

Combinatoire sur des structures aléatoires

Charles Bordenave

Slide

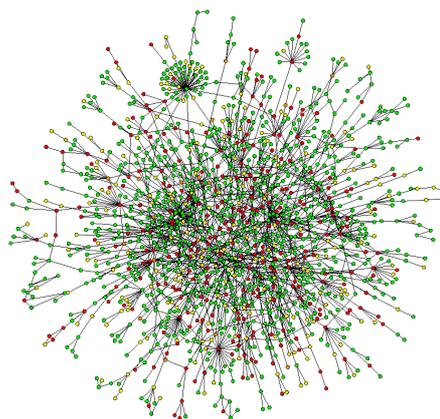
Institut de Mathématiques de Toulouse
CNRS & Université de Toulouse

Travail en collaboration avec Marc Lelarge et Justin Salez (INRIA & ENS).

GRAPHES ET RÉSEAUX COMPLEXES

On s'intéresse à des problèmes de combinatoire sur des graphes $G = (V, E)$ comme celui-ci :

Slide



Réseau d'interaction d'une protéine de levure, Rick Durrett.

GRAPHES ET RÉSEAUX COMPLEXES

⇒ Grands graphes : $|G| \rightarrow \infty$.

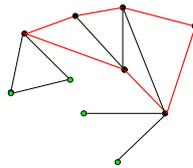
Slide

⇒ Dilués : $\deg(v; G) = O(1)$.

⇒ Géométrie pauvre : localement comme des arbres = peu de cycles courts.

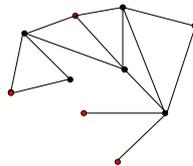
OPTIMISATION COMBINATOIRE

⇒ Grands cycles : Plus grand cycle dans le graphe.



Slide

⇒ Ensemble stable : Plus grand sous-ensemble de sommets non-voisins.



⇒ ...

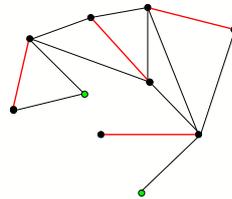
COUPLAGE OU DIMÈRE OU APPARIEMENT D'UN GRAPHE

Un **appariement** de G est un sous-ensemble d'arêtes sans sommet incident commun.

- Un appariement est **parfait** si chaque sommet est incident à une arête de l'appariement.

Slide

- Un appariement est **maximal** si le nombre d'arêtes appariées est maximal.



MESURE DE BOLTZMANN SUR LES APPARIEMENTS (MONOMÈRE-DIMÈRE)

(Zdeborová & Mézard 2006)

Soit $\sigma \subset E$ un appariement de $G = (V, E)$.

L'énergie de σ est définie comme

Slide

$$H_G(\sigma) = |\{\text{sommets non appariés dans } \sigma\}| = |V| - 2|\sigma|.$$

La **mesure de probabilité de Boltzmann** à température $-1/\ln z$, $0 < z < 1$, est

$$\mu_G^z(\sigma) = \frac{z^{H_G(\sigma)}}{P_G(z)}$$

où P_G est la **fonction de partition**.

Pour $z \rightarrow 0$, la mesure converge vers la mesure uniforme sur les appariements maximaux.

FONCTION DE PARTITION ET FONCTION GÉNÉRATRICE

Slide

$$P_G(z) = \sum_{\sigma : \text{appariements}} z^{H_G(\sigma)} = z^{|V|} M_G(z^2)$$

$$M_G(z) = \sum_{\sigma : \text{appariements}} z^{|\sigma|} = \sum_{k=0}^{\infty} |\{\sigma : \text{appariements de taille } k\}| z^k.$$

FONCTIONS D'ÉTAT DE LA PHYSIQUE STATISTIQUE

Slide

- $\varepsilon_G(z) = |V|^{-1} \sum_{\sigma} \mu_G^z(\sigma) H_G(\sigma)$ est la densité d'énergie interne.
- $f_G(z) = |V|^{-1} \ln P_G(z)$ est la densité d'énergie libre.
- $s_G(z) = -|V|^{-1} \sum_{\sigma} \mu_G^z(\sigma) \ln \mu_G^z(\sigma)$ est la densité d'entropie.

FONCTIONS D'ÉTAT DE LA PHYSIQUE STATISTIQUE

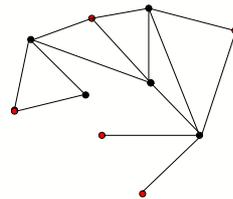
$$f_G(z) = (\ln z) \varepsilon_G(z) + s_G(z).$$

Slide

$$\lim_{z \rightarrow 0} \varepsilon_G(z) = |V|^{-1} |\{\text{sommets non appariés dans un appariement maximal}\}|$$

$$\lim_{z \rightarrow 0} s_G(z) = |V|^{-1} \ln |\{\text{appariements maximaux}\}|.$$

ENSEMBLES STABLES



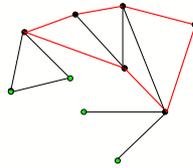
Slide 1

Soit $\sigma \subset V$ un ensemble stable de $G = (V, E)$. On pose

$$H_G(\sigma) = |V| - |\sigma|.$$

\Rightarrow l'énergie est minimisée pour les ensembles stables maximaux.

GRANDS CYCLES



Slide 1 Pour σ un cycle de G , on pose également

$$H_G(\sigma) = |V| - |\sigma|.$$

⇒ **Problème** : la contrainte d'appariement ou d'ensemble stable est **locale**. Pas celle de cycle.

⇒ (*Marinari & Semerjian 2006*) : on regarde plutôt les unions de cycles = sous-graphes de G où chaque sommet est de degré 0 ou 2.

MÉTHODE DE LA CAVITÉ

(*Parisi, Mézard et al., 80's-90's-00's*)

Dans la limite des grands réseaux aléatoires, des outils de **physique statistique** ont été développés pour étudier ces mesures de Boltzmann.

Slide 1

⇒ Importantes conséquences algorithmiques (*belief propagation, survey propagation, ...*).

⇒ Premières justifications rigoureuses de ces méthodes : (*Aldous 2000*), (*Talagrand 2005*)

CONVERGENCE VERS DES STRUCTURES LIMITES

(Aldous et al., 90's-00's)

⇒ On plonge le graphe dans un **espace métrique**.

Slide 1

⇒ On tire un graphe "**au hasard**".

⇒ Dans la limite des grands graphes, on cherche une convergence vers une **structure limite**.

CONVERGENCE VERS DES STRUCTURES LIMITES

⇒ On espère que la convergence dans la topologie choisie implique la convergence des fonctions du graphe qui nous intéressent ...

Slide 1

⇒ On espère que la structure limite est suffisamment simple pour faire des calculs dessus !

⇒ On espère que la structure limite est suffisamment universelle pour éclairer notre compréhension sur des graphes "**réels**".

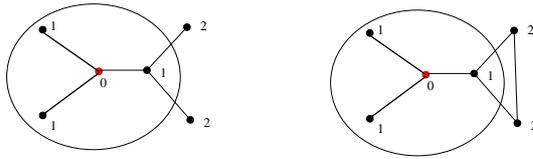
L'ESPACE DES GRAPHES ENRACINÉS

(Benjamini & Schramm 2001, Aldous & Steele 2004)

Grphe enraciné (G, o) = un sommet distingué $o \in V$ + composante connexe de $G = (V, E)$ qui contient o .

$(G, o)_t$ est le sous-graphe de (G, o) des sommets à distance au plus t de la racine o .

Slide 1



$$\text{dist}((G_1, o_1), (G_2, o_2)) = \sum_{t \in \mathbb{N}} 2^{-t} \mathbb{I}\{(G_1, o_1)_t \text{ et } (G_2, o_2)_t \text{ sont isomorphes}\}.$$

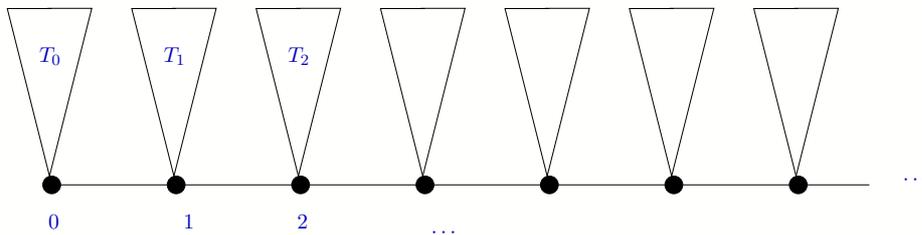
\implies L'espace des classes d'isomorphismes des graphes enracinés localement finis connexe est un espace métrique complet séparable.

ARBRE UNIFORME

Soit G_n un arbre aléatoire à n sommets tiré uniformément.

(Grimmett 1980), G_n converge localement faiblement vers l'arbre squelette :

Slide 1



$(T_i)_{i \in \mathbb{N}}$ arbres de Galton-Watson i.i.d. avec distribution des descendants $\text{Poi}(1)$.

ARBRE UNIFORME

Slide 1 \implies Dans un autre espace et une autre normalisation, G_n converge vers l'arbre brownien (CRT) (Aldous, 1993).

Tout dépend des fonctions du graphe que l'on souhaite étudier !

ENSEMBLES DE GRAPHES

Trois exemples importants de graphes aléatoires sur $\{1, \dots, n\}$,

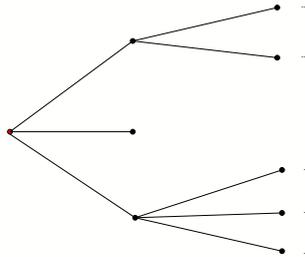
- Graphe d'Erdős-Rényi de paramètre $p = \lambda/n$.
 - Graphe r -régulier uniforme (avec rn pair !).
 - Graphe aléatoire avec suite des degrés prescrits : graphe tiré uniformément sur les graphes avec suite de degrés (d_1, \dots, d_n) .
- On suppose que la mesure empirique des degrés converge,

$$F_n = \frac{1}{n} \sum_{k=1}^n \delta_{d_k} \xrightarrow{n \rightarrow \infty} F.$$

CONVERGENCE LOCALE VERS DES ARBRES ALÉATOIRES

Ces graphes aléatoires convergent localement vers des arbres de Galton-Watson avec distribution des degrés F .

Slide 1



La racine a distribution des descendants F et tous les autres générateurs ont pour distribution des descendants \hat{F} ,

$$\hat{F}(k-1) = \frac{kF(k)}{\sum_{\ell} \ell F(\ell)}.$$

CONVERGENCE LOCALE VERS DES ARBRES ALÉATOIRES

La racine o est tirée uniformément sur $V = \{1, \dots, n\}$: on s'intéresse à la mesure

$$\rho_n = \frac{1}{n} \sum_{v=1}^n \delta_{(G_n, v)}.$$

Slide 2

Alors ρ_n converge presque sûrement

- *graphe r -régulier uniforme* : vers un AGW de distribution des degrés $F = \delta_r$.
- *graphe d'Erdős-Rényi de paramètre λ/n* : vers un AGW de distribution des degrés $F = \text{Poi}_\lambda$.
- *graphe avec mesure empirique des degrés F_n* : vers un AGW de distribution des degrés F .

LA CONVERGENCE LOCALE N'EST PAS GLOBALE

La **structure limite locale** de ces graphes aléatoires est simple et sans cycle.

Slide 2

Le nombre de cycles de longueur ℓ converge faiblement vers Poi_θ où

$$\theta = \frac{1}{2\ell} \left(\frac{\sum_{i \geq 2} i(i-1)F(i)}{\sum_{i \geq 1} iF(i)} \right)^\ell.$$

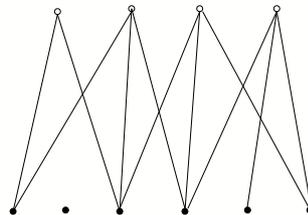
PLUS DE GRAPHES ALÉATOIRES

⇒ De nombreuses variantes et généralisations.

⇒ **Poids** sur les arêtes.

⇒ **Graphes bipartis**.

Slide 2



RETOUR À LA MESURE DE BOLTZMANN DES APPARIEMENTS

$$\mu_G^z(\sigma) = \frac{z^{H_G(\sigma)}}{P_G(z)} = \frac{z^{|V|-2|\sigma|}}{P_G(z)}.$$

Slide 2

$$\varepsilon_G(z) = |V|^{-1} \sum_{\sigma} \mu_G^z(\sigma) H_G(\sigma).$$

$$f_G(z) = |V|^{-1} \ln P_G(z).$$

$$s_G(z) = -|V|^{-1} \sum_{\sigma} \mu_G^z(\sigma) \ln \mu_G^z(\sigma).$$

CONVERGENCE DE LA MESURE DE BOLTZMANN

Théorème 1. Soit G_1, G_2, \dots une suite de graphes convergeant sous enracinement uniforme vers un AGW avec distribution des degrés F . Si $\sum_k kF(k) < \infty$, alors pour $z > 0$

Slide 2

$$\begin{aligned} \lim_n \varepsilon_{G_n}(z) &= \int \frac{z^2}{z^2 + x^2} \nu(dx) \\ \lim_n f_{G_n}(z) &= \frac{1}{2} \int \ln(z^2 + x^2) \nu(dx) \\ \lim_n s_{G_n}(z) &= \frac{1}{2} \int \ln(z^2 + x^2) \nu(dx) - \ln z \int \frac{z^2}{z^2 + x^2} \nu(dx). \end{aligned}$$

où ν est une mesure de probabilité qui ne dépend que de F définie par \dots .

TAILLE ASYMPTOTIQUE DES APPARIEMENTS MAXIMAUX

Théorème 2. Soit G_1, G_2, \dots une suite de graphes convergeant sous enracinement uniforme vers un AGW avec distribution des degrés F . Si $\sum_k kF(k) < \infty$, alors

Slide 2

$$\lim_n \frac{|\{\text{sommets non-appariés dans un appariement maximal}\}|}{|G_n|} = \nu(\{0\}) = \max_{x \in [0,1]} \left\{ -1 + x\varphi'(1-x) + \varphi(1-x) + \varphi\left(1 - \frac{\varphi'(1-x)}{\varphi'(1)}\right) \right\},$$

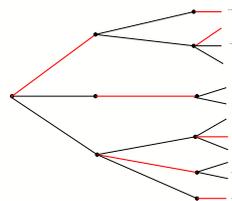
où $\varphi(x) = \sum F(k)x^k$ est la fonction génératrice de F .

⇒ Dans le cas Erdős-Rényi, (Karp & Sipser, 1981) résultat sous une autre forme.

APPARIEMENT DANS UN ARBRE INFINI

La limite ne dépend que de l'arbre limite. Si l'arbre n'a pas de feuille, il existe un appariement parfait sur cet arbre.

Slide 2



Cependant, pour certains choix de la distribution F , p.e.

$$F(d) = \frac{d}{d+1}, \quad F(d^3) = \frac{1}{d+1},$$

la proportion asymptotique de sommets non appariés est **strictement plus petite** que 1...

⇒ But : un appariement maximal **unimodulaire** sur l'arbre aléatoire (Aldous & Lyons 2007).

EXTENSION À D'AUTRES GRAPHES LIMITES

Plus généralement, à un graphe G , on associe son **arbre des chemins** T_G .

On définit le **nombre de branchement** d'un arbre T

Slide 2

$$\text{br}(T) = \sup \left\{ \lambda > 0, \inf_S \sum_{e \in \partial S} \lambda^{-\text{hauteur}(e)} > 0 \right\}.$$

⇒ Si T un AGW de distribution de descendants F alors p.s. $\text{br}(T) = \int x dF$.

CONTINUITÉ DE LA MESURE DE BOLTZMANN

Théorème 3. Soit G_1, G_2, \dots une suite de graphes convergeant sous enracinement uniforme vers un graphe aléatoire G de loi ρ . Si $\text{br}(T_G) < \infty$ p.s. alors

Slide 2

$$\lim_n \varepsilon_{G_n}(z) = \int \frac{z^2}{z^2 + x^2} \nu_\rho(dx)$$
$$\lim_n \frac{|\{\text{sommets non-appariés dans un appariement maximal}\}|}{|G_n|} = \nu_\rho(\{0\}).$$

Si $\text{br}(T_G) < \lambda$ p.s. et si $\ln(1 + \deg(o; G_n))$ est uniformément intégrable alors

$$\lim_n f_{G_n}(z) = \frac{1}{2} \int \ln(z^2 + x^2) \nu_\rho(dx)$$
$$\lim_n s_{G_n}(z) = \frac{1}{2} \int \ln(z^2 + x^2) \nu_\rho(dx) - \ln z \int \frac{z^2}{z^2 + x^2} \nu_\rho(dx).$$

STRATÉGIE DE PREUVE

Il existe une **équation de récursion** satisfaite par $P_G(z)$:

$$P_G(z) = zP_{G-o}(z) + \sum_{v \sim o} P_{G-o-v}(z).$$

Slide 2

$$f_{(G,o)}(z) = \mu_G^z(o \text{ n'est pas appariée}) = \frac{zP_{G-o}(z)}{P_G(z)}.$$

⇒ Permet de définir une mesure de Boltzmann pour $z > 0$ sur tout arbre de branchement fini.

⇒ La formule sur les appariements maximaux vient de l'**équation de récursion en distribution** satisfaite par la v.a. $f_{(T,o)}$ pour $z = 0$ et T un AGW. Il peut exister **plusieurs solutions** à cette équation de récursion.

STRATÉGIE DE PREUVE

Une autre conséquence fondamentale :

Slide 3

Théorème 4 (Heilmann et Lieb, 1972). *Les racines de P_G sont dans $i\mathbb{R}$.*

⇒ Absence de transition de phase pour $z > 0$.

⇒ Par analyticit , il suffit de montrer la convergence pour z suffisamment grand.

NOMBRE D'APPARIEMENTS MAXIMAUX

$$\lim_{z \rightarrow 0} \varepsilon_{G_n}(z) = |G_n|^{-1} |\{\text{sommets non appariés dans un appariement maximal}\}|$$

Slide 3

$$\lim_{z \rightarrow 0} s_{G_n}(z) = |G_n|^{-1} \ln |\{\text{appariements maximaux}\}|.$$

Nos résultats permettent d'échanger les limites $z \downarrow 0$ et $n \rightarrow \infty$ pour l'énergie mais pas pour l'entropie.

ENSEMBLES STABLES

Slide 3

Les racines complexes de la fonction de partition ne sont pas contrôlées. Il existe une transition de phase pour $z > 0$.

⇒ Peu de résultats connus sur la convergence de la mesure de Boltzmann, (*Weitz, 2006*), (*Bayati, Gamarnik & Tetali, 2009*).

ENSEMBLES STABLES

Dans le cas d'un **graphe biparti**, $G = (V, E)$, le théorème de König (Steinitz ?) implique

$$|\{\text{sommets dans un ensemble stable maximal}\}| \\ + |\{\text{arêtes dans un appariement maximal}\}| = |V|.$$

Slide 3

\Rightarrow Soit G_1, G_2, \dots une suite de graphes **bipartis** convergeant sous enracinement uniforme vers un AGW avec distribution des degrés F , alors

$$\lim_n \frac{|\{\text{sommets dans un ensemble stable maximal}\}|}{|G_n|} \\ = \frac{1}{2} \max_{x \in [0,1]} \left\{ x\varphi'(1-x) + \varphi(1-x) + \varphi\left(1 - \frac{\varphi'(1-x)}{\varphi'(1)}\right) \right\}.$$

\Rightarrow **Faux** pour une suite de graphes non-bipartis !

GRANDS CYCLES

Les racines complexes de la fonction de partition ne sont pas contrôlées.

Slide 3

\Rightarrow on peut modifier la fonction de partition de façon à avoir un analogue d'Heilmann-Lieb (Ruelle 1999, Wagner 2009).

(Semerjian) Conjecture sur la **taille asymptotique du plus grand cycle** d'un graphe aléatoire avec mesure empirique des degrés $F_n \Rightarrow F$.