

ESTIMER DES PARAMÈTRES DE GRAPHE AVEC DES MARCHES ALÉATOIRES

Anna BEN-HAMOU (Sorbonne Université)

En collaboration avec Roberto OLIVEIRA et Yuval PERES

Journée de Statistique Mathématique

Institut Henri Poincaré

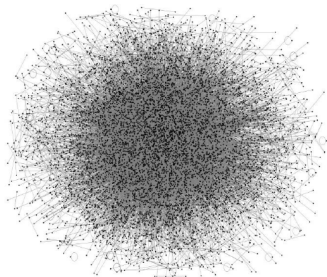
10 janvier 2019



Que peut-on apprendre d'un graphe par exploration locale ?



Que peut-on apprendre d'un graphe par exploration locale ?



Que peut-on apprendre *avant* d'avoir exploré tout le graphe ?

ÉCHANTILLONNAGE PAR MARCHES ALÉATOIRES

Soit $G = (V, E)$ un graphe connexe fini.

$$n = |V| \quad m = |E|$$

On part d'un sommet $x \in V$ fixé.

ÉCHANTILLONNAGE PAR MARCHES ALÉATOIRES

Soit $G = (V, E)$ un graphe connexe fini.

$$n = |V| \quad m = |E|$$

On part d'un sommet $x \in V$ fixé.

- ▶ $X^{(1)}, \dots, X^{(K)}$: MA indép. de longueur t , parties de x .
- ▶ on observe les **labels** et les **degrés** des sommets visités.

ÉCHANTILLONNAGE PAR MARCHES ALÉATOIRES

Soit $G = (V, E)$ un graphe connexe fini.

$$n = |V| \quad m = |E|$$

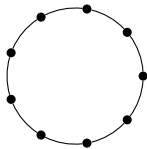
On part d'un sommet $x \in V$ fixé.

- ▶ $X^{(1)}, \dots, X^{(K)}$: MA indép. de longueur t , parties de x .
- ▶ on observe les labels et les degrés des sommets visités.

Objectif : construire un estimateur $\hat{n}_t = \hat{n}_t(X^{(1)}, \dots, X^{(K)})$ tel que pour tout $G = (V, E)$, pour tout $t \geq t(\varepsilon, G)$ et $K \geq K(\varepsilon, G)$, pour tout $x \in V$,

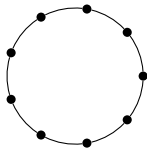
$$\mathbb{P}_x^G \left(\left| \frac{\hat{n}_t}{n} - 1 \right| > \frac{1}{2} \right) \leq \varepsilon,$$

avec $K(\varepsilon, G) \times t(\varepsilon, G)$ aussi petit possible.



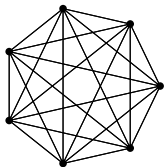
Le cycle

- ↪ Avant $\Omega(n^2)$, impossible de différencier C_n et \mathbb{Z} .
- ↪ Après $O(n^2)$ pas, le graphe est complètement connu.



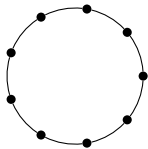
Le cycle

- ↪ Avant $\Omega(n^2)$, impossible de différencier \mathcal{C}_n et \mathbb{Z} .
- ↪ Après $O(n^2)$ pas, le graphe est complètement connu.



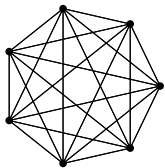
Le graphe complet

- ↪ Jusqu'à $\Omega(\sqrt{n})$, la marche ne voit que des sommets différents.
- ↪ À partir de $O(\sqrt{n})$, le nombre de collisions permet d'estimer n .



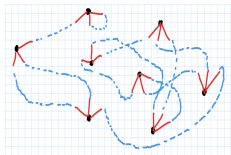
Le cycle

- ↪ Avant $\Omega(n^2)$, impossible de différencier \mathcal{C}_n et \mathbb{Z} .
- ↪ Après $O(n^2)$ pas, le graphe est complètement connu.



Le graphe complet

- ↪ Jusqu'à $\Omega(\sqrt{n})$, la marche ne voit que des sommets différents.
- ↪ À partir de $O(\sqrt{n})$, le nombre de collisions permet d'estimer n .



Graphe 3-régulier « typique »

- ↪ Jusqu'à $\Omega(\sqrt{n})$, le sous-graphe exploré par la marche est un arbre.
- ↪ On verra que $O(\sqrt{n})$ est suffisant.

CONVERGENCE DE MARCHES ALÉATOIRES

$X = (X_t)_{t \geq 0}$ une MA *paresseuse* sur G de matrice de transition P :

$$\forall (x, y) \in V^2, P^t(x, y) \xrightarrow[t \rightarrow \infty]{} \pi(y) = \frac{\deg(y)}{2m}$$

CONVERGENCE DE MARCHES ALÉATOIRES

$X = (X_t)_{t \geq 0}$ une MA *paresseuse* sur G de matrice de transition P :

$$\forall (x, y) \in V^2, P^t(x, y) \xrightarrow[t \rightarrow \infty]{} \pi(y) = \frac{\deg(y)}{2m}$$

\hookrightarrow **Reversibilité** : $\pi(x)P(x, y) = \pi(y)P(y, x)$

CONVERGENCE DE MARCHES ALÉATOIRES

$X = (X_t)_{t \geq 0}$ une MA *paresseuse* sur G de matrice de transition P :

$$\forall (x, y) \in V^2, P^t(x, y) \xrightarrow{t \rightarrow \infty} \pi(y) = \frac{\deg(y)}{2m}$$

\hookrightarrow **Reversibilité** : $\pi(x)P(x, y) = \pi(y)P(y, x)$

\hookrightarrow Vitesse de convergence mesurée par le **temps de mélange uniforme** :

$$t_{\text{unif}} = \inf \left\{ t \geq 0, \max_{x, y \in V} \left| \frac{P^t(x, y)}{\pi(y)} - 1 \right| \leq \frac{1}{4} \right\}$$

CONVERGENCE DE MARCHES ALÉATOIRES

$X = (X_t)_{t \geq 0}$ une MA *paresseuse* sur G de matrice de transition P :

$$\forall (x, y) \in V^2, P^t(x, y) \xrightarrow{t \rightarrow \infty} \pi(y) = \frac{\deg(y)}{2m}$$

↪ **Reversibilité** : $\pi(x)P(x, y) = \pi(y)P(y, x)$

↪ Vitesse de convergence mesurée par le **temps de mélange uniforme** :

$$t_{\text{unif}} = \inf \left\{ t \geq 0, \max_{x, y \in V} \left| \frac{P^t(x, y)}{\pi(y)} - 1 \right| \leq \frac{1}{4} \right\}$$

↪ Ou par le **temps de relaxation** :

$$t_{\text{rel}} = \frac{1}{1 - \lambda_2} \quad \text{où} \quad \lambda_2 = \max\{\lambda \in \text{Sp}(P), \lambda \neq 1\}$$

SE RAMENER À UN ÉCHANTILLON I.I.D.

↔ K MAP indép. parties de $x \in V$, de longueur $t \geq t_{\text{unif}}$.

↔ L'échantillon $(X_t^{(1)}, \dots, X_t^{(K)})$ est (presque) I.I.D. de loi π .

SE RAMENER À UN ÉCHANTILLON I.I.D.

↔ K MAP indép. parties de $x \in V$, de longueur $t \geq t_{\text{unif}}$.

↔ L'échantillon $(X_t^{(1)}, \dots, X_t^{(K)})$ est (presque) I.I.D. de loi π .

Graphes réguliers ($\pi = 1/n$)

Compter le nombre de collisions :

$$C_K = \sum_{i < j} \mathbb{1}_{\{X_t^{(i)} = X_t^{(j)}\}}.$$

↔ *Paradoxe des anniversaires* : $K \asymp \sqrt{n}$

SE RAMENER À UN ÉCHANTILLON I.I.D.

↔ K MAP indép. parties de $x \in V$, de longueur $t \geq t_{\text{unif}}$.

↔ L'échantillon $(X_t^{(1)}, \dots, X_t^{(K)})$ est (presque) I.I.D. de loi π .

Graphes réguliers ($\pi = 1/n$)

Compter le nombre de collisions :

$$C_K = \sum_{i < j} \mathbb{1}_{\{X_t^{(i)} = X_t^{(j)}\}}.$$

↔ *Paradoxe des anniversaires* : $K \asymp \sqrt{n} \implies$ Complexité : $t_{\text{unif}} \sqrt{n}$.

SE RAMENER À UN ÉCHANTILLON I.I.D.

↪ K MAP indép. parties de $x \in V$, de longueur $t \geq t_{\text{unif}}$.

↪ L'échantillon $(X_t^{(1)}, \dots, X_t^{(K)})$ est (presque) I.I.D. de loi π .

Graphes réguliers ($\pi = 1/n$)

Compter le nombre de collisions :

$$C_K = \sum_{i < j} \mathbb{1}_{\{X_t^{(i)} = X_t^{(j)}\}}.$$

↪ *Paradoxe des anniversaires* : $K \asymp \sqrt{n} \implies$ Complexité : $t_{\text{unif}} \sqrt{n}$.

On non-regular graphs

Katzir, Liberty, Somekh, Cosma (2014) ; Kanade, Mallmann-Trenn, Verdugo (2017)

$$K \asymp \sqrt{n} + \frac{m}{n}$$

SE RAMENER À UN ÉCHANTILLON I.I.D.

↔ K MAP indép. parties de $x \in V$, de longueur $t \geq t_{\text{unif}}$.

↔ L'échantillon $(X_t^{(1)}, \dots, X_t^{(K)})$ est (presque) I.I.D. de loi π .

Graphes réguliers ($\pi = 1/n$)

Compter le nombre de collisions :

$$C_K = \sum_{i < j} \mathbb{1}_{\{X_t^{(i)} = X_t^{(j)}\}}.$$

↔ *Paradoxe des anniversaires* : $K \asymp \sqrt{n} \implies$ Complexité : $t_{\text{unif}} \sqrt{n}$.

On non-regular graphs

Katzir, Liberty, Somekh, Cosma (2014) ; Kanade, Mallmann-Trenn, Verdugo (2017)

$$K \asymp \sqrt{n} + \frac{m}{n} \implies \text{Complexité : } t_{\text{unif}} \left(\sqrt{n} + \frac{m}{n} \right)$$

SE RAMENER À UN ÉCHANTILLON I.I.D.

↪ K MAP indép. parties de $x \in V$, de longueur $t \geq t_{\text{unif}}$.

↪ L'échantillon $(X_t^{(1)}, \dots, X_t^{(K)})$ est (presque) I.I.D. de loi π .

Graphes réguliers ($\pi = 1/n$)

Compter le nombre de collisions :

$$C_K = \sum_{i < j} \mathbb{1}_{\{X_t^{(i)} = X_t^{(j)}\}}.$$

↪ *Paradoxe des anniversaires* : $K \asymp \sqrt{n} \implies$ Complexité : $t_{\text{unif}} \sqrt{n}$.

On non-regular graphs

Katzir, Liberty, Somekh, Cosma (2014) ; Kanade, Mallmann-Trenn, Verdugo (2017)

$$K \asymp \sqrt{n} + \frac{m}{n} \implies \text{Complexité : } t_{\text{unif}} \left(\sqrt{n} + \frac{m}{n} \right)$$

↪ Peut-on améliorer la complexité en utilisant toute la trajectoire ?

↪ Le facteur t_{unif} est-il nécessaire ?

ESTIMATION DU NOMBRE D'ARÊTES : INTERSECTIONS PONDÉRÉES

Soient X et Y deux MAP indép. parties de $x \in V$.

$$\mathbf{I}_t = \sum_{i,j=0}^{t-1} \frac{1}{\deg(X_i)} \mathbb{1}_{\{X_i=Y_j\}}$$

ESTIMATION DU NOMBRE D'ARÊTES : INTERSECTIONS PONDÉRÉES

Soient X et Y deux MAP indép. parties de $x \in V$.

$$\mathbf{I}_t = \sum_{i,j=0}^{t-1} \frac{1}{\deg(X_i)} \mathbb{1}_{\{X_i=Y_j\}}$$

Proposition (B., Oliveira et Peres, 2018)

$$\frac{t^2}{2m} \leq \mathbb{E}_{x,x} \mathbf{I}_t \leq \frac{t^2}{2m} + \frac{16t_{\text{rel}}^{3/2}}{d_{\text{MIN}}},$$

et

$$\max_{x \in V} \mathbb{E}_{x,x} \mathbf{I}_t^2 \lesssim \left(\max_{u \in V} \mathbb{E}_{u,u} \mathbf{I}_t \right)^2.$$

\hookrightarrow Concentration de \mathbf{I}_t dès que $t \gtrsim t_{\text{rel}}^{3/4} \sqrt{m/d_{\text{MIN}}}$.

INTERSECTIONS PONDÉRÉES

$$\frac{t^2}{2m} \leq \mathbb{E}_{x,x} \mathbf{I}_t \leq \frac{t^2}{2m} + \frac{16t_{\text{rel}}^{3/2}}{d_{\text{MIN}}},$$

Utiliser la réversibilité

$$\mathbb{E}_{x,x} \mathbf{I}_t = \sum_{i,j=0}^{t-1} \frac{P^{i+j}(x,x)}{\deg(x)} \lesssim \sum_{i+j \leq t_{\text{unif}}} \frac{P^{i+j}(x,x)}{\deg(x)} + \frac{t^2}{2m}.$$

Borne sur les probabilités de retour

$$\forall s \leq t_{\text{unif}}, P^s(x,x) \lesssim \frac{\deg(x)}{d_{\text{MIN}} \sqrt{s+1}} \quad \text{Oliveira et Peres, 2018}$$

Aldous et Fill

UN PREMIER ESTIMATEUR

On considère K paires de MAP $\{(X^{(k)}, Y^{(k)})\}_{k=1}^K$ et l'on définit

$$\widehat{m}_t = \frac{t^2}{\frac{2}{K} \sum_{k=1}^K \mathbf{I}_t^{(k)}}$$

UN PREMIER ESTIMATEUR

On considère K paires de MAP $\{(X^{(k)}, Y^{(k)})\}_{k=1}^K$ et l'on définit

$$\widehat{m}_t = \frac{t^2}{\frac{2}{K} \sum_{k=1}^K \mathbf{I}_t^{(k)}}$$

Proposition (B., Oliveira and Peres, 2018)

Pour tout $G = (V, E)$ connexe et pour tout $x \in V$,

$$\forall t \geq t_{\text{rel}}^{3/4} \sqrt{m/d_{\text{MIN}}}, \quad \mathbb{P}_x \left(\left| \frac{\widehat{m}_t}{m} - 1 \right| > \frac{1}{2} \right) = O \left(\frac{1}{K} \right).$$

UN PREMIER ESTIMATEUR

On considère K paires de MAP $\{(X^{(k)}, Y^{(k)})\}_{k=1}^K$ et l'on définit

$$\widehat{m}_t = \frac{t^2}{\frac{2}{K} \sum_{k=1}^K \mathbf{I}_t^{(k)}}$$

Proposition (B., Oliveira and Peres, 2018)

Pour tout $G = (V, E)$ connexe et pour tout $x \in V$,

$$\forall t \geq t_{\text{rel}}^{3/4} \sqrt{m/d_{\text{MIN}}}, \quad \mathbb{P}_x \left(\left| \frac{\widehat{m}_t}{m} - 1 \right| > \frac{1}{2} \right) = O \left(\frac{1}{K} \right).$$

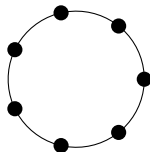
Dans le cas des graphes réguliers,

$$\forall t \geq t_{\text{rel}}^{3/4} \sqrt{n}, \quad \mathbb{P}_x \left(\left| \frac{\widehat{n}_t}{n} - 1 \right| > \frac{1}{2} \right) = O \left(\frac{1}{K} \right).$$

BORNE INFÉRIEURE POUR LES GRAPHES RÉGULIERS

La borne $t_{\text{rel}}^{3/4} \sqrt{n}$ est atteinte par le cycle :

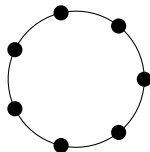
$$t_{\text{rel}}^{3/4} \sqrt{n} = (n^2)^{3/4} \sqrt{n} = n^2$$



BORNE INFÉRIEURE POUR LES GRAPHES RÉGULIERS

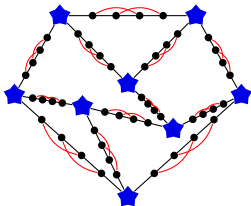
La borne $t_{\text{rel}}^{3/4} \sqrt{n}$ est atteinte par le cycle :

$$t_{\text{rel}}^{3/4} \sqrt{n} = (n^2)^{3/4} \sqrt{n} = n^2$$



Borne inférieure pour tous les temps de relaxation

1. Soit \mathcal{E}_k un graphe 3-régulier « typique » à k sommets ;
2. On remplace chaque arête de \mathcal{E}_k par un chemin de longueur $\ell \geq 1$;
3. On s'arrange pour que le graphe obtenu soit 3-régulier.

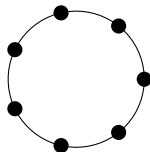


$$\begin{aligned} n &\asymp kl \\ t_{\text{rel}} &\asymp \ell^2 \end{aligned}$$

BORNE INFÉRIEURE POUR LES GRAPHES RÉGULIERS

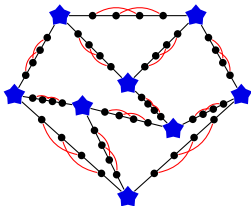
La borne $t_{\text{rel}}^{3/4} \sqrt{n}$ est atteinte par le cycle :

$$t_{\text{rel}}^{3/4} \sqrt{n} = (n^2)^{3/4} \sqrt{n} = n^2$$



Borne inférieure pour tous les temps de relaxation

1. Soit \mathcal{E}_k un graphe 3-régulier « typique » à k sommets ;
2. On remplace chaque arête de \mathcal{E}_k par un chemin de longueur $\ell \geq 1$;
3. On s'arrange pour que le graphe obtenu soit 3-régulier.



$$\begin{aligned} n &\asymp k\ell \\ t_{\text{rel}} &\asymp \ell^2 \end{aligned}$$

↪ Pour estimer n , il faut pouvoir estimer k , et cela prend au moins

$$\ell^2 \sqrt{k} \asymp t_{\text{rel}}^{3/4} \sqrt{n}$$

UN DEUXIÈME ESTIMATEUR

$$\mathbf{J}_t = \sum_{i,j=t_{\text{unif}}}^{t_{\text{unif}}+t-1} \frac{1}{\deg(X_i)} \mathbb{1}_{\{X_i=Y_j\}}$$

UN DEUXIÈME ESTIMATEUR

$$\mathbf{J}_t = \sum_{i,j=t_{\text{unif}}}^{t_{\text{unif}}+t-1} \frac{1}{\deg(X_i)} \mathbb{1}_{\{X_i=Y_j\}}$$

Proposition

$$\mathbb{E}_{x,x} \mathbf{J}_t \asymp \frac{t^2}{2m} \quad \text{et} \quad \mathbb{E}_{x,x} \mathbf{J}_t^2 \lesssim \left(\frac{t^2}{2m} + \frac{nt_{\text{rel}}^{5/3}}{m} \right)^2.$$

UN DEUXIÈME ESTIMATEUR

$$\mathbf{J}_t = \sum_{i,j=t_{\text{unif}}}^{t_{\text{unif}}+t-1} \frac{1}{\deg(X_i)} \mathbb{1}_{\{X_i=Y_j\}}$$

Proposition

$$\mathbb{E}_{x,x} \mathbf{J}_t \asymp \frac{t^2}{2m} \quad \text{et} \quad \mathbb{E}_{x,x} \mathbf{J}_t^2 \lesssim \left(\frac{t^2}{2m} + \frac{nt_{\text{rel}}^{5/3}}{m} \right)^2.$$

Borne sur la somme des probabilités de retour

$$\sum_{x \in V} P^t(x, x) \leq 1 + \frac{13n}{t^{1/3}} \quad \text{Lyons et Oveis Gharan, 2017}$$

UN DEUXIÈME ESTIMATEUR

On considère K paires de MAP $\{(X^{(k)}, Y^{(k)})\}_{k=1}^K$ et l'on définit

$$\tilde{m}_t = \frac{t^2}{\frac{2}{K} \sum_{k=1}^K \mathbf{J}_t^{(k)}}$$

UN DEUXIÈME ESTIMATEUR

On considère K paires de MAP $\{(X^{(k)}, Y^{(k)})\}_{k=1}^K$ et l'on définit

$$\tilde{m}_t = \frac{t^2}{\frac{2}{K} \sum_{k=1}^K \mathbf{J}_t^{(k)}}$$

Proposition (B., Oliveira and Peres, 2018)

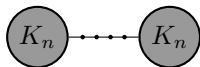
Pour tout $G = (V, E)$ connexe et pour tout $x \in V$,

$$\forall t \geq t_{\text{rel}}^{5/6} \sqrt{n}, \quad \mathbb{P}_x \left(\left| \frac{\tilde{m}_t}{m} - 1 \right| > \frac{1}{2} \right) = O \left(\frac{1}{K} \right).$$

BORNE INFÉRIEURE

↪ La borne $t_{\text{rel}}^{5/6} \sqrt{n}$ est atteinte par l'haltère :

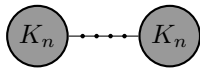
$$t_{\text{rel}}^{5/6} \sqrt{n} = (n^3)^{5/6} \sqrt{n} = n^3$$



BORNE INFÉRIEURE

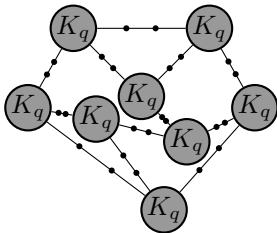
↪ La borne $t_{\text{rel}}^{5/6} \sqrt{n}$ est atteinte par l'haltère :

$$t_{\text{rel}}^{5/6} \sqrt{n} = (n^3)^{5/6} \sqrt{n} = n^3$$



Borne inférieure pour tous les temps de relaxation

1. Soit \mathcal{E}_k un graphe 3-régulier « typique » à k sommets ;
2. On remplace chaque sommet de \mathcal{E}_k par une clique K_q de taille q ;
3. On remplace chaque arête de \mathcal{E}_k par un chemin de longueur q .

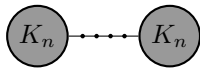


$$\begin{aligned} n &\asymp kq \\ m &\asymp kq^2 \\ t_{\text{rel}} &\asymp q^3 \end{aligned}$$

BORNE INFÉRIEURE

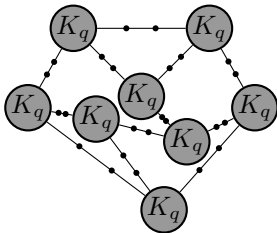
↪ La borne $t_{\text{rel}}^{5/6} \sqrt{n}$ est atteinte par l'haltère :

$$t_{\text{rel}}^{5/6} \sqrt{n} = (n^3)^{5/6} \sqrt{n} = n^3$$



Borne inférieure pour tous les temps de relaxation

1. Soit \mathcal{E}_k un graphe 3-régulier « typique » à k sommets ;
2. On remplace chaque sommet de \mathcal{E}_k par une clique K_q de taille q ;
3. On remplace chaque arête de \mathcal{E}_k par un chemin de longueur q .



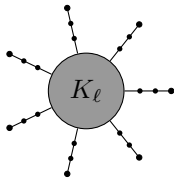
$$\begin{aligned} n &\asymp kq \\ m &\asymp kq^2 \\ t_{\text{rel}} &\asymp q^3 \end{aligned}$$

↪ Pour estimer m , il faut pouvoir estimer k et cela prend au moins

$$q^3 \sqrt{k} \asymp t_{\text{rel}}^{5/6} \sqrt{n}$$

DES ARÊTES AUX SOMMETS

$\hookrightarrow t_{\text{unif}}^{5/6} \sqrt{n}$ n'est pas toujours suffisant pour le nombre de sommets.



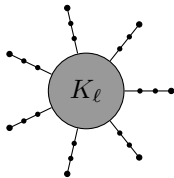
1. On part d'un graphe complet K_ℓ ;
2. On rajoute des chemins de longueur q à chaque sommet, avec $q \ll \ell$.

$$n \asymp \ell q \quad m \asymp \ell^2 \quad t_{\text{unif}} \asymp q^2$$

$$T(n) \asymp q\ell \gg t_{\text{unif}}^{5/6} \sqrt{n} \text{ pour } q \text{ assez petit.}$$

DES ARÊTES AUX SOMMETS

↪ $t_{\text{unif}}^{5/6} \sqrt{n}$ n'est pas toujours suffisant pour le nombre de sommets.



1. On part d'un graphe complet K_ℓ ;
2. On rajoute des chemins de longueur q à chaque sommet, avec $q \ll \ell$.

$$n \asymp \ell q \quad m \asymp \ell^2 \quad t_{\text{unif}} \asymp q^2$$

$$T(n) \asymp q\ell \gg t_{\text{unif}}^{5/6} \sqrt{n} \text{ pour } q \text{ assez petit.}$$

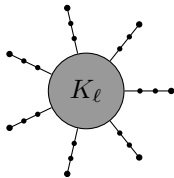
↪ Cependant, si l'on sait estimer m , il suffit d'estimer le **degré moyen**, ce qui peut se faire en

$$O\left(t_{\text{unif}} \frac{m}{n}\right)$$

(et cette borne est atteinte par l'exemple précédent : $T(n) \asymp q\ell \asymp t_{\text{unif}} \frac{m}{n}$.)

DES ARÊTES AUX SOMMETS

↪ $t_{\text{unif}}^{5/6} \sqrt{n}$ n'est pas toujours suffisant pour le nombre de sommets.



1. On part d'un graphe complet K_ℓ ;
2. On rajoute des chemins de longueur q à chaque sommet, avec $q \ll \ell$.

$$n \asymp \ell q \quad m \asymp \ell^2 \quad t_{\text{unif}} \asymp q^2$$

$$T(n) \asymp q\ell \gg t_{\text{unif}}^{5/6} \sqrt{n} \text{ pour } q \text{ assez petit.}$$

↪ Cependant, si l'on sait estimer m , il suffit d'estimer le **degré moyen**, ce qui peut se faire en

$$O\left(t_{\text{unif}} \frac{m}{n}\right)$$

(et cette borne est atteinte par l'exemple précédent : $T(n) \asymp q\ell \asymp t_{\text{unif}} \frac{m}{n}$.)

↪ Pour résumer, on obtient une complexité en

$$t_{\text{rel}}^{5/6} \sqrt{n} + t_{\text{unif}} \frac{m}{n}$$

ALGORITHMES AVEC TEMPS D'ARRÊT

↪ Supposons que l'on dispose d'une borne sup T sur t_{unif} .

↪ Pour $\varepsilon > 0$, et $q = 0, 1, 2, \dots$

1. soit $\hat{m} = 2^q$ la supposition courante pour m ;
2. on fait une série d'expériences indépendantes avec les intersections ;
3. si l'on obtient assez de succès, on s'arrête : $\hat{m} = 2^q$;
4. sinon $q \leftarrow q + 1$.

ALGORITHMES AVEC TEMPS D'ARRÊT

↔ Supposons que l'on dispose d'une borne sup T sur t_{unif} .

↔ Pour $\varepsilon > 0$, et $q = 0, 1, 2, \dots$

1. soit $\hat{m} = 2^q$ la supposition courante pour m ;
2. on fait une série d'expériences indépendantes avec les intersections ;
3. si l'on obtient assez de succès, on s'arrête : $\hat{m} = 2^q$;
4. sinon $q \leftarrow q + 1$.

↔ Cet algorithme est tel que

- avec probabilité $1 - \varepsilon$, l'output \hat{m} vérifie $\left| \frac{\hat{m}}{m} - 1 \right| \leq \frac{1}{2}$;
- le temps d'exécution moyen est $O\left(T^{3/4} \sqrt{m} \log \log m\right)$.

ALGORITHMES AVEC TEMPS D'ARRÊT

Supposons que l'on connaisse m (ou une bonne approximation).

On veut estimer $t_x(1/4) = \inf \{t \geq 0, d_x(t) \leq 1/4\}$, où

$$d_x(t) = \left(\sum_y \pi(y) \left(\frac{P^t(x, y)}{\pi(y)} - 1 \right)^2 \right)^{1/2} .$$

ALGORITHMES AVEC TEMPS D'ARRÊT

Supposons que l'on connaisse m (ou une bonne approximation).

On veut estimer $t_x(1/4) = \inf \{t \geq 0, d_x(t) \leq 1/4\}$, où

$$d_x(t) = \left(\sum_y \pi(y) \left(\frac{P^t(x, y)}{\pi(y)} - 1 \right)^2 \right)^{1/2}.$$

Connection with intersections

$$\mathbb{E}_x \mathbf{I}_t = \sum_{i,j=0}^{t-1} \frac{d_x\left(\frac{i+j}{2}\right)^2 + 1}{2m}$$

ALGORITHMES AVEC TEMPS D'ARRÊT

Supposons que l'on connaisse m (ou une bonne approximation).

On veut estimer $t_x(1/4) = \inf \{t \geq 0, d_x(t) \leq 1/4\}$, où

$$d_x(t) = \left(\sum_y \pi(y) \left(\frac{P^t(x, y)}{\pi(y)} - 1 \right)^2 \right)^{1/2}.$$

Connection with intersections

$$\mathbb{E}_x \mathbf{I}_t = \sum_{i,j=0}^{t-1} \frac{d_x\left(\frac{i+j}{2}\right)^2 + 1}{2m}$$

↪ On peut construire un algorithme tel que

- avec probabilité $1 - \varepsilon$, l'output $\widehat{t_x(1/4)}$ vérifie

$$\frac{t_x(1/4)}{2} \leq \widehat{t_x(1/4)} \leq t_x(1/8)$$

- le temps moyen d'exécution est $O\left(t_x(1/8)^{3/4} \sqrt{m} \log \log(t_x(1/8))\right)$.

MERCI!