

# Le problème d'assignation aléatoire euclidienne: état de l'art et quelques problèmes ouverts en dimension $d \leq 2$

Groupe de travail Probabilités LAMA-UPEC  
**Matteo D'Achille**

LAMA, UParis-Est Créteil et LIPN, USorbonne Paris Nord



Collaborateurs: Dario Benedetto, Emanuele Caglioti, Sergio Caracciolo, Vittorio Erba, Gabriele Sicuro et Andrea Sportiello

mardi 10 novembre 2020

# Plan de l'exposé

- ▶ Introduction
- ▶ Le problème d'assignation aléatoire euclidienne (ou ERAP)
- ▶ Diagramme de phase: état de l'art
- ▶ ERAP en dimension de Hausdorff  $d_H \in (1, 2)$
- ▶ Conclusions

Papiers pour plus de détails:

- P ⑤ **MD**, A. Sportiello, *Euclidean Random Assignment Problems at non-integer Hausdorff dimension  $d_H \in (1, 2)$* , en préparation
- P ④ D. Benedetto, E. Caglioti, S. Caracciolo, **MD**, G. Sicuro, A. Sportiello, [2008.01462](#)
- P ③ S. Caracciolo, **MD**, V. Erba, A. Sportiello, J. Phys. A 53, [1904.10867](#)
- P ② S. Caracciolo, **MD**, G. Sicuro, J. Stat. Phys. 174, [1803.04723](#)
- P ① S. Caracciolo, **MD**, G. Sicuro, Phys. Rev. E 96, [1707.05541](#)

# Introduction

Les problèmes d'optimisation combinatoire sont omniprésents dans la vie quotidienne et consistent à trouver l'extrema d'une fonction d'intérêt sur des larges ensembles discrets

Le **problème d'assignation** en est un exemple majeur et trouve application dans une grande variété de situations, par exemple :

- ▶ Suivi des particules d'objets identiques qui se déplacent en régime diffusive (gouttes d'eau dans un nuage, oiseaux dans un troupeau...)
- ▶ Conception des réseaux cellulaires, compte tenu de la géographie + la densité spatiale des utilisateurs ( e.g. à partir des données historiques )

L'importance de ce problème a été discutée dans divers contextes interdisciplinaires :

1. Economie (Koopmans–Beckmann 1957)
2. Théorie de jeux: stratégie mixte optimale dans un jeu à deux joueurs (Von Neumann 1953, 1954)

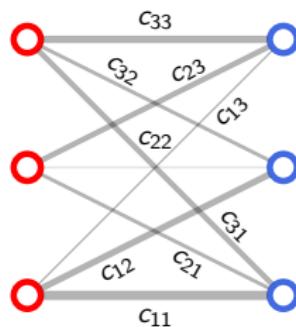
## Le problème d'assignation (AP)

On a une matrice  $n \times n c$ , on cherche une bijection  $\pi$  de sorte que

$$\mathcal{H}(\pi, c) = \sum_i c_{i\pi(i)}$$

soit minimal. Appelons-le  $\pi_{\text{opt}}$  et soit  $\mathcal{H}(\pi_{\text{opt}}; c) = \mathcal{H}_{\text{opt}}(c)$ .

Ex.  $n = 3$ :



$$c = \begin{pmatrix} 5 & 3.5 & 1 \\ 2 & 1.2 & 3 \\ 3 & 2 & 4 \end{pmatrix}$$

AP est **P-complete**: tout autre problème dans cette classe ( e.g. inversion d'une matrice ou déterminer si un nombre naturel donné est premier ) peut être formulé comme un AP.  $c$  étant connue,  $\pi_{\text{opt}} = \pi_{\text{opt}}(c)$  est trouvée en temp  $\mathcal{O}(n^3)$  (Munkres 1957)

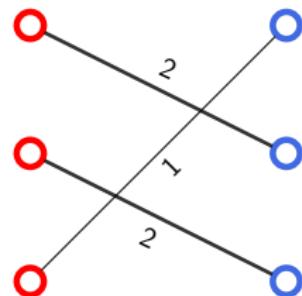
## Le problème d'assignation (AP)

On a une matrice  $n \times n c$ , on cherche une bijection  $\pi$  de sorte que

$$\mathcal{H}(\pi, c) = \sum_i c_{i\pi(i)}$$

soit minimal. Appelons-le  $\pi_{\text{opt}}$  et soit  $\mathcal{H}(\pi_{\text{opt}}; c) = \mathcal{H}_{\text{opt}}(c)$ .

Ex.  $n = 3$ :



$$c = \begin{pmatrix} 5 & 3.5 & 1 \\ 2 & 1.2 & 3 \\ 3 & 2 & 4 \end{pmatrix}$$

$$\mathcal{H}_{\text{opt}}(c) = 5$$

AP est **P-complete**: tout autre problème dans cette classe ( e.g. inversion d'une matrice ou déterminer si un nombre naturel donné est premier ) peut être formulé comme un AP.  $c$  étant connue,  $\pi_{\text{opt}} = \pi_{\text{opt}}(c)$  est trouvée en temp  $\mathcal{O}(n^3)$  (Munkres 1957)

## AP: un vieux problème



König  
1916



Egérvary  
1931



Munkres  
1957

von Neumann  
1953



Kuhn  
1955

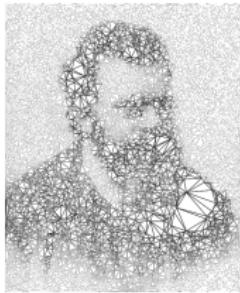


*“De investigando or-  
dine systematis ae-  
quationum...”*  
Jacobi 1860  
See Ollivier 2009



AP est bien compris et aujourd'hui résolu efficacement même sur les smartphones... pourquoi l'étudier ?

# Classe de problèmes d'OC $\equiv$ système désordonné



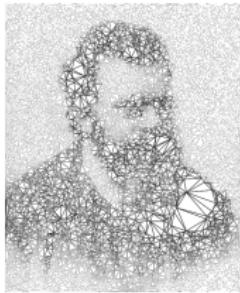
le "Boltzmann désordonné"



Physique: limite de température zéro ( état fondamental ) d'un seul système physique désordonné (Kirkpatrick *et al.* 1983).

*Mécanique Statistique*

# Classe de problèmes d'OC $\equiv$ système désordonné



le "Boltzmann désordonné"

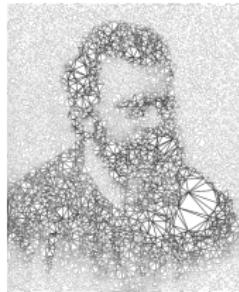


Physique: limite de température zéro ( état fondamental ) d'un seul système physique désordonné (Kirkpatrick *et al.* 1983).

*Optimisation Combinatoire*

*Mécanique Statistique*

# Classe de problèmes d'OC $\equiv$ système désordonné



le "Boltzmann désordonné"



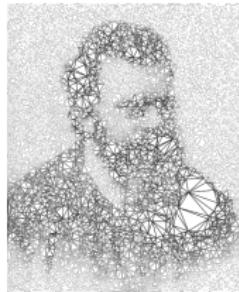
Physique: limite de température zéro ( état fondamental ) d'un seul système physique désordonné (Kirkpatrick *et al.* 1983).

*Optimisation Combinatoire*

*Mécanique Statistique*

- ▶ Solution optimale

# Classe de problèmes d'OC $\equiv$ système désordonné



le "Boltzmann désordonné"



Physique: limite de température zéro ( état fondamental ) d'un seul système physique désordonné (Kirkpatrick *et al.* 1983).

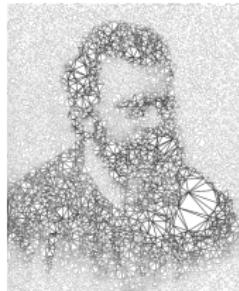
*Optimisation Combinatoire*

*Mécanique Statistique*

- ▶ Solution optimale

- ▶ État fondamental

# Classe de problèmes d'OC $\equiv$ système désordonné



le "Boltzmann désordonné"



Physique: limite de température zéro ( état fondamental ) d'un seul système physique désordonné (Kirkpatrick *et al.* 1983).

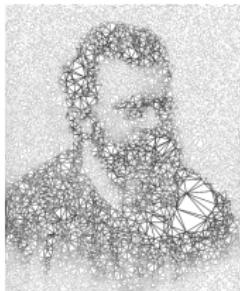
## *Optimisation Combinatoire*

- ▶ Solution optimale
- ▶ Propriétés typiques de la solution optimale

## *Mécanique Statistique*

- ▶ État fondamental

# Classe de problèmes d'OC $\equiv$ système désordonné



le "Boltzmann désordonné"



Physique: limite de température zéro ( état fondamental ) d'un seul système physique désordonné (Kirkpatrick *et al.* 1983).

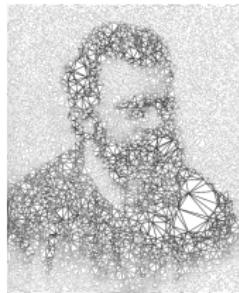
## *Optimisation Combinatoire*

- ▶ Solution optimale
- ▶ Propriétés typiques de la solution optimale

## *Mécanique Statistique*

- ▶ État fondamental
- ▶ Moyenne "quenched" sur le désordre

# Classe de problèmes d'OC $\equiv$ système désordonné



le “Boltzmann désordonné”



Physique: limite de température zéro ( état fondamental ) d'un seul système physique désordonné (Kirkpatrick *et al.* 1983).

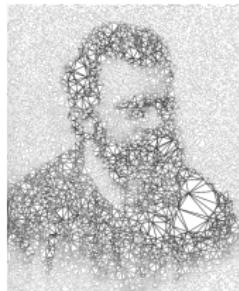
## *Optimisation Combinatoire*

- ▶ Solution optimale
- ▶ Propriétés typiques de la solution optimale
- ▶ Contraintes géométriques et topologiques

## *Mécanique Statistique*

- ▶ État fondamental
- ▶ Moyenne “quenched” sur le désordre

# Classe de problèmes d'OC $\equiv$ système désordonné



le "Boltzmann désordonné"



Physique: limite de température zéro ( état fondamental ) d'un seul système physique désordonné (Kirkpatrick *et al.* 1983).

## *Optimisation Combinatoire*

- ▶ Solution optimale
- ▶ Propriétés typiques de la solution optimale
- ▶ Contraintes géométriques et topologiques

## *Mécanique Statistique*

- ▶ État fondamental
- ▶ Moyenne "quenched" sur le désordre
- ▶ Frustration

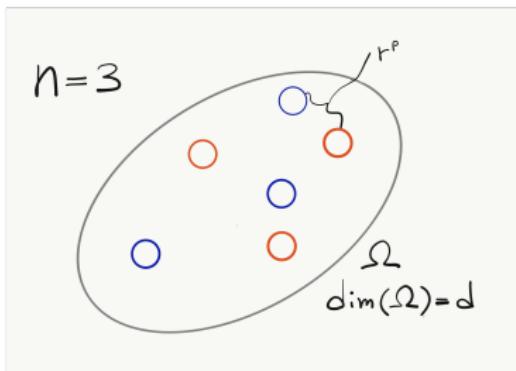
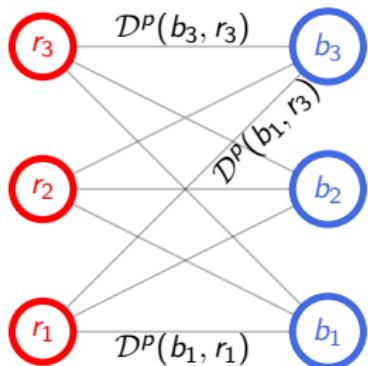
## Le problème d'assignation aleatoire (RAP)

Dans un problème d'assignation aléatoire la matrice des coûts  $c$  vient d'un ensemble statistique. On s'intéresse aux propriétés statistiques de  $\mathcal{H}_{\text{opt}}$  ( asymptotique de la valeur d'espérance mathématique pour grand  $n$ , loi de  $\mathcal{H}_{\text{opt}}$  ) dépendant du choix de l'ensemble.

Si les  $c_{ij}$  sont des RV i.i.d., on parle de "problème d'assignation aléatoire ( ou RAP )". Le RAP est un problème bien connu surtout en physique statistique ( Mézard–Parisi, Orland ), et dans les probabilités ( Aldous, Wästlund, Salez ).

Insatisfaisant du point de vue physique: il s'agit d'un modele de champ moyen, i.e. pas de frustration!

# Problème d'assignation aléatoire euclidienne: motivation



Un ensemble statistique différent: deux échantillons aléatoires  $\mathcal{B}$  et  $\mathcal{R}$  ( i.e. sommets du graphe biparti complet ), dans une variété  $\Omega$  de dimension  $d$ . Cout d'une couple bleu-rouge:  $D^P(b_i, r_j)$ , où  $D$  est la distance euclidienne

Principe d'universalité: tout autre détail de la fonction de coût serait éliminé dans la grande limite de  $n$

Motivation de la physique des systèmes critiques: invariance sous translations, rotations et transformation d'échelle

## Problème d'assignation aléatoire euclidienne: motivation

Deux connections importants:

- ▶ **Transport optimal** - C'est le problème de Monge-Kantorovitch entre les mesures empiriques  $\rho_B$  et  $\rho_R$  associées aux points bleues et rouges:

$$\mathcal{H}_{\text{opt}} = nW_p^p(\rho_B, \rho_R)$$

où  $W_p$  est la distance de  $p$ -Wasserstein (Ledoux, Talagrand, Brezis, Ambrosio et al.)

- ▶ **Verres de spin** - C'est un modèle-jouet de verre de spin en dimension finie. Il modélise les propriétés essentielles (frustration, en raison de l'inégalité triangulaire) tout en restant plus facile à traiter, par rapport au verre de spin de Edwards–Anderson (Mézard–Parisi)

## Problème d'assignation aléatoire euclidienne: définition

Nous nous intéressons à la valeur moyenne de l'énergie et en particulier à son expansion asymptotique :

$$E_{p,d}(n) = \mathbb{E}_{n,p,d} [\mathcal{H}_{\text{opt}}] \stackrel{?}{=} K_{p,d} n^{\gamma_{p,d}} (\ln n)^{\gamma'_{p,d}} (1 + o(1))$$

pour grand  $n$ , en fonction de  $d, p$  et du choix du désordre ( loi de prob. associée aux points bleus et rouges ).

On veut étudier  $(\gamma_{p,d}, \gamma'_{p,d})$  et on s'attende qu'ils soient largement indépendants des détails microscopiques du problème ( qui affectent la constante  $K_{p,d}$  ) → diagramme de phase

Remarque: les cas non uniformes sont plus subtils

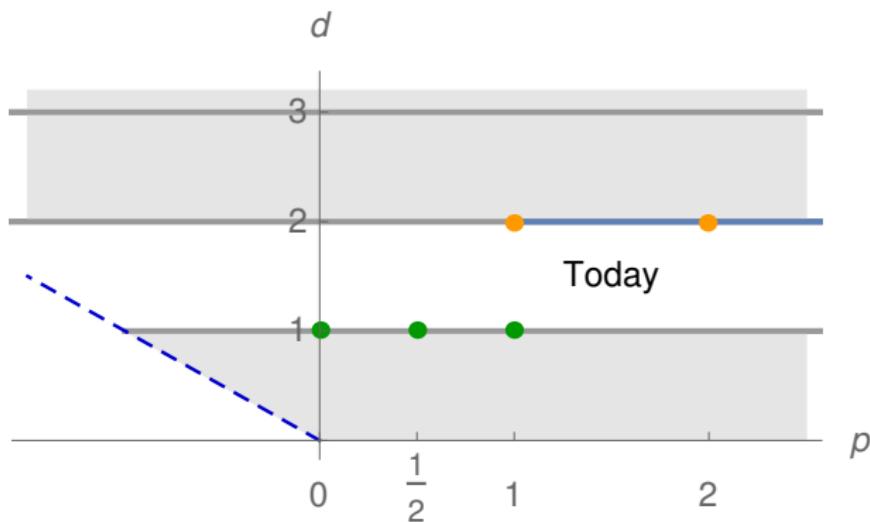
Example:  $d = 1, p = 2$ , et loi gaussienne, on a

$$E_{2,1}(n) \sim 2 \ln \ln n \quad (\gamma_{2,1} = \gamma'_{2,1} = 0)$$

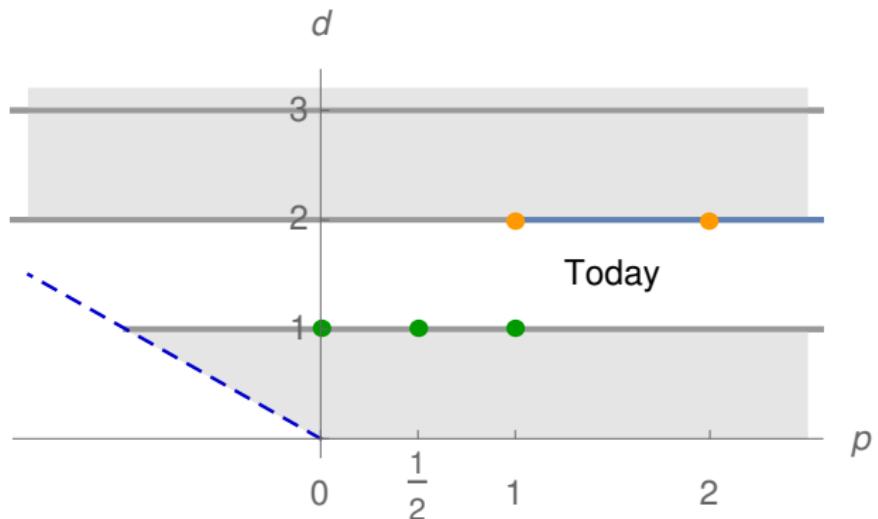
(P ②, Bobkov–Ledoux 2019 + Berthet–Fort 2020)

## Diagramme de phase: état de l'art

Je résume ici les faits connus sur le diagramme de phase dans le cas de la mesure uniforme.



## Diagramme de phase: état de l'art

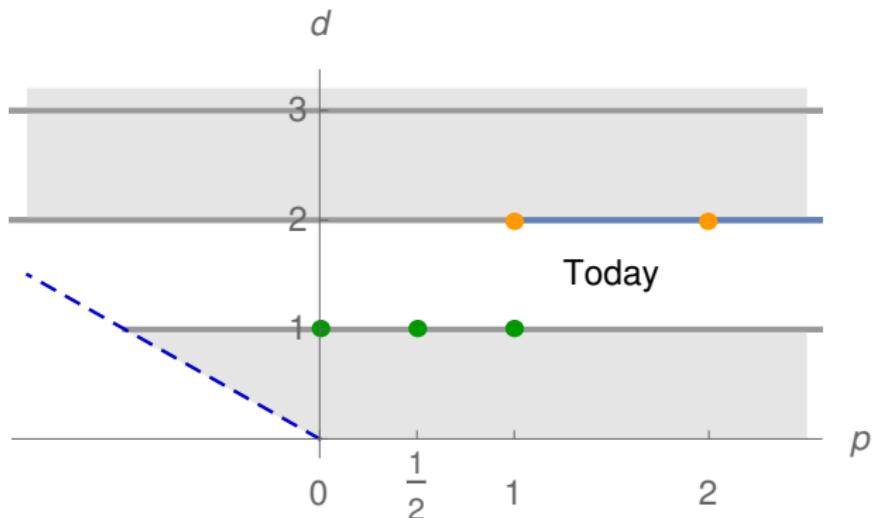


$d \geq 3$  “simple” ou sans “aucune géométrie” (Mézard–Parisi 1988)

$$E_{p,d}(n) \sim E_{\text{LB}}(n) = cn^{1-p/d}$$

Pour  $\exists! K_{p,d}$  voir e.g. Barthe–Bordenave 2013, Goldman–Trevisan 2020

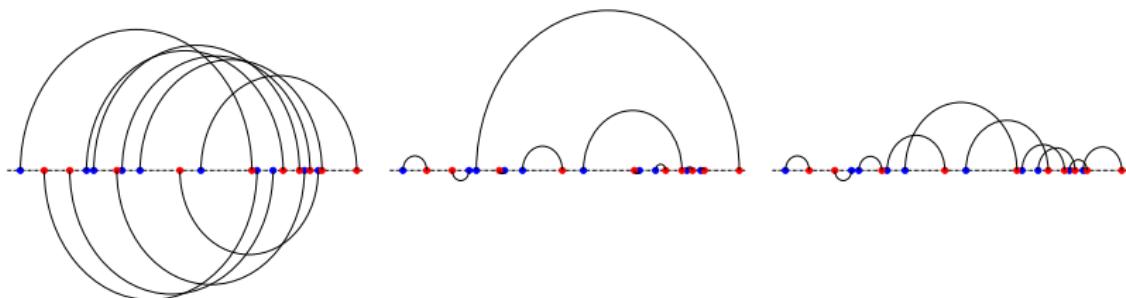
## Diagramme de phase: état de l'art



$d = 1$  “simple” ( mais riche ! ) en raison de la structure mathématique de la solution

## $d = 1$ : propriétés qualitatives de la solution

$p = 0$  et  $p = 1$  séparent trois régimes qualitativement différents:



Une fois bleus et rouges triés dans l'ordre naturel, optimalité implique que la permutation optimale est:

$p < 0$  cyclique:  $\pi_{\text{opt}}(i) = i + k \pmod n$  (P ①)

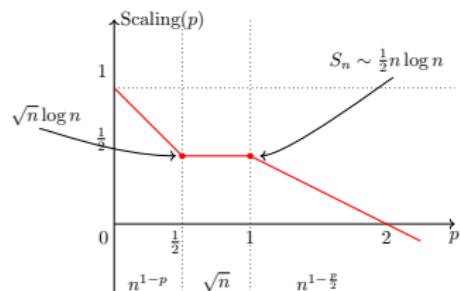
$p \in (0, 1)$  non-crossing: pour  $AB, CD$  des intervalles dans  $\pi_{\text{opt}}$ , soit  $AB \subset CD$  ou  $AB \cap CD = \emptyset$  (McCann 1999)

$p > 1$  solution triée:  $\pi_{\text{opt}}$  est la permutation identique ( convexité + stricte croissance );

Remarque: ces propriétés combinatoires sont vraies  $\forall$  instance

## Tranche du diagramme de phase à $d = 1$

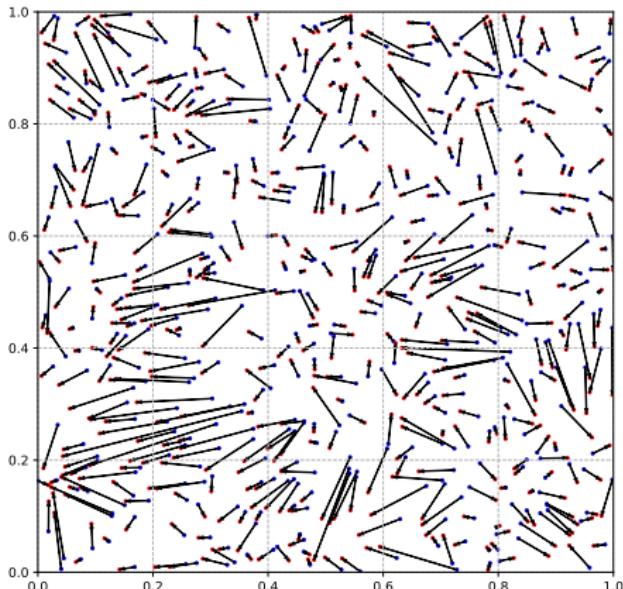
En utilisant ces propriétés (P ① + P ③) on a montré



$$(\gamma_{p,d}, \gamma'_{p,d}) = \begin{cases} (1 - \frac{p}{2}, 0) & p \geq 1 \\ (\frac{1}{2}, 0) & p \in (\frac{1}{2}, 1) \\ (\frac{1}{2}, 1) & p = \frac{1}{2} \\ (1 - p, 0) & p \in (0, \frac{1}{2}) \end{cases}$$

et  $K_{p,1}|_{p \geq 1} = \frac{\Gamma(1+p/2)}{p+1}$ , mais  $K_{p,1}|_{p \in (0,1)}$  est inconnue.

## $d = 2$ : un problème “ancien” et stimulante



$$(\gamma_{p,d}, \gamma'_{p,d}) = (\gamma_{\text{LB}}, \frac{p}{2}) \text{ à } p \geq 1 \text{ (Ajtai-Komlós-Tusnády 1984)}$$

Typiquement, l'assignation ne se réalise pas dans le voisinage euclidien d'un point, mais plutôt avec le  $\sim (\ln n)$ -ième voisin

## Quelques remarques sur le cas $(p, d) = (2, 2)$

$$E_\Omega(n) \sim K_{p,2} n^{1-\frac{p}{2}} (\ln n)^{\frac{p}{2}}$$

La constante  $K_{p,2}$  était inconnue pour toute valeur  $p$  et toute variété riemannienne bi-dimensionnelle  $\Omega$

Caracciolo–Lucibello–Parisi–Sicuro ont montré par un calcul de QFT qu'à  $p = 2$ , dans un  $\Omega$  t.q.  $|\Omega| = 1$

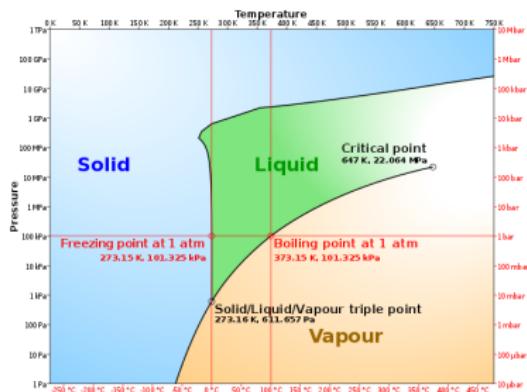
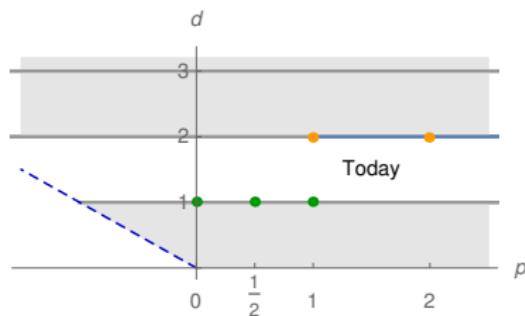
$$K_{2,2} = \frac{1}{2\pi}$$

Preuve de  $K_{2,2}$  en (Ambrosio–Stra–Trevisan 2019)

On peut calculer analytiquement la différence asymptotique  $E_\Omega(n) - E_{\Omega'}(n)$  entre deux variétés ( voir P ④ ), mais je n'aborderai pas ce sujet aujourd'hui ( voir aussi cet recent exposé par A. Sportiello )

# ERAP en dimension de Hausdorff $d_H \in (1, 2)$ P ⑤

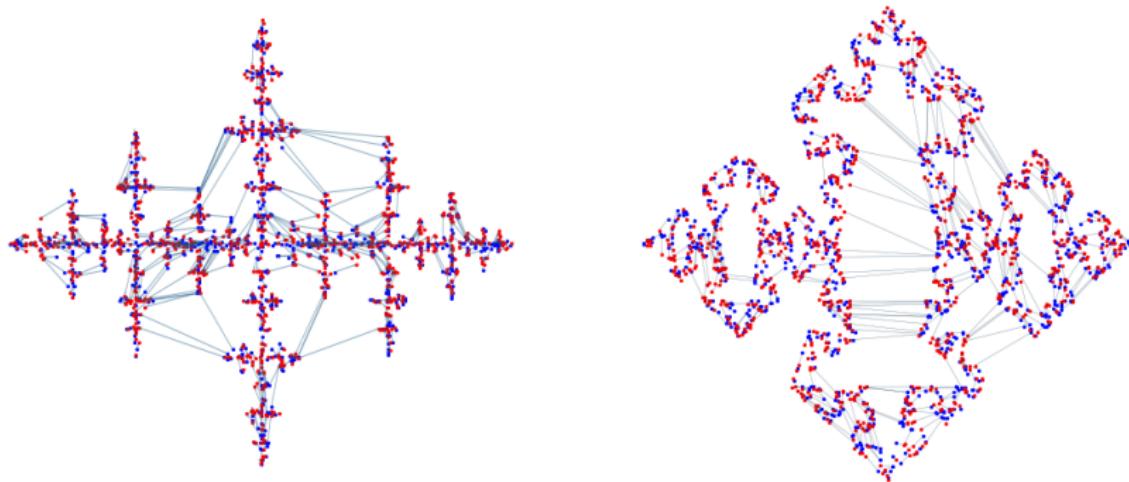
Motivation: on a observé plusieurs indications de “points critiques” dans le diagramme de phase



Dans la physique des **phénomènes critiques**, lorsque des points critiques sont observés, il est naturel de se demander s'ils appartiennent à une ou plusieurs **lignes critiques**. Les propriétés des observables statistiques peuvent changer de manière non-analytique à travers une ligne critique.

Comment les points critiques à  $d = 1$  “entrent-ils” dans le diagramme de phase ?

## ERAP en dimension de Hausdorff $d_H \in (1, 2)$ P ⑤



IDÉE: on considère des points **bleus** et **rouges** iid sur un ensemble fractal de dimension de Hausdorff  $d_H \in (1, 2)$  immergé dans le plan et on calcule l'énergie en utilisant la distance euclidienne bi-dimensionnelle.

On considere deux ensembles fractals afin de tester l'universalité

## Fractal de Peano

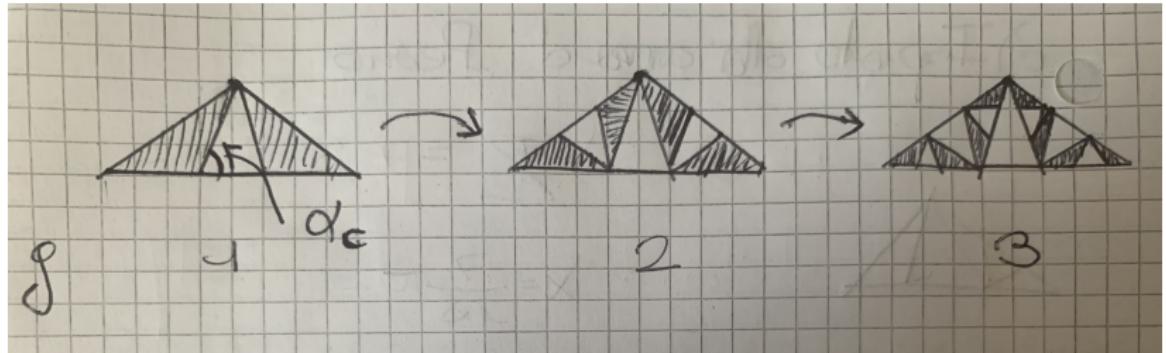
Pour  $P = (\xi, z) \in \mathbb{C}^2$ ,  $\alpha_P = (2^{d_H-1} - 1)^{\frac{1}{d_H}}$  et

$$\begin{aligned}\mathbf{P} &= \frac{1}{2(1 + \alpha_P^{d_H})} \left(1, \alpha_P^{d_H}, \alpha_P^{d_H}, 1\right), \\ \mathbf{\Lambda} &= \frac{1}{2} \left(1, \alpha_P^{d_H}, \alpha_P^{d_H}, 1\right), \\ \mathbf{\Phi} &= (1, i, -i, 1), \\ \mathbf{V} &= \frac{1}{2} (0, 1, 1 + i\alpha_P, 1)\end{aligned}\tag{1}$$

et pour  $u \in (0, 1)$ ,  $j = \lfloor R^{-1}(u) \rfloor$ ,  $j \in \{1, 2, 3, 4\}$  on itère la fonction aléatoire

$$f(P) := (\lambda_j \phi_j \xi, z + v_j \xi)$$

# Fractal de Cesàro



Quand  $\alpha_C = \frac{\pi}{3}$  on retrouve la courbe de Koch. On montre que

$$d_H = \frac{2}{1 + \frac{\ln(1+\cos \alpha_C)}{\ln 2}} \in (1, 2) \quad \text{pour} \quad \alpha_C \in \left(0, \frac{\pi}{2}\right)$$

## Peano vs Cesàro à la même dimension de Hausdorff

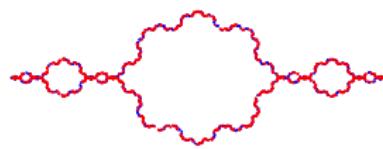
$g = 10$  générations pour éviter les problèmes dus à la précision de virgule flottante

$$d_H = 1.1, \quad n = 2^{12}$$

Peano



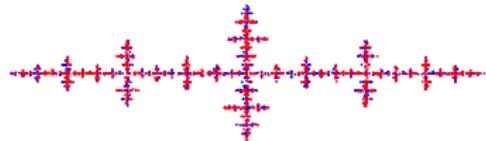
Cesàro



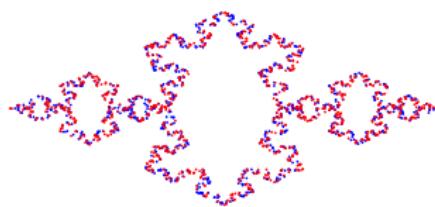
## Peano vs Cesàro à la même dimension de Hausdorff

$$d_H = 1.4, \quad n = 2^{12}$$

Peano



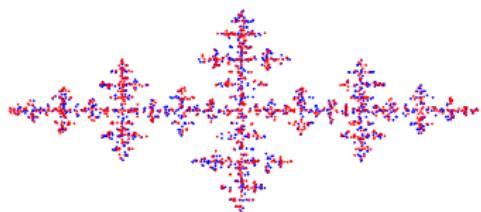
Cesàro



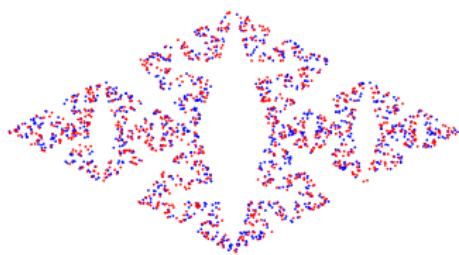
## Peano vs Cesàro à la même dimension de Hausdorff

$$d_H = 1.6, \quad n = 2^{12}$$

Peano



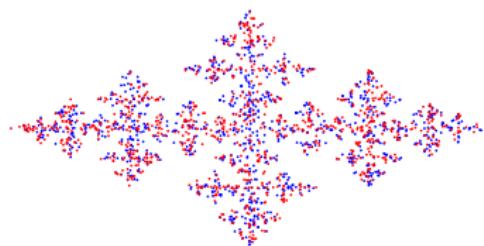
Cesàro



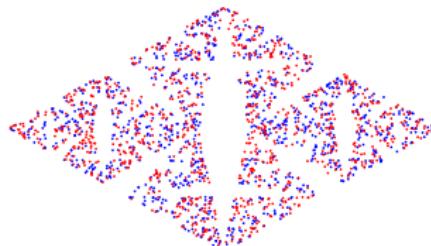
# Peano vs Cesàro à la même dimension de Hausdorff

$$d_H = 1.7, \quad n = 2^{12}$$

Peano



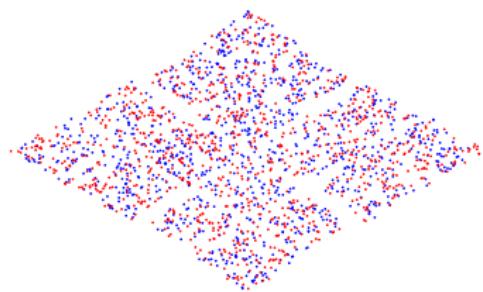
Cesàro



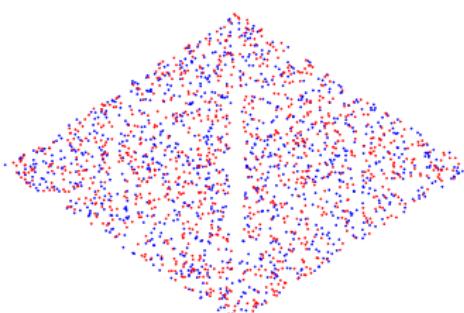
## Peano vs Cesàro à la même dimension de Hausdorff

$$d_H = 1.9, \quad n = 2^{12}$$

Peano



Cesàro



# L'observable

Hypothèse: nous négligeons  $\gamma'_{p,d_H}$  et supposons<sup>1</sup>

$$E_{p,d_H}^{\text{P/C}}(n) \sim c_{(p,d_H)}^{\text{P/C}} n^{\gamma_{(p,d_H)}^{\text{P/C}}}$$

Est-ce que les exposants  $\gamma_{(p,d_H)}^{\text{P/C}}$  dépendent-ils de l'ensemble fractal considéré ?

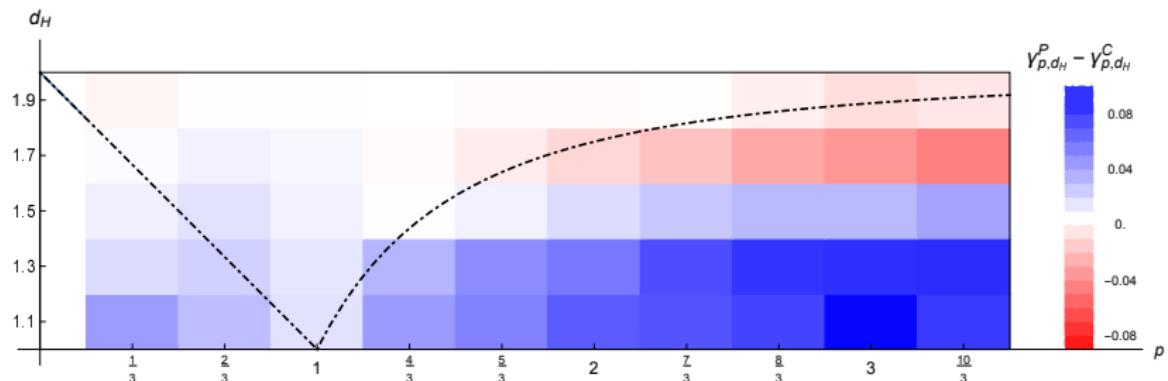
Démarche expérimentale: nous avons estimé numériquement  $\gamma_{(p,d_H)}^{\text{P/C}}$  dans la région du plan  $(p, d_H)$  :  $p \in (.33, 1.33)$  et  $d_H \in (1.1, 1.9)$ , pour un total de 50 points  $(p, d_H)$ . On a estimé  $\gamma_{(p,d_H)}^{\text{P/C}}$  par deux méthodes statistiques différentes qui ont donné des résultats cohérents

---

<sup>1</sup>L'estimation de  $\gamma'_{p,d_H}$  reste hors de portée de notre protocole numérique.

## Resultats [P ⑤]

Résumé: pour Peano et Cesàro on a  $\gamma_{(p,d_H)}^{\text{P/C}} > 1 - \frac{p}{d_H}$



Signature d'universalité: dans toute une sous-région du domaine considéré "comportant" la ligne  $d_H = 2$

$\gamma_{(p,d_H)}^{\text{P}} - \gamma_{(p,d_H)}^{\text{C}}$  est compatible avec zero entre l'erreur statistique

## Conclusions et quelques perspectives

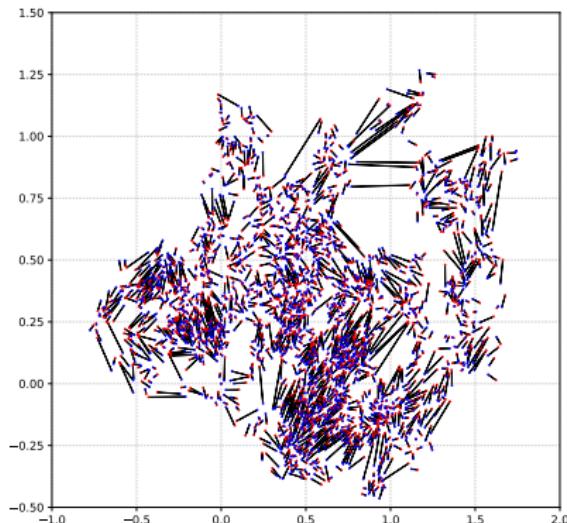
Motivé par une analogie avec la physique des transitions de phase, nous avons introduit et étudié une notion d'ERAP en “dimension intermédiaire”  $d_H \in (1, 2)$ .

De façon surprenante, l'exposant  $\gamma_{p,d_H}$  semble être universel d'après un étude experimental des fractals de Peano et Cesàro.

- ▶ Comment établir une borne inférieure pour  $\gamma_{p,d_H}$  ?
- ▶ Y a-t-il une ligne critique ?
- ▶ Des autres ensembles fractals, pour explorer aussi  $d_H \in (2, 3)$

## Une autre perspective: ERAP sur la boucle brownienne

Quand  $(p, d) = (1, 2)$ , on peut montrer que la solution optimal est non-crossing: les arêtes ne se croisent pas ( inégalité triangulaire ) Qualitativement, on observe de longs "peignes" des arêtes parallèles qui assignent des régions éloignées. Exemple, sur la boucle brownienne, est-ce que  $(\gamma, \gamma')$  sont déterminés par la masse ou par le comportement anormal autour des temps de coupure



MERCI DE VOTRE ATTENTION !

