Further work

Long version available at arXiv:2009.12127. For three different models of random digraphs, we looked at the probability that the random digraph

- is acyclic,
- is elementary, i.e., the strong components are either single vertices or cycles,
- has one complex strong component.

This is a joint work with Élie de Panafieu, Sergey Dovgal, Vonjy Rasendrahasina, and Stephan Wagner

Next, we will be looking at the parameters of random acyclic digraphs (DAGs),e.g., number of sources/sinks.