

Graphes inhomogènes : comportement au seuil critique

Lucas MERCIER

ALEA 2010

24 mars 2010

Modèle d'Erdős-Rényi

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

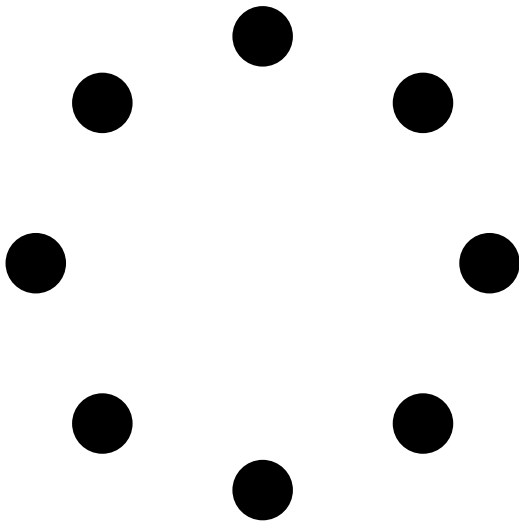
Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du seuil
critique

Processus auxiliaires
utilisés
Renormalisation
Dérive
Covariance



On considère
 n sommets.

Modèle d'Erdős-Rényi

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

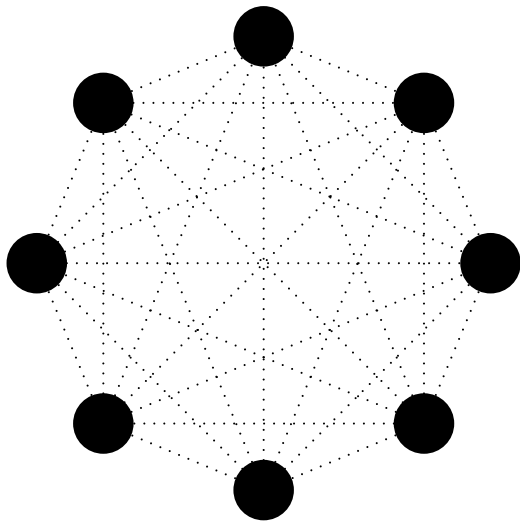
Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du seuil
critique

Processus auxiliaires
utilisés
Renormalisation
Dérive
Covariance



On considère
toutes les
arêtes
possibles.

Modèle d'Erdős-Rényi

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

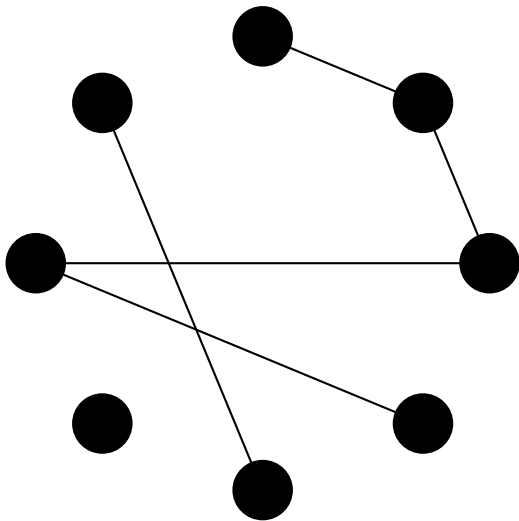
Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du seuil
critique

Processus auxiliaires
utilisés
Renormalisation
Dérive
Covariance



On garde
chaque arête
indépendam-
ment avec
probabilité p .

Limitations du modèle

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du seuil
critique

Processus auxiliaires
utilisés

Renormalisation

Dérive

Covariance

- ▶ Tous les sommets sont équivalents. La loi du degré d'un sommet est proche d'une loi de Poisson.
- ▶ On veut pouvoir modéliser des graphes plus généraux, par exemple un graphe bipartite, ou un graphe avec des serveurs centraux très connectés et des sommets peu connectés.
- ▶ On va donc introduire différents types de sommets, et faire dépendre la probabilité de présence d'une arête du type de ses extrémités.

Graphe inhomogène

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

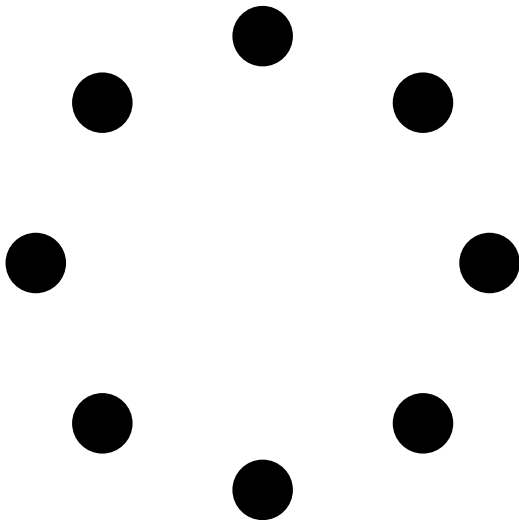
Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du seuil
critique

Processus auxiliaires
utilisés
Renormalisation
Dérive
Covariance



On considère
 n sommets.

Graphe inhomogène

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

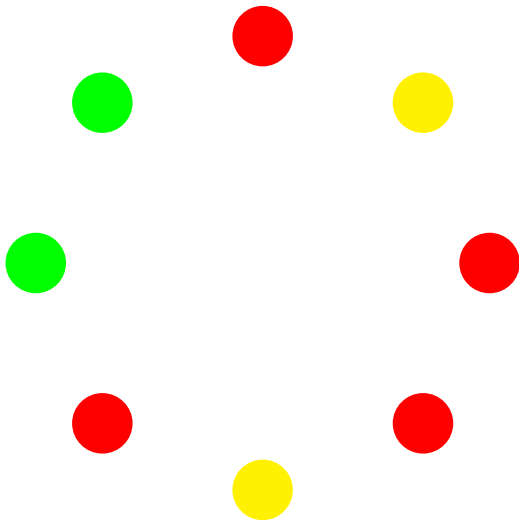
Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du seuil
critique

Processus auxiliaires
utilisés
Renormalisation
Dérive
Covariance



On choisit
pour chaque
sommet un
type.

Graphes
inhomogènes :
comportement au
seuil critique

Graphe inhomogène

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

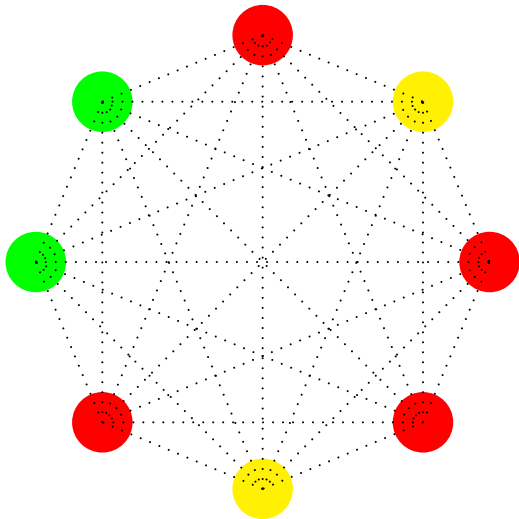
Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du seuil
critique

Processus auxiliaires
utilisés
Renormalisation
Dérive
Covariance



On considère
toutes les
arêtes
possibles.

Graphe inhomogène

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

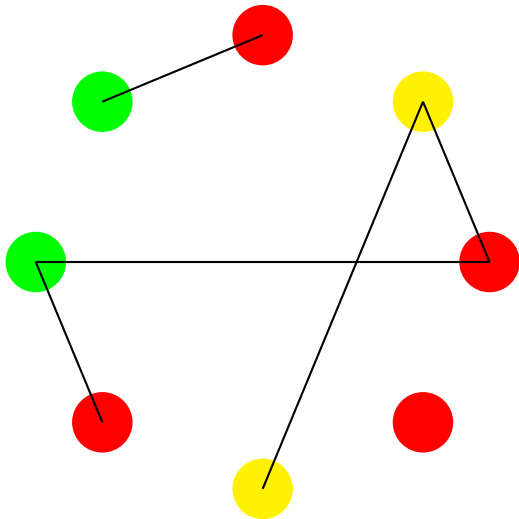
Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du seuil
critique

Processus auxiliaires
utilisés
Renormalisation
Dérive
Covariance



On garde
chaque arête
indépendam-
ment avec
probabilité
ne dépendant
que du type
des extrémités
de l'arête.

Dans ce modèle, si l'on a ℓ types autorisés, on a deux paramètres :

- ▶ Un vecteur r à ℓ coordonnées positives et de somme 1, qui sera la proportion de chaque type.
- ▶ Une matrice $\ell \times \ell$ symétrique réelle à coefficients positifs, qui donnera la probabilité d'existence des arêtes, que l'on notera C .

Processus d'exploration des composantes connexes.

On choisit
un sommet
au hasard,
dont on va
explorer la
composante
connexe.



Graphes
inhomogènes :
comportement au
seuil critique

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

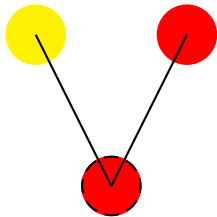
Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du seuil
critique

Processus auxiliaires
utilisés
Renormalisation
Dérive
Covariance



On considère
tous les
sommets
voisins, qui
n'ont pas
encore été
découverts,
et on les
place comme
fils du
premier
sommets.

Graphes
inhomogènes :
comportement au
seuil critique

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

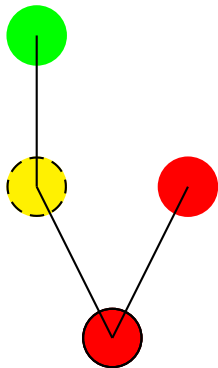
Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du
seuil critique

Processus auxiliaires
utilisés
Renormalisation
Dérive
Covariance



On fait de
même avec
le premier
sommet que
l'on vient de
découvrir.

Graphes
inhomogènes :
comportement au
seuil critique

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

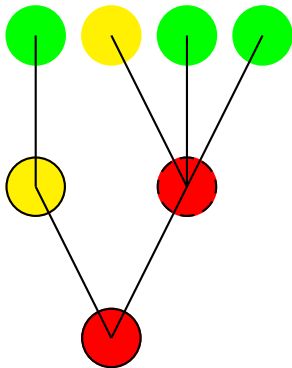
Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du
seuil critique

Processus auxiliaires
utilisés
Renormalisation
Dérive
Covariance



Graphes
inhomogènes :
comportement au
seuil critique

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

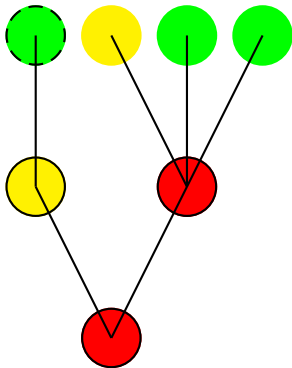
Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du seuil
critique

Processus auxiliaires
utilisés
Renormalisation
Dérive
Covariance



Graphes
inhomogènes :
comportement au
seuil critique

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

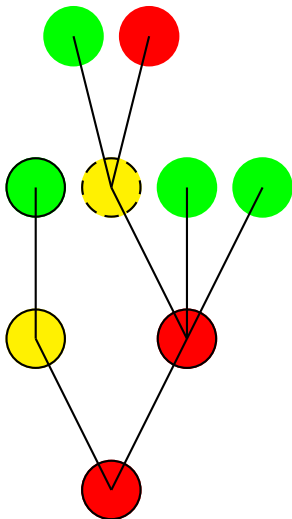
Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du seuil
critique

Processus auxiliaires
utilisés
Renormalisation
Dérive
Covariance



Graphes
inhomogènes :
comportement au
seuil critique

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

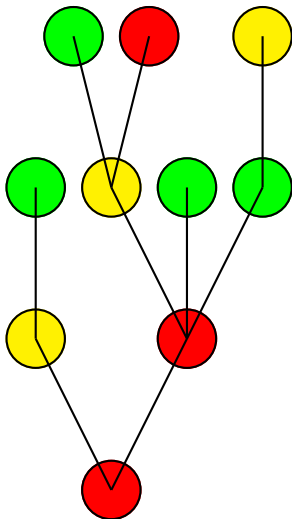
Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du seuil
critique

Processus auxiliaires
utilisés
Renormalisation
Dérive
Covariance



Lorsque l'on
n'a plus de
sommets en
réserve (fin
d'une
composante
connexe), on
rajoute un
nouveau
sommet.

Graphes
inhomogènes :
comportement au
seuil critique

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

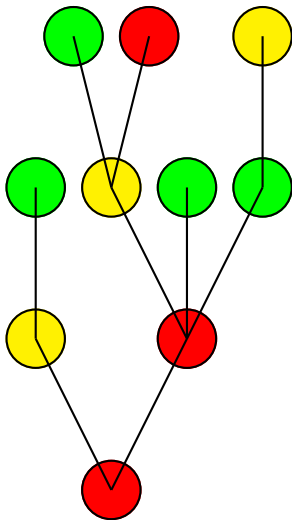
Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du seuil
critique

Processus auxiliaires
utilisés
Renormalisation
Dérive
Covariance



Et on
recommence
le processus.



Graphes
inhomogènes :
comportement au
seuil critique

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

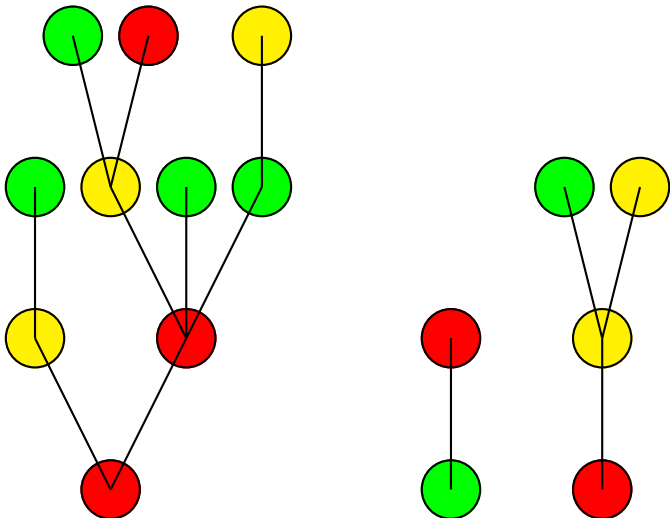
Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du
seuil critique

Processus auxiliaires
utilisés
Renormalisation
Dérive
Covariance



Intérêt de cet algorithme

- ▶ La taille des composantes connexes est aussi la taille des arbres du processus d'exploration.
- ▶ On ne considère jamais deux fois la même arête, ce qui permet d'avoir des propriétés d'indépendance.

Intérêt de cet algorithme

- ▶ La taille des composantes connexes est aussi la taille des arbres du processus d'exploration.
- ▶ On ne considère jamais deux fois la même arête, ce qui permet d'avoir des propriétés d'indépendance.

On veut savoir si l'on a de « grands » arbres. Pour cela, on a besoin de connaître la loi du nombre de fils d'un sommet.

Si l'on considère un sommet de type i , quelle est la loi du nombre de fils de type j ?

Intérêt de cet algorithme

- ▶ La taille des composantes connexes est aussi la taille des arbres du processus d'exploration.
- ▶ On ne considère jamais deux fois la même arête, ce qui permet d'avoir des propriétés d'indépendance.

On veut savoir si l'on a de « grands » arbres. Pour cela, on a besoin de connaître la loi du nombre de fils d'un sommet.

Si l'on considère un sommet de type i , quelle est la loi du nombre de fils de type j ?

$$\mathcal{B}(n_j - \nu_j, C_{i,j})$$

Avec n_j le nombre de sommets de type j initialement présents et ν_j le nombre de sommets de type j déjà découverts.

Intérêt de cet algorithme

- ▶ La taille des composantes connexes est aussi la taille des arbres du processus d'exploration.
- ▶ On ne considère jamais deux fois la même arête, ce qui permet d'avoir des propriétés d'indépendance.

On veut savoir si l'on a de « grands » arbres. Pour cela, on a besoin de connaître la loi du nombre de fils d'un sommet.

Si l'on considère un sommet de type i , quelle est la loi du nombre de fils de type j ?

$$\mathcal{B}(n_j - \nu_j, C_{i,j}) \simeq \mathcal{B}(n_j, C_{i,j})$$

Avec n_j le nombre de sommets de type j initialement présents et ν_j le nombre de sommets de type j déjà découverts.

Au début de l'exploration, on peut donc approcher l'arbre d'exploration par un arbre de Galton-Watson multitype, dont la matrice d'espérance vaut :

$$C_{i,j}n_j \simeq C_{i,j}r_jn =: Q_{i,j}$$

La bonne renormalisation est donc de prendre

$$C = \frac{1}{n}\mathcal{C}$$

Ainsi Q ne dépend plus de n .

Théorème (Söderberg (2002))

- ▶ *Si la plus grande valeur propre de Q est inférieure à 1, alors la taille de la plus grande composante connexe est $o_p(n)$.*
- ▶ *Si une valeur propre de Q est strictement plus grande que 1, il existe une composante connexe de taille $\Theta(n)$.*

Le cas critique

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du seuil
critique

Processus auxiliaires
utilisés

Renormalisation

Dérive

Covariance

On se donne \mathcal{C} et r tels que la plus grande valeur propre de Q soit 1.

On prend $C = \mathcal{C}(\frac{1}{n} + tn^{-\frac{4}{3}})$ comme renormalisation.

Dans le cas simple (Erdős-Rényi), Aldous a montré que l'on pouvait relier la taille des plus grandes composantes connexes, renormalisées par un facteur $n^{\frac{2}{3}}$, à la longueur d'excursions browniennes.

Et dans le cas multitype ?

On sait, qu'au premier ordre, on est au seuil critique pour l'arbre d'exploration. Nous allons introduire les processus suivants Z^j pour mesurer l'écart à l'espérance :

- ▶ $Z_0^j = 0$.
- ▶ Si à la k ième étape, on explore le sommet x_k de type i , alors
$$Z_{k+1}^j = Z_k^j - Q_{i,j} + N_k^j$$
avec N_k^j le nombre de fils de x_k de type j .

Remarque

Dans le cas simple ($\ell = 1$), il s'agit de la marche de Łukasiewicz.

Marche de Łukasiewicz

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du seuil
critique

Processus auxiliaires
utilisés

Renormalisation

Dérive

Covariance

Lors de l'exploration d'une forêt, on définit L_k la marche de Łukasiewicz de la manière suivante :

$$L_0 = 0$$

$L_{k+1} = L_k - 1 + N_k$ avec N_k le nombre de fils du k ème sommet.

L caractérise l'arbre

On a en particulier les deux propriétés suivantes :

- ▶ $-\min_{j < k} L_j$ est le nombre de composantes entièrement explorées avant le sommet k .
- ▶ $L_k - \min_{j < k} L_j$ est le nombre de sommets en réserve au temps k (sans compter le sommet utilisé à la date k).

Pourquoi considérer ces processus ?

Ils permettent aussi d'estimer le nombre de composantes et de sommets en réserve.

Soit v le vecteur propre à droite de Q pour la valeur propre 1.

$$Y_k = \sum_{j=1}^{\ell} v_j Z_k^j$$

- ▶ Si x_k est de type i , $Y_{k+1} = Y_k - v_i + \sum v_j N_k^j$.
- ▶ Si k et k' sont les premiers sommets de deux composantes consécutives, et si x_k est de type i , alors $Y_{k'} = Y_k - v_i$.
- ▶ Si k est le premier sommet d'une composante, et k' un autre sommet de cette composante, $Y_{k'} = Y_k - v_i + \sum v_j R^j$ avec R^j le nombre de sommets de type j en réserve à l'étape k' .

Renormalisation

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du seuil
critique

Processus auxiliaires
utilisés

Renormalisation

Dérive

Covariance

$$\tilde{Z}^j(s) = n^{-\frac{1}{3}} (Z^j(n^{\frac{2}{3}}s))$$

Renormalisation

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du seuil
critique

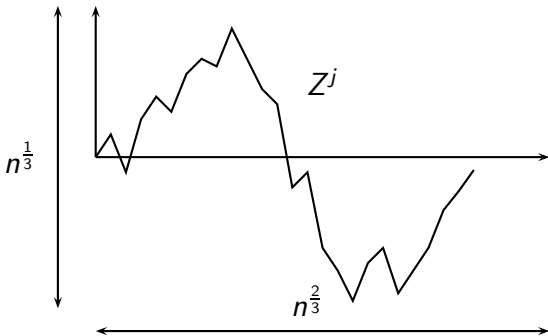
Processus auxiliaires
utilisés

Renormalisation

Dérive

Covariance

$$\tilde{Z}^j(s) = n^{-\frac{1}{3}} (Z^j(n^{\frac{2}{3}}s))$$



Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du seuil
critique

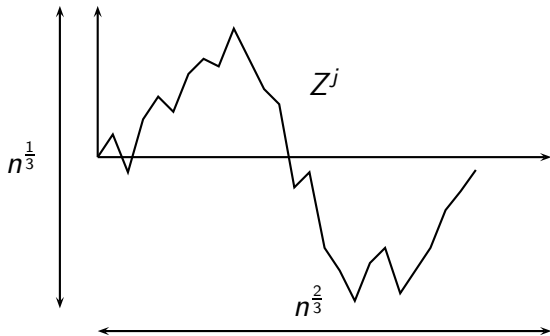
Processus auxiliaires
utilisés

Renormalisation

Dérive

Covariance

$$\tilde{Z}^j(s) = n^{-\frac{1}{3}} (Z^j(n^{\frac{2}{3}}s))$$



Le but est de montrer que cela converge vers un brownien avec une dérive.

On s'intéresse à l'évolution entre k et $k + 1$, avec $k = \lfloor sn^{\frac{2}{3}} \rfloor$ (ce qui fait une variation en temps de $n^{-\frac{2}{3}}$ pour le processus renormalisé).

$$Z_{k+1}^j - Z_k^j = -Q_{i,j} + N_k^j$$

avec N_k^j le nombre de fils de type j qu'a le sommet x_k .
La loi de N_k^j est une binomiale de paramètres :

- ▶ $n_j - \nu_j$ (le nombre d'arêtes potentielles)
- ▶ $\mathcal{C}_{i,j}(\frac{1}{n} + tn^{-\frac{4}{3}})$ (la probabilité d'une arête).

L'espérance vaut donc :

$$- Q_{i,j} + (n_j - \nu_j) \mathcal{C}_{i,j} \left(\frac{1}{n} + tn^{-\frac{4}{3}} \right)$$

L'espérance vaut donc :

$$- Q_{i,j} + (n_j - \nu_j) \mathcal{C}_{i,j} \left(\frac{1}{n} + tn^{-\frac{4}{3}} \right)$$

Le terme du premier ordre est annulé (par construction).

L'espérance vaut donc :

$$- Q_{i,j} + (n_j - \nu_j) \mathcal{C}_{i,j} \left(\frac{1}{n} + tn^{-\frac{4}{3}} \right)$$

Le terme du premier ordre est annulé (par construction).

L'ordre suivant est en $n^{-\frac{1}{3}}$ avec deux termes :

- ▶ le terme en $tn^{-\frac{4}{3}}$ dans la probabilité d'existence d'une arête, multiplié par $n_j \simeq r_j n = \Theta(n)$.

L'espérance vaut donc :

$$- Q_{i,j} + (n_j - \nu_j) \mathcal{C}_{i,j} \left(\frac{1}{n} + tn^{-\frac{4}{3}} \right)$$

Le terme du premier ordre est annulé (par construction).

L'ordre suivant est en $n^{-\frac{1}{3}}$ avec deux termes :

- ▶ le terme en $tn^{-\frac{4}{3}}$ dans la probabilité d'existence d'une arête, multiplié par $n_j \simeq r_j n = \Theta(n)$.
- ▶ le termes dû aux sommets déjà explorés (ν_j), qui est de l'ordre de k , avec $k = \Theta(n^{\frac{2}{3}})$ d'après la renormalisation, multiplié par $\frac{1}{n}$.

L'espérance vaut donc :

$$- Q_{i,j} + (n_j - \nu_j) \mathcal{C}_{i,j} \left(\frac{1}{n} + tn^{-\frac{4}{3}} \right)$$

Le terme du premier ordre est annulé (par construction).

L'ordre suivant est en $n^{-\frac{1}{3}}$ avec deux termes :

- ▶ le terme en $tn^{-\frac{4}{3}}$ dans la probabilité d'existence d'une arête, multiplié par $n_j \simeq r_j n = \Theta(n)$.
- ▶ le termes dû aux sommets déjà explorés (ν_j), qui est de l'ordre de k , avec $k = \Theta(n^{\frac{2}{3}})$ d'après la renormalisation, multiplié par $\frac{1}{n}$.

La dérive est du bon ordre pour le processus renormalisé (une variation de \tilde{Z} en $\Theta(n^{-\frac{2}{3}})$ pour une variation en temps de $n^{-\frac{2}{3}}$).

La covariance

Lucas MERCIER

Présentation des
modèles

Modèle d'Erdős-Rényi
Graphe inhomogène

Processus
d'exploration des
composantes
connexes

Tailles des
composantes
connexes

Existence d'un
transition de phase

Comportement au
voisinage du seuil
critique

Processus auxiliaires
utilisés
Renormalisation
Dérive
Covariance

Supposons que x_k est de type i . Intéressons-nous au terme principal de la matrice de covariance de Z_{k+1} .

- ▶ $j \neq j'$. Les accroissements respectifs de Z^j et $Z^{j'}$ dépendent d'événements indépendants (la présence d'arêtes distinctes), donc la covariance est nulle.
- ▶ la variance de Z_{k+1}^j . La loi de l'accroissement de Z^j est une binomiale de paramètres $n_j - \nu_j = \Theta(n)$ et $\mathcal{C}_{i,j}(\frac{1}{n} + tn^{-\frac{4}{3}}) = \Theta(\frac{1}{n})$, donc de variance $\Theta(1)$.

Une fois renormalisée, on a, que pour un accroissement de $n^{-\frac{2}{3}}$ en temps, la matrice de covariance est diagonale, et ses termes diagonaux sont $\Theta(n^{-\frac{2}{3}})$.

Cela est bien compatible avec un brownien en dimension ℓ , dont les composantes sont indépendantes.