

Back and forth between the beta distribution and edge stochastic domination in ERAPs

DYOGENE Seminar

INRIA de Paris, ENS and CNRS

Monday, 20 June 2022, 14h00-15h00 Paris time

Matteo D'Achille



UNIVERSITÉ
PARIS-EST CRÉTEIL
VAL DE MARNE



Based on two recent works in progress in collaboration with:

- Andrea Sportiello (CNRS, université Sorbonne Paris Nord)
- Yuqi Liu (M2 internship, université Paris-Est Créteil)

and on several papers in collaboration with:

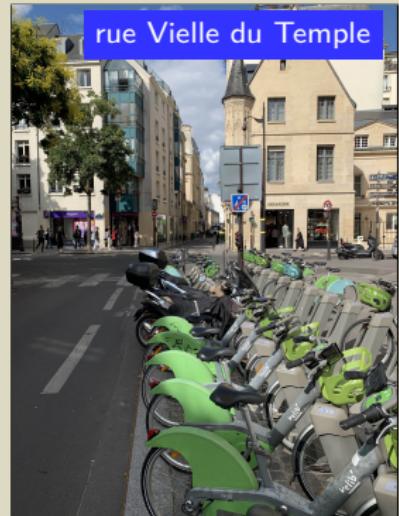
- Dario Benedetto (Rome La Sapienza)
- Emanuele Caglioti (Rome La Sapienza)
- Sergio Caracciolo (Milan University, INFN)
- Vittorio Erba (Lausanne EPFL)
- Gabriele Sicuro (London King's College)
- Andrea Sportiello (CNRS and Université Sorbonne Paris Nord)

References

Main references: (here and afterwards all paper references are clickable) preprint at

- J. Stat. Phys. **183**, 34, 2021 2008.01462
- PhD Thesis, Université Paris-Saclay, 2020 tel:03098672
- J. Phys. A Math. Theor. **53**, 6, 2020 1904.10867
- J. Stat. Phys. **174**, 4, 2019 1803.04723
- Phys. Rev. E **96**, 4, 2017 1707.05541

The assignment problem: A current trouble



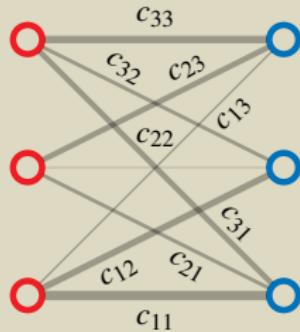
The assignment problem: Some definitions

Definition (assignment problem). Consider a $n \times n$ real (cost) matrix c . For any n -permutation $\pi \in \mathcal{S}_n$, define the total cost

$$E(\pi) = \sum_{i=1}^n c_{i\pi(i)}.$$

Problem: Find $\pi_{\text{opt}} := \arg \min_{\pi \in \mathcal{S}_n} E(\pi)$.

Example at $n = 3$:



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

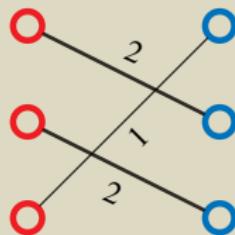
The assignment problem: Some definitions

Definition (assignment problem). Consider a $n \times n$ real (cost) matrix c . For any n -permutation $\pi \in \mathcal{S}_n$, define the total cost

$$E(\pi) = \sum_{i=1}^n c_{i\pi(i)}. \quad \text{Problem: Find } \pi_{\text{opt}} := \arg \min_{\pi \in \mathcal{S}_n} E(\pi).$$

Example at $n = 3$:

$$\pi_{\text{opt}} = (3 \ 1 \ 2), E_{\text{opt}} := E(\pi_{\text{opt}}) = 5$$



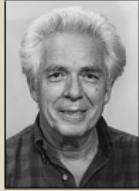
$$c = \begin{pmatrix} 5 & 3.5 & \textcircled{1} \\ \textcircled{2} & 1.2 & 3 \\ 3 & \textcircled{2} & 4 \end{pmatrix}$$

- Optimization of a linear function over the **convex Birkhoff polytope**;
- **P-complete**, $\mathcal{O}(n^3)$ complexity (Munkres 1957);
- Equivalent to a two player zero-sum game (von Neumann 1953, 1954).

Assignment problem: Historical remarks



von
Neumann
1953



Kuhn
1955



König
1916



Egérvary
1931

Assignment problem: Historical remarks



von
Neumann
1953



Kuhn
1955

König
1916



Egérvary
1931



Canon simplicissimus.							
	I	II	III	IV	V	VI	VII
I	25*	21	20	18	20	18	25
II	21	22*	21	21	13	21	22
III	16	19	23*	22	17	14	16
IV	21	12	18	27*	18	14	24
V	25	22	22	27	31*	16	31
VI	10	18	23	21	19	23*	21
VII	5	14	10	27	31	20	40*

“De investigando ordine systematis aequationum differentialium vulgarium cuiuscunque”

See also (Ollivier 2009)

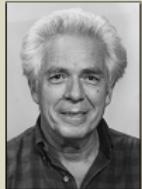
Jacobi
1860



Assignment problem: Historical remarks



von
Neumann
1953

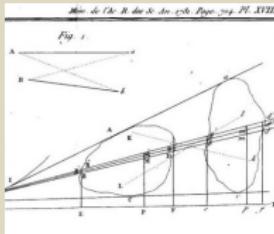


Kuhn
1955

König
1916



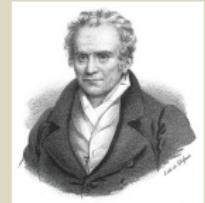
Egérvary
1931



“Le mémoire sur les déblais et les remblais”

See this introduction at images.math.cnrs.fr (Ghys 2012)

Monge
1784



Random Assignment Problems: The case of independent weights

Here c is a random matrix, E_{opt} a **(non-trivial) random variable**.

- Pioneered in Physics in the 80s by Mézard–Parisi and Orland;
- Entered Probability Theory mostly via Aldous in the 90s.

Basic result: if $(c_{ij})_{i,j=1}^n$ are i.i.d. r.v.s of pdf $\rho(l) = l^r + o(l^r)$, then

$$\mathbb{E}[E_{\text{opt}}]_n \underset{n \rightarrow \infty}{\sim} c_r n^{1 - \frac{1}{r+1}}.$$

Only “infinitely short” edges are relevant for large n . r plays the role of a “**universal exponent**”.

Nice fact: At $r = 0$ (i.e. ρ is e.g. **uniform** or $\text{Exp}(\lambda)$ distribution),

$$c_0 = \zeta(2) = \frac{\pi^2}{6}.$$

The Parisi conjecture

If $c_{ij} \sim \text{Exp}(1)$, Parisi conjectured (1998):

$$\mathbb{E}[E_{\min}]_n = \sum_{k=1}^n \frac{1}{k^2} = \frac{\pi^2}{6} - \frac{1}{n} + o\left(\frac{1}{n}\right).$$

- Extension to rectangular cost matrices (Coppersmith-Sorkin 1998);
- Proof of $\zeta(2)$ limit (among other things) (Aldous 2001);
- Proof of Parisi conjecture (Nair–Prabhakar–Sharma 2001);
- Extension to k -partite graphs (which is **NP-hard** for $k \geq 3$) (Martin–Mézard–Rivoire 2004,2005);
- Prof of $\exists!$ of Mézard–Parisi order parameter $\forall r \geq 0$ (Wästlund 2012, Larsson 2014, Salez 2015).

NOT discussed today...

Euclidean Random Assignment Problems (ERAPs)

Let $\mathcal{B} = (B_1, \dots, B_n)$ be blue points and $\mathcal{R} = (R_1, \dots, R_n)$ be red ones: n -samples of i.i.d. r.v. with pdf $v_{\mathcal{B}(\mathcal{R})} : \Omega \rightarrow \mathbb{R}$ (**disorder**). Let (Ω, \mathcal{D}) be a metric space (mostly an **Euclidean** space with \mathcal{D} **Euclidean** distance). For $p \in \mathbb{R}$ and an assignment (n -permutation) $\pi \in \mathcal{S}_n$, consider the *Hamiltonian*

$$\mathcal{H}(\pi) = \sum_{i=1}^n \mathcal{D}^p(B_i, R_{\pi(i)})$$

and the random variable “**ground state energy**”

$$\mathcal{H}_{\text{opt},(n,d)}^{(p)} = \min_{\pi \in \mathcal{S}_n} \mathcal{H}(\pi) \quad (\pi_{\text{opt}} = \arg \min_{\pi \in \mathcal{S}_n} \mathcal{H}(\pi)).$$

Definition (Euclidean Random Assignment Problem).

Understand the statistical properties of $\mathcal{H}_{\text{opt},(n,d)}^{(p)}$ (exact or asymptotic law, moments, etc.) depending on (Ω, p, d) and $v_{\mathcal{B}(\mathcal{R})}$.

Three main motivations for ERAPs

- **Spin Glasses.** ERAP provides a toy-model of **spin-glass in finite dimension**. Besides disorder, the assignment constraint provides **frustration**. But the model is numerically simpler than e.g. Edwards–Anderson spin glass (Mézard–Parisi 1988).
- **Optimal Transport.** ERAP = **Monge–Kantorovitch** transportation problem on Ω ($\dim(\Omega) = d$) associated to the empirical measures $\rho_{\mathcal{B}(\mathcal{R})} = \frac{1}{n} \sum_j \delta_{B_j(R_j)}$. In particular

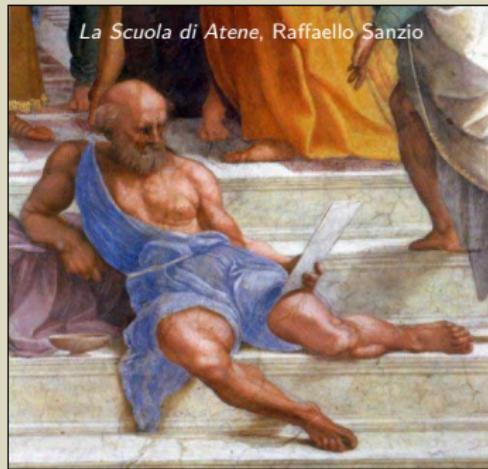
$$\mathcal{H}_{\text{opt},(n,d)}^{(p)} = n W_p^p(\rho_{\mathcal{B}}, \rho_{\mathcal{R}}),$$

with W_p is **p -Wasserstein dist.** (Villani 2009, Brezis 2018).

- **Computational Complexity Theory.** ERAPs are small modifications of random TSPs; but TSP is **NP-complete**.

Simplifications required: A minimalist approach

The exemple of Diogenes (412-323 BC)



A tool for understanding ERAPs: The phase diagram

Let us put $v_{\mathcal{B}} = v_{\mathcal{R}} := v$. We can start studying

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

as $n \rightarrow \infty$, depending on (p, d) and the choice of v .

Phase diagram: $(\gamma_{p,d}, \gamma_{p,d})$ are expected to be **largely insensitive** on the choice of v (which may alter the constant $K_{p,d}$).

Remark: Non-uniform disorder v is more subtle!

Example: Take $v = \text{standard Gaussian}$ and $(p, d) = (2, 1)$. Then

$$E_{2,1}(n) \underset{n \rightarrow \infty}{\sim} 2 \ln \ln n \quad (\text{i.e. } \gamma_{2,1} = \gamma'_{2,1} = 0).$$

(Caracciolo–D’A–Sicuro 2019, Bobkov–Ledoux 2019, Berthet–Fort 2020)

See (Benedetto–Caglioti 2020) for non-uniform case at $d = 2$.

$d \geq 3, p \geq 1, \Omega = \text{a bounded domain}$

“Simple”: Solution is realized at the scale of nearest-neighbors

$$E_{p,d}(n) \Big|_{d \geq 3, p \geq 1} \underset{n \rightarrow \infty}{\sim} K_{p,d} n^{\gamma_{LB}},$$

where

$$\gamma_{p,d} = \gamma_{LB} := 1 - \frac{p}{d}, \quad \gamma'_{p,d} = 0 \quad (\text{Mézard--Parisi 1988})$$

(if the disorder v is uniform on Ω , otherwise **unknown**).

Remark: The constants $K_{p,d}$ are **universal** (Barthe--Bordenave 2013 and refs. therein for $p < \frac{d}{2}$, and Goldman--Trevisan 2020 for an extension to $p \geq 1$) but **unknown explicitly**. Upper and lower bounds on some $K_{p,d}$ for $\Omega = [0, 1]^d$ are in (Talagrand 1992), numerical estimates for $\Omega = [0, 1]^d, \mathbb{T}^d$ are in (Caracciolo--Sicuro 2015, D'A MSc Thesis 2016).

$d = 2$: A challenge for both mathematicians and physicists

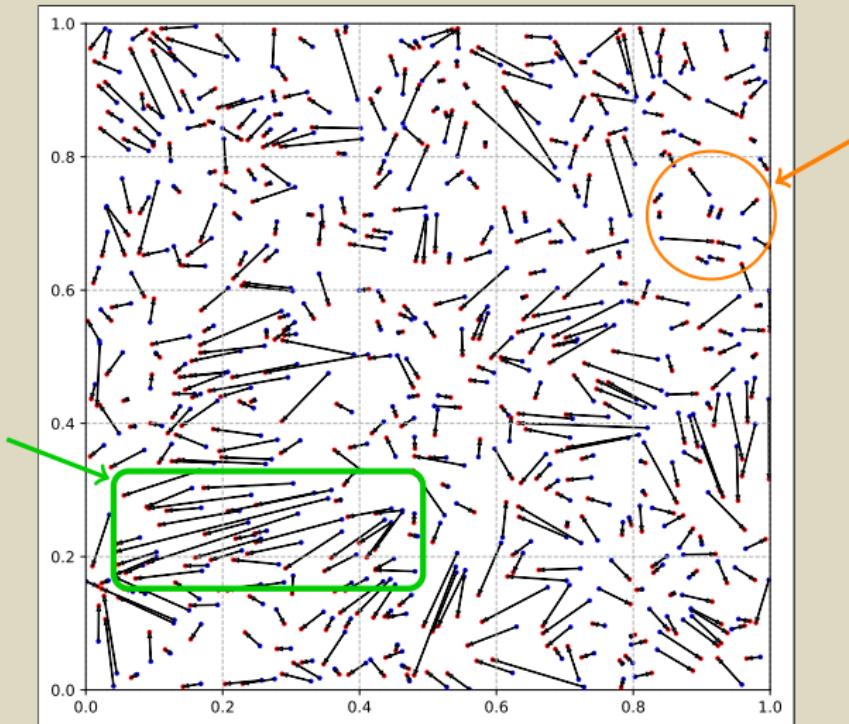
Example: $\Omega = [0, 1]^2$, $\mathcal{D} = |\cdot|$. A solution at $p = 2$:

The solution π_{opt} connects typically $O(\ln n)$ -nearest-neighbors.

$$\gamma_{p,d} = \gamma_{\text{LB}} = 1 - \frac{p}{2}$$

$$\gamma_{p,d} = \frac{p}{2}$$

(Ajtai–Komlós–Tusnády 1984).



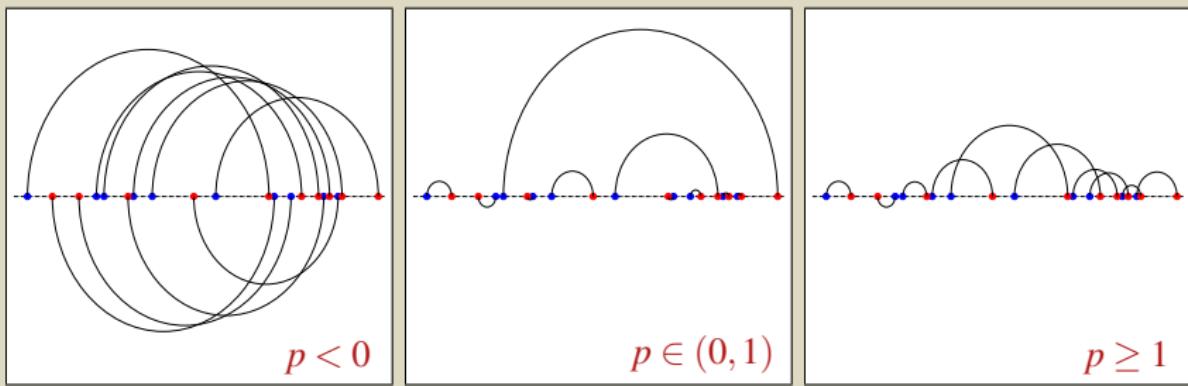
A few recent developments at $(p,d) = (2,2)$

- 2014 Caracciolo–Lucibello–Parisi–Sicuro (Phys. Rev. E): Using a **classical field-theoretical approach**, first predicted $K_{2,2} = \frac{1}{2\pi}$.
- 2019 Ambrosio–Stra–Trevisan (PTRF): **Proof** of $K_{2,2} = \frac{1}{2\pi}$
- 2020 Ambrosio–Glaudo (JEP): Refinement on the **remainder term**
- 2020 Benedetto–Caglioti (J. Stat. Phys.): non-constant densities (UB)
- 2021 Benedetto–Caglioti–Caracciolo–D’A–Sicuro–Sportiello
(J. Stat. Phys.): **Exact formula** for $\lim_{n \rightarrow \infty} [E_\Omega(n) - E_{\Omega'}(n)]$, where Ω, Ω' are two manifolds.
- 2021 Caracciolo *et al.*: Link with loop-erased **SARW**.
- 2022 Ambrosio–Goldman–Trevisan : **Proof** of universality of $K_{2,2}$.

...

$d = 1$: Qualitative properties of a solution π_{opt}

For any v , $p = 0$ and $p = 1$ separate **three qualitative regimes**:



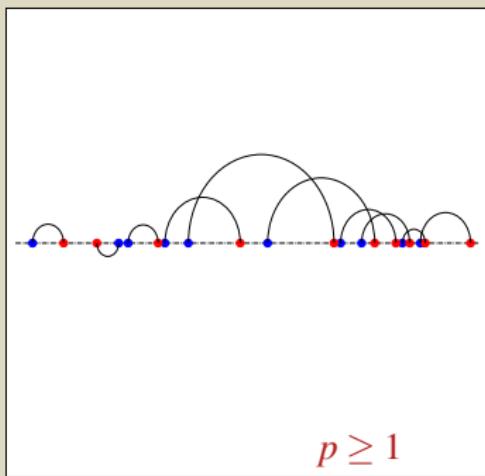
If (b_1, \dots, b_n) and (r_1, \dots, r_n) are sorted in natural order π_{opt} is:

- **Cyclical** for $p < 0$: $\pi_{\text{opt}} = i + k \pmod{n}$ (Caracciolo–D’A–Sicuro 2017);
- **Non-crossing** for $p \in (0, 1)$: Intervals “covered by edges” are either disjoint or one is included into the other (McCann 1999);
- **Ordered** for $p \geq 1$: π_{opt} is the identity permutation.

$$\Omega = [0, 1] \text{ (or } \mathbb{R}, \mathbb{R}^+, \dots\text{)}, \mathcal{D} = || \text{ and } p \geq 1$$

Take $B_{j+1} \geq B_j$, $R_{j+1} \geq R_j$, for $j = 1, \dots, n-1$.

Optimality + **(strict) convexity** + (strict) monotonicity of \mathcal{D}^p



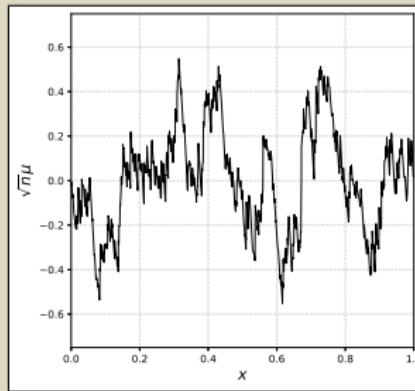
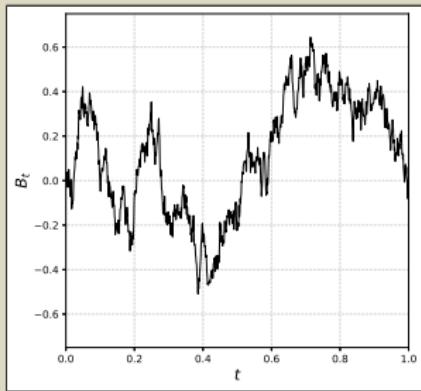
$$\mathcal{H}_{\text{opt},(n,p)}|_{p \geq 1} = \sum_{i=1}^n |B_i - R_i|^p.$$

⇒ Stairway to the **Brownian world!**

Brownian Bridge for $\Omega = [0, 1]$, $\mathcal{D} = |\cdot|$, $p \geq 1$

Let the **transport field** be $\mu_i := b_i - r_i$, for $i = 1, \dots, n$ and put $i = nt + \frac{1}{2}$. Then by **Donsker's Theorem**,

$\sqrt{n} \mu_i \xrightarrow{\text{weakly}} B_t$, the Brownian Bridge.



Recall $\mathcal{H}_{\text{opt},(n,p)} = \sum_{i=1}^n |\mu_i|^p$. Then $E_{p,1}(n)|_{p \geq 1} \underset{n \rightarrow \infty}{\sim} \mathbb{E}[B_t^p] n^{1-\frac{p}{2}}$.

(Boniolio–Caracciolo–Sportiello 2014, Caracciolo–Sicuro 2014, Caracciolo–D’A–Sicuro 2017)

$\Omega = \mathbb{S}_1$, \mathcal{D} = arc-distance, $p = 2$: The limit distribution of \mathcal{H}_{opt}

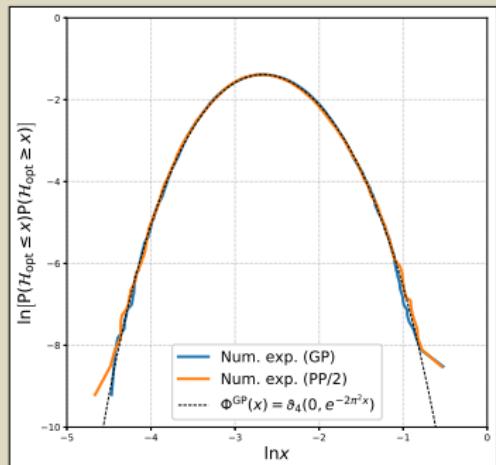
The $n \rightarrow \infty$ cdf of $\mathcal{H}_{\text{opt},(n,2)}$ on \mathbb{S}_1 is **explicit**. Sketch of proof:

1. At $p = 2$ we have **Parseval identity** for μ ;
2. The mgf $\mathbb{E} \left[e^{-w \sum_{s \neq 0} |\hat{\mu}_s|^2_s} \right] = \prod_{s \geq 1} \frac{1}{1 + \frac{w}{2\pi^2 s^2}}$ can be

inverse-Laplace transformed in closed form (Watson 1961).

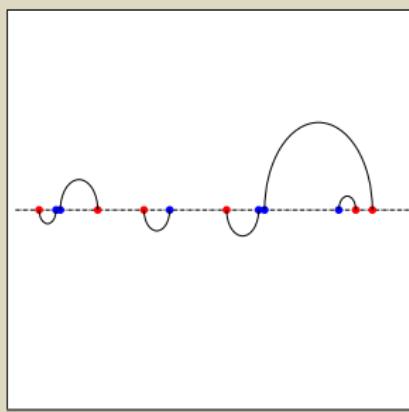
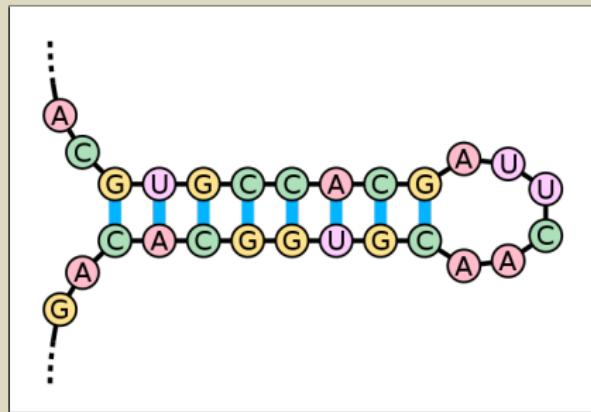
We have (D'A, 2020):

$$\begin{aligned} \mathbb{P}[\mathcal{H}_{\text{opt},(n,2)} \leq x] &\xrightarrow[n \rightarrow \infty]{} \vartheta_4(0, e^{-2\pi^2 x}) \\ &:= \sum_{s \in \mathbb{Z}} e^{i\pi s} e^{-2\pi^2 s^2 x}. \end{aligned}$$



ERAPs at $p \in (0, 1)$: Further motivation in biology

Besides economics, interest due to the non-crossing property of the solution: Toy-model for the **secondary structure of RNA** (discarding pseudo-knots).

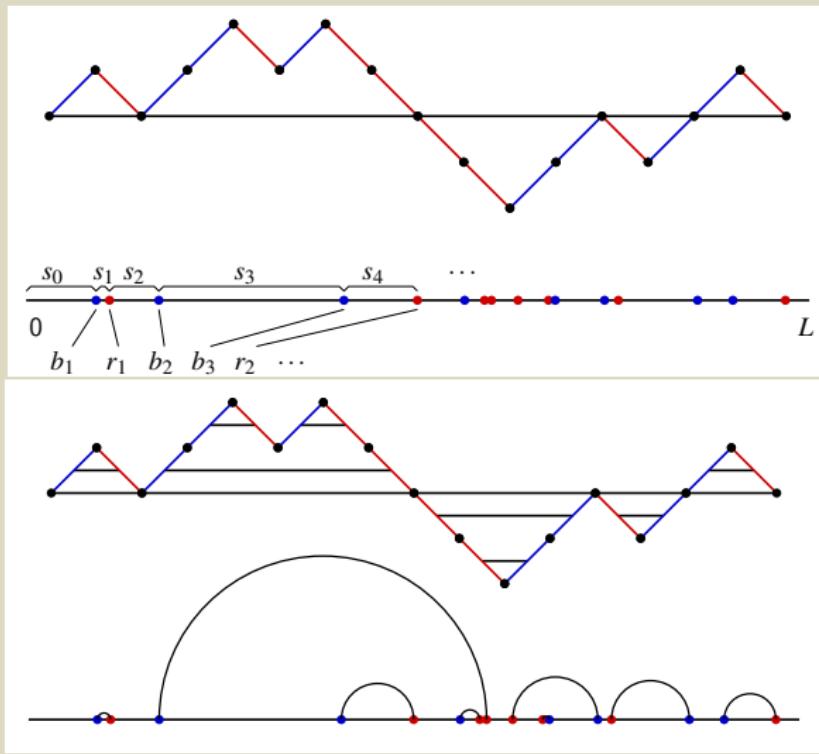


Despite this, poorly understood (McCann 1999).

IDEA: Try a canonical construction, an approximate solution sharing non-crossing property with $\pi_{\text{opt}} \rightarrow$ Dyck matchings!

The Dyck matching (Caracciolo–D’A–Erba–Sportiello 2020)

Construction:



The Dyck Conjecture at $p \in (0,1)$

Expected energy of Dyck matchings grows asymptotically as

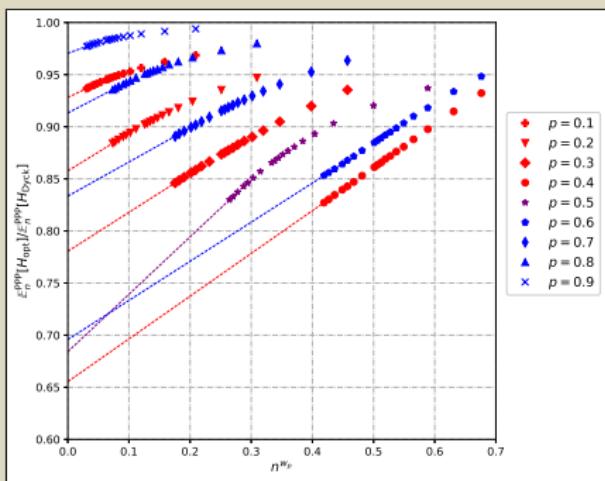
$$\mathbb{E}_n(\mathcal{H}_{\text{Dyck}}) \underset{n \rightarrow \infty}{\sim} \begin{cases} n^{1-p} & \text{if } 0 \leq p < \frac{1}{2} \\ \sqrt{n} \ln n & \text{if } p = \frac{1}{2} \\ n^{\frac{1}{2}} & \text{if } \frac{1}{2} < p \leq 1 \end{cases}.$$

Conjecture (Caracciolo–D’A–Erba–Sportiello 2020):

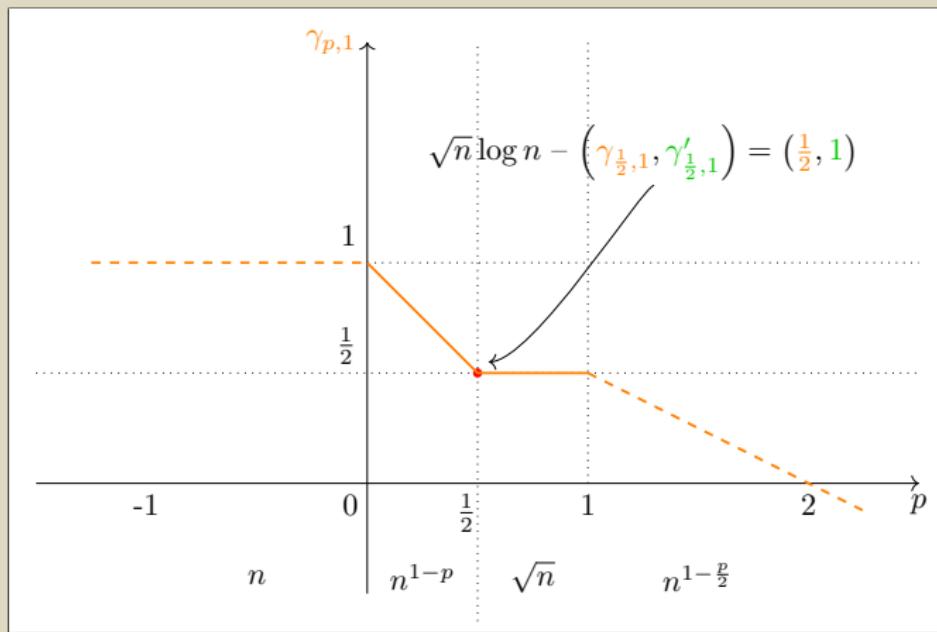
$$\lim_{n \rightarrow \infty} \frac{\mathbb{E}_n(\mathcal{H}_{\text{opt}})}{\mathbb{E}_n(\mathcal{H}_{\text{Dyck}})} = k_p,$$

for some $k_p \in (0,1)$.

Upper bound implied by (Fournier–Guillin 2014).



Section of the Phase Diagram at $d = 1$



Product formula for number of solutions at $d = p = 1$ (Caracciolo–Erba–Sportiello 2021).

Beyond uniform disorder: Bulk and anomalous scaling

Recall that for $\nu = \mathbf{1}_{[0,1]}$ and $p \geq 1$ (Caracciolo *et al.* 2019):

$$\begin{aligned} E_{p,1}^U(n) &= n \frac{\Gamma\left(1 + \frac{p}{2}\right)}{p+1} \frac{\Gamma(n+1)}{\Gamma\left(n+1 + \frac{p}{2}\right)} \\ &= c_p n^{1 - \frac{p}{2}} (1 + o(1)) = c_p n^{\gamma_{p,1}} (1 + o(1)). \end{aligned}$$

Thus $E_{p,1}^U(n)$ exhibits **bulk scaling**.

What if ν is non-uniform?

Definition. If $\gamma_{p,1} \neq 1 - \frac{p}{2}$ or if $\gamma_{p,1} \neq 0$, we say that $E_{p,1}(n)$ exhibits an **anomalous scaling**.

“Reduction to quadratures” in the bulk scaling

For R the cdf of v , let $\Psi^{(v)} := v \circ R_v^{-1}$. Bobkov–Ledoux 2019 get

$$E_{p,1}^{(v)}(n) = \frac{2^p}{\sqrt{\pi}} \Gamma\left(\frac{p+1}{2}\right) \int_0^1 \left[\frac{\sqrt{s(1-s)}}{\Psi^{(v)}(s)} \right]^p ds n^{1-p/2} + o(n^{1-p/2}).$$

Caracciolo–D’A–Sicuro 2018: **Regularize the integral** inspired by cutoff regularization in Quantum Field Theory.

Example: $v(x) = e^{-x}$, $\Psi^{\exp}(s) = 1 - s$.

Cutoff method: Stop integration at distance $\frac{c}{n}$ away from the singularity:

$$E_{p,1}^{\exp}(n) = \frac{2^p}{\sqrt{\pi}} \Gamma\left(\frac{p+1}{2}\right) \int_0^{1-c/n} \left(\frac{s}{1-s}\right)^{\frac{p}{2}} ds.$$

At $p = 2$, this gives:

$$E_{2,1}^{\exp}(n) = 2 \ln n - 2 \log c - 2 + o(1).$$

Exact result, $p = 2$ (Beta integrals):

$$\begin{aligned} E_{2,1}^{\exp}(n) &= 2 \sum_{k=1}^n \frac{1}{k} \\ &= 2 \ln n + 2\gamma_E + o(1). \end{aligned}$$

Rigorous approach to anomalous scaling

(D'A-Sportiello 2020-)

Guiding principle: Only the local properties of v in a neighbourhood of the zero will determine the leading anomalous behaviour. So we decided to estimate contrib. of k -th edge at **fixed n** in

$$\mathcal{H}_{\text{opt},(n,1)}^{(p)}|_{p \geq 1} = \sum_{k=1}^n |b_k - r_k|^p.$$

The general strategy:

1. **Binomial expand** $\mathcal{H}_{\text{opt},(n,1)}^{(p)}|_{p \geq 1}$ for $p > 1$ even and fixed n ;
2. Use **linearity** to compute expected contrib. of k -th edge;
3. **Asymptotic analysis** for $n \rightarrow \infty$ (\implies local properties of v);
4. **Analytic continuation** of result $\forall p \geq 1$.

A few details on the general strategy

For a disorder v , we have for Step 1-2:

$$E_{p,1}^{(v)}(n) = \sum_{k=1}^n \mathbb{E}[(b_k - r_k)^p] = \sum_{k=1}^n \sum_{q=0}^p \binom{p}{q} (-1)^{p-q} M_{n,k,q}^{(v)} M_{n,k,p-q}^{(v)},$$

where the ℓ -moment of the quantile function is

$$M_{n,k,\ell}^{(v)} = \langle R_v^{-1}(u)^\ell \rangle_{P_{n,k}}$$

and $\langle \dots \rangle_{P_{n,k}}$ denotes expectation w.r.t. $\text{Beta}(k, n-k+1)$:

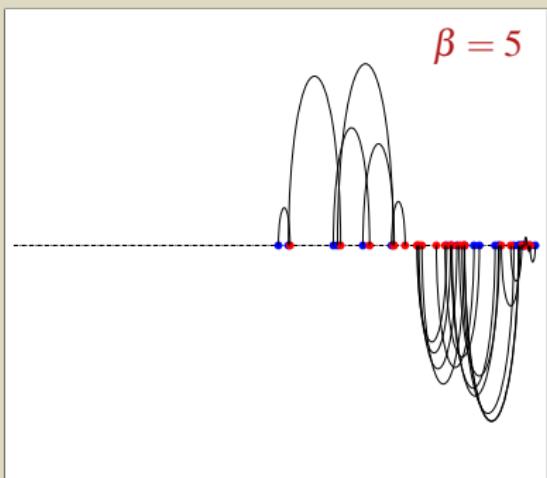
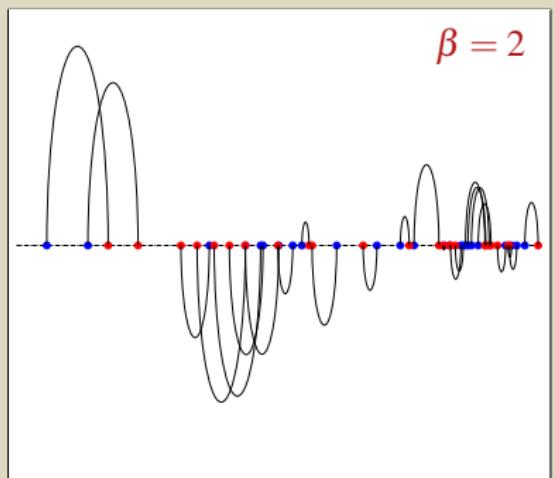
$$P_{n,k}(u)du := \frac{n!}{(k-1)!(n-k)!} u^{k-1} (1-u)^{n-k} du.$$

Example: Internal endpoint, algebraic zero $v_{fa,\beta}$

Consider the following pdf:

$$v_{fa,\beta}(x) = \beta x^{\beta-1}, \quad x \in [0, 1].$$

Example: Solutions at $n = 25$:



$$\text{Thus } R_{fa,\beta}(x) = x^\beta \implies R_{fa,\beta}^{-1}(u) = u^{\frac{1}{\beta}}.$$

Origin of bulk and anomalous regimes

Here the ℓ -moment of the quantile function is

$$M_{k,n;\ell}^{(\text{fa}),\beta} = \left\langle (u^{\frac{1}{\beta}})^\ell \right\rangle_{P_{n,k}} = \frac{\Gamma(k + \frac{\ell}{\beta})\Gamma(n+1)}{\Gamma(k)\Gamma(n+1 + \frac{\ell}{\beta})}$$

$$\underset{n \rightarrow \infty}{\sim} \begin{cases} \frac{\Gamma(k + \frac{\ell}{\beta})}{\Gamma(k)} n^{-\ell/\beta} & k \text{ small (**anomalous**)} \\ \exp\left(-\left(\frac{\ell^2}{\beta}\right)\left(\frac{1-x^{-1}}{2n}\right)\right) x^{\ell/\beta} & k = xn' \text{ (**bulk**)} \end{cases}.$$

A useful lemma in asymptotic analysis

Lemma (see e.g. Whittaker–Watson, Tricomi)

As $n \rightarrow \infty$, for β finite, the function $\psi_n(\beta) := \ln \left[\frac{n^\beta \Gamma(n)}{\Gamma(n+\beta)} \right]$ has the series expansion

$$\psi_n(\beta) = \sum_{\substack{k \geq 1 \\ 0 \leq \ell \leq k}} A_{k,\ell} n^{-k} \beta^{\ell+1} = \frac{\beta - \beta^2}{2n} + \frac{\beta - 3\beta^2 + 2\beta^3}{12n^2} + \dots$$

where

$$A_{k,\ell} = (-1)^k B_{k-\ell} \frac{\Gamma(k)}{\Gamma(\ell)\Gamma(k-\ell+1)}, \quad k \geq 1, \quad 0 \leq \ell \leq k,$$

and B_s the *s-th Bernoulli number*.

Leading asymptotics in the bulk regime

Result

Lemma 1 (bulk regime, D'A-Sportiello 2020)

For the family of distributions $v_{fa,\beta} = \beta x^{\beta-1}$, $\beta \geq 1$,

$$E_p^\beta(n) \underset{n \rightarrow \infty}{\sim} b_{\beta,p} n^{1-p/2}, \quad 2\beta + 2p - p\beta > 0,$$

with

$$b_{\beta,p} = \frac{1}{\beta^p} \frac{\Gamma(1+p)\Gamma\left(1 - \frac{p(\beta-2)}{2\beta}\right)}{\Gamma\left(2 + \frac{p}{\beta}\right)}.$$

Leading asymptotics in the bulk regime

Sketch of proof

1. Lemma implies

$$E_{(\beta,p),n}(k) \underset{n \rightarrow \infty}{\sim} \frac{1}{\Gamma(p/2+1)} x^{p/\beta} \left(\frac{-1+x^{-1}}{n\beta^2} \right)^{p/2}, \quad k = xn' \text{ (bulk)}.$$

2. Put $\Lambda = \mathcal{O}(\sqrt{n})$. Then $E_p^{\beta,\text{bulk}}(n) = \sum_{k=\Lambda}^n E_{(\beta,p),n}(k)$ can be transformed into an integral **without affecting the leading asymptotics**. This gives

$$E_p^{\beta,\text{bulk}}(n) \underset{n \rightarrow \infty}{\sim} \int_{\Lambda/n}^1 dx x^{\frac{p}{\beta} - \frac{p}{2}} P(x) n^{1-p/2},$$

for a polynomial $P(x)$. **Remark:** Convergence if $2\beta + 2p - p\beta > 0$.

3. Change of variables + Beta- Γ identity.

Leading asymptotics in the anomalous regime

Result

Lemma 2 (anomalous regime, D'A–Sportiello 2020).

For the family of distributions $v_{fa,\beta} = \beta x^{\beta-1}$, for $\beta \geq 1$,

$$E_p^{\beta, \text{anom}}(n) \underset{n \rightarrow \infty}{\sim} a_{\beta,p} n^{-p/\beta}, \quad 2\beta + 2p - p\beta < 0,$$

with

$$a_{\beta,p} = \Gamma(-1 - p/\beta) \sum_{q=0}^p \binom{p}{q} (-1)^q \frac{\Gamma(1 + \frac{q}{\beta}) \Gamma(1 + \frac{p-q}{\beta})}{\Gamma(-\frac{q}{\beta}) \Gamma(-\frac{p-q}{\beta})}.$$

Leading asymptotics in the anomalous regime

Sketch of proof

- Recall that for $k = o(\sqrt{n})$ (**anomalous regime**)

$$\sum_{k=1}^{\Lambda} E_{(\beta,p),n}(k) \sim n^{-p/\beta} \sum_{k=1}^{\Lambda} \frac{1}{\Gamma^2(k)} \sum_{q=0}^p \binom{p}{q} (-1)^q \Gamma\left(k + \frac{q}{\beta}\right) \Gamma\left(k + \frac{p-q}{\beta}\right).$$

- Remark:** \sum_k^{Λ} converges if $2\beta + 2p - p\beta < 0$. In this case we can take $\Lambda \rightarrow \infty$ (we “remove the infrared cutoff”).

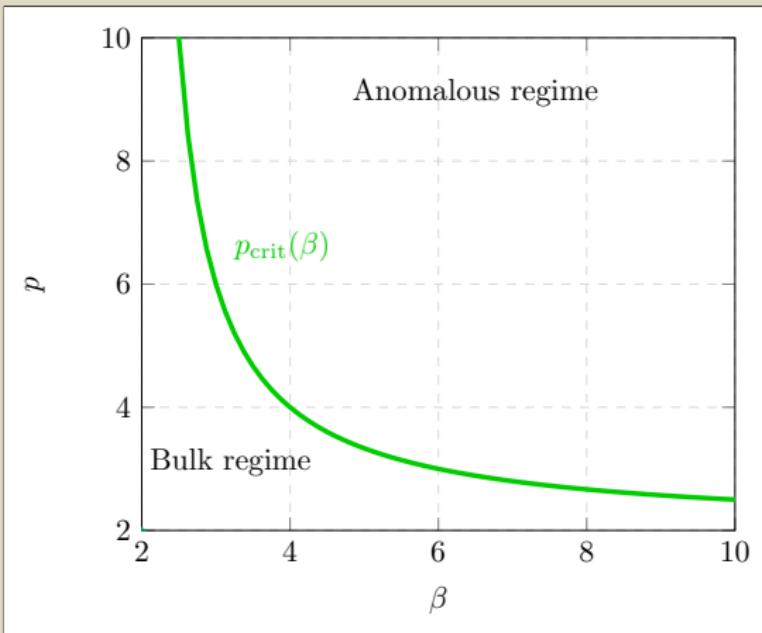
- Recall the hypergeometric identity

$$\sum_{k=1}^{\infty} \frac{\Gamma(k+a)\Gamma(k+b)}{\Gamma^2(k)} = \Gamma(-1-a-b) \frac{\Gamma(a+1)}{\Gamma(-a)} \frac{\Gamma(b+1)}{\Gamma(-b)}.$$

Bulk meets anomalous: The critical hyperbola

Equating leading exponents gives **the critical hyperbola**:

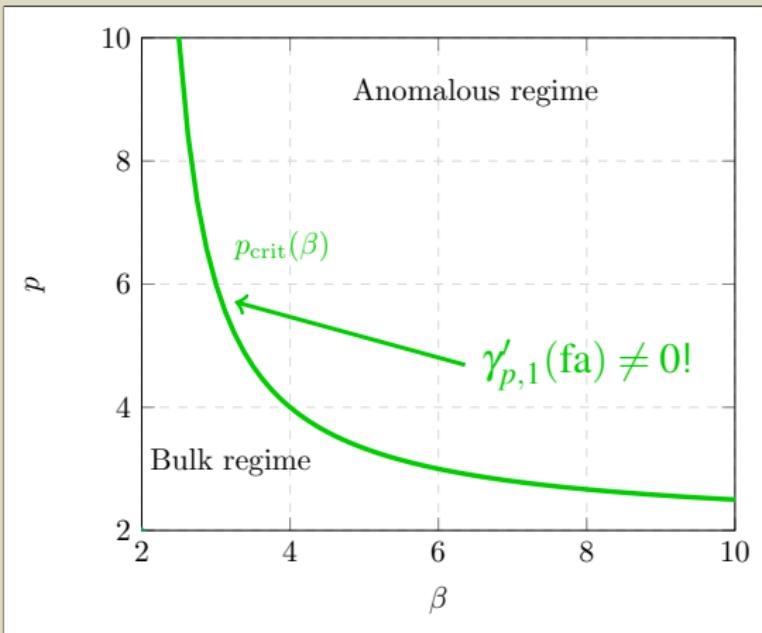
$$2\beta + 2p = p\beta, \quad p \geq 1, \beta \geq 1.$$



Bulk meets anomalous: The critical hyperbola

Equating leading exponents gives **the critical hyperbola**:

$$2\beta + 2p = p\beta, \quad p \geq 1, \beta \geq 1.$$



$\gamma'_{p,1} \neq 0$ along the critical hyperbola

Main result

Theorem 1 (D'A-Sportiello 2020)

For the family of distributions $v_{fa,\beta}(x) = \beta x^{\beta-1}$, for $\beta \geq 1$,

$$E_p^\beta(n) \underset{n \rightarrow \infty}{\sim} \frac{p!}{(p/2)!} \left(\frac{p-2}{2p} \right)^p n^{1-p/2} \log n, \quad 2\beta + 2p = p\beta.$$

Thus along the critical hyperbola:

$$\gamma_{p,1} = 1 - \frac{p}{2}, \quad \gamma'_{p,1} = 1$$

independently on β !

$\gamma'_{p,1} \neq 0$ along the critical hyperbola

Sketch of proof

1. Here $E_{(n,p)} \sim E_{(n,p)}^{\text{bulk}} + E_{(n,p)}^{\text{anom}}$ at leading order in n . We need to keep a **finite cutoff** $1 \ll \Lambda(n) \ll n$ throughout the calculation.

2.

$$E_{(n,p)}^{\text{bulk}} = \int_{\Lambda/n}^1 dx \frac{1+xR(x)}{x} \underset{n \rightarrow \infty}{\sim} \int_{\Lambda/n}^1 dx \frac{1}{x} + \int_0^1 dx R(x) ,$$

where $R(x)$ is a polynomial.

3. One proves (after some work involving 1) hypergeometric identities and 2) the Gauss digamma theorem) that

$$E_{(n,p)}^{\text{anom}} = o(n^{1-p/2} \log n).$$

On edge stochastic domination

Main result

Stochastic order of random variables: Two random variables X and Y are in *stochastic order*, written $X \leq_{\text{st}} Y$, if

$$\mathbb{P}(X \geq t) \leq \mathbb{P}(Y \geq t), \quad \forall t \in \mathbb{R}.$$

Theorem (D'A–Liu 2022)

Let $\Omega_1, \Omega_2, \Omega_3, \Omega_4$ be four domains and $\mathcal{B}^{(\Omega_l)}$ ($\mathcal{R}^{(\Omega_k)}$) n i.i.d. r.v.s over Ω_l (Ω_k), for $l = 1, 3$ ($k = 2, 4$).

If $c\left(\mathcal{B}^{(\Omega_3)}, \mathcal{R}^{(\Omega_4)}\right) \leq_{\text{st}} c\left(\mathcal{B}^{(\Omega_1)}, \mathcal{R}^{(\Omega_2)}\right)$ then

$$\mathbb{E}[\mathcal{H}_{\text{opt}}^{\Omega_3, \Omega_4}] \leqslant \mathbb{E}[\mathcal{H}_{\text{opt}}^{\Omega_1, \Omega_2}], \quad \forall n \geqslant 1.$$

On edge stochastic domination

Sketch of proof

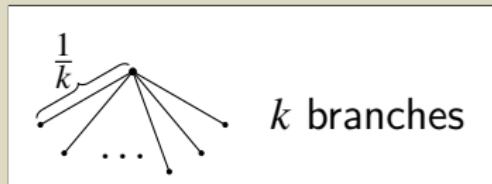
6 steps:

1. **Positivity:** $c \geq 0 \implies \mathbb{E}[c(B^{(\Omega_1)}, R^{(\Omega_2)})] \geq \mathbb{E}[c(B^{(\Omega_3)}, R^{(\Omega_4)})]$
2. **Monotonicity :** if for $i = 1, \dots, m$, $X_i \leq_{\text{st}} Y_i \implies \sum_{i=1}^m X_i \leq_{\text{st}} \sum_{i=1}^m Y_i$
3. **Apagogical argument:** If for $i = 1, \dots, m$, $X_i \leq_{\text{st}} Y_i \implies \min X_i \leq_{\text{st}} \min Y_i$
4. **Independence** of \mathcal{B} and \mathcal{R}
5. **Equidistribution** in product space
6. Conclusion.

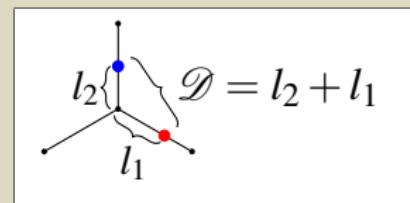
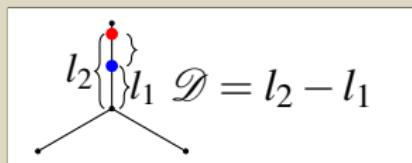
Application 1: ERAP on the k -star graph

Choice of $((\Omega, \mathcal{D}), (v_{\mathcal{B}}, v_{\mathcal{R}}), p)$:

- $\Omega = \mathbb{K}_{1,k}$ (i.e. the star graph with k branches)



with distance \mathcal{D} :

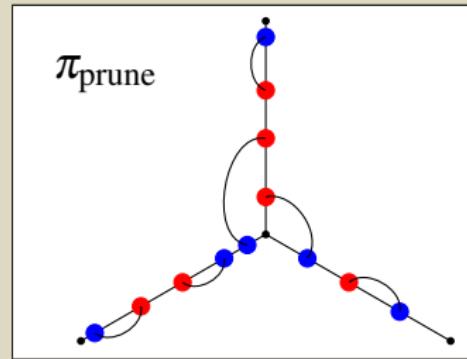
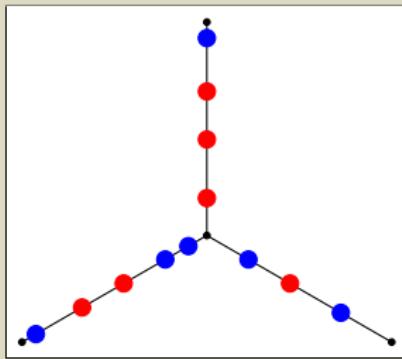


- $v_{\mathcal{B}} = v_{\mathcal{R}} = v$, $v \sim U[0, \frac{1}{k}] \otimes \mathcal{U}_{\{1, \dots, k\}}$;
- $p \geq 1$.

On the configuration π_{prune}

IDEA: Try a canonical construction, an approximate solution sharing local ordering property with $\pi_{\text{opt}} \rightarrow \underline{\pi_{\text{prune}}}$!

Example: $k = 3$ and $n = 6$



On the configuration π_{prune}

Main result : $\mathbb{E}_v \left[\mathcal{H}_{\text{opt}}^{(p)} \right] \underset{n \rightarrow \infty}{\sim} \mathbb{E}_v \left[\mathcal{H}_{\text{prune}}^{(p)} \right], \quad p \geq 1.$

The phase diagram: $\gamma_{p,1}^{(k)} = 1 - \frac{p}{2}$, $\gamma'_{p,1}^{(k)} = 0$.

And we get explicit (and simple) formulas for limit constants (partial moments of Brownian Bridge).

Example: At $p = 2$,

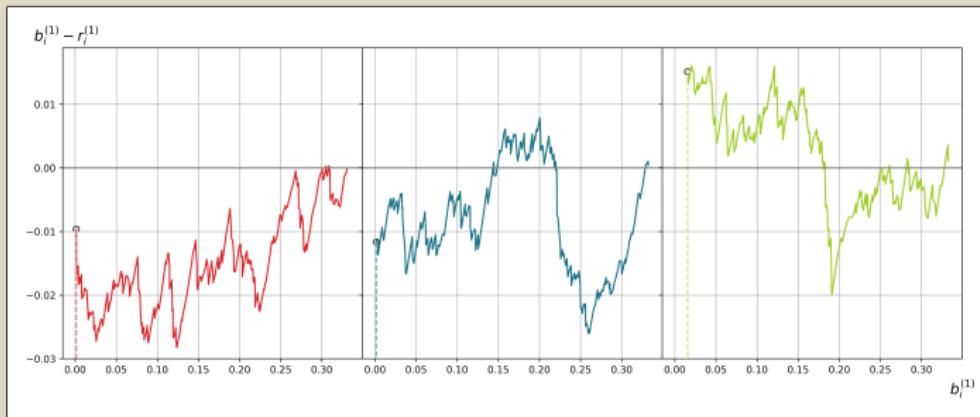
$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\mathcal{H}_{\text{opt}}^{(2)} \right] = \frac{3k - 2}{3k^2}.$$

On the configuration π_{prune}

Example: At $p = 2$, $\forall k \geq 1$,

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\mathcal{H}_{\text{opt}}^{(2)} \right] = \frac{3k - 2}{3k^2}.$$

Sketch of proof: Donsker's theorem for k independent Brownian Bridges (stopped at $\frac{1}{k}$).



A result for the star graph

Proposition (D'A–Liu 2022)

Let Ω_b, Ω_r be identical 3-star graphs, i.e. $\Omega^* := \Omega_b = \Omega_r$, and let $c = D^2(B^{(\Omega^*)}, R^{(\Omega^*)})$. Then

$$\frac{4n}{27(n+1)} \leq \mathbb{E}[\mathcal{H}_{\text{opt}}^{\Omega^*, \Omega^*}] \leq \frac{n}{3(n+1)}, \quad \forall n \geq 1.$$

Application 2: Family of area 1 ellipses

Consider the 1-parameter family of ellipses

$$\mathcal{E}_\lambda = \{(x, y) \in \mathbb{R}^2, \text{ s.t. } (\lambda x)^2 + \left(\frac{y}{\lambda}\right)^2 \leq \frac{1}{\pi}, \lambda \in (0, 1]\}$$

and let

$$\mathcal{H}_{\text{opt}}^\lambda := \min_{\pi \in \mathcal{S}_n} \sum_{i=1}^n D^2(\mathbf{B}_i, \mathbf{R}_{\pi(i)})$$

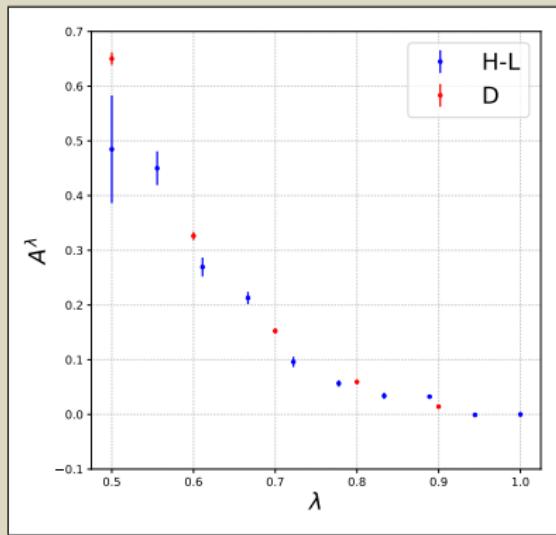
where $\mathbf{B}_i, \mathbf{R}_j$ i.i.d. $\sim \mathcal{U}(\mathcal{E}_\lambda)$ rvs, $1 \leq i, j \leq n$.

Then

$$\mathbb{E}[\mathcal{H}_{\text{opt}}^\lambda] \geq \mathbb{E}[\mathcal{H}_{\text{opt}}^1], \forall n \geq 1, \forall \lambda \in (0, 1].$$

Remark: special case of a much more general theorem.

Bonus: Numerical verification



Assumption

$$\overline{\mathcal{H}_{\text{opt}}^\lambda} - \overline{\mathcal{H}_{\text{opt}}^1} \underset{n \rightarrow \infty}{\sim} A^\lambda + O\left(\frac{1}{n}\right)$$

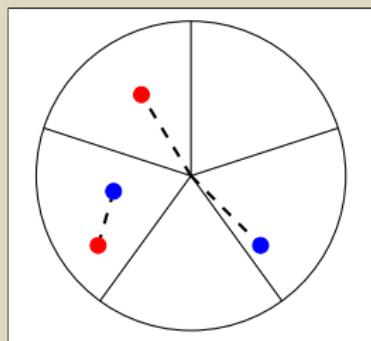
(Simulations performed students Mouad HAI et Yilun LI during their M1 TER and D'A)

1. ERAPs on convex domains with wormholes at $d \leq 2$

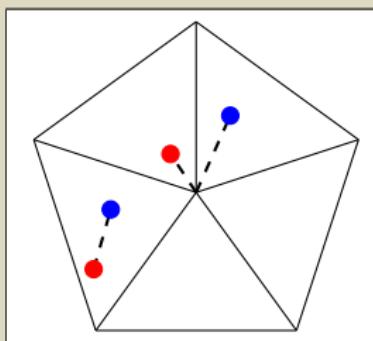
(W.i.p. with Sergio Caracciolo, Yuqi Liu and Andrea Sportiello, 2021-)

Example 2: k -identical $2d$ domains $(\Omega_i)_{i=1}^k$ “glued” by a point (wormhole) on the border.

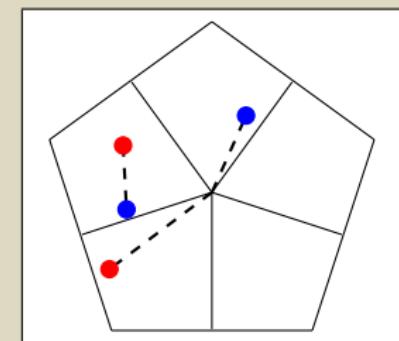
At $k = 5$:



Model 1



Model 2



Model 3

$$\nu_{\mathcal{B}} = \nu_{\mathcal{R}} \sim U(\Omega) \otimes \mathcal{U}_{\{1, \dots, k\}}; \quad p = 2. \quad \text{Remark: } \sum_{j=1}^k |\Omega_j| = 1.$$

1. ERAPs on convex domains with wormholes at $d \leq 2$

(W.i.p. with Sergio Caracciolo, Yuqi Liu and Andrea Sportiello, 2021-)

Is the exp. contribution due to “swapping points” negligible as $n \rightarrow \infty$?

From Caracciolo–Lucibello–Parisi–Sicuro (Phys. Rev. E 2014), Ambrosio–Stra–Trevisan (PTRF 2019) and Benedetto–Caglioti (J. Stat. Phys. 2020)+ Ambrosio–Goldman–Trevisan (EJP 2022), one expects for a uniform disorder:

$$\mathbb{E} [\mathcal{H}_{\text{opt}}] \underset{n \rightarrow \infty}{\sim} \frac{\sum_{j=1}^k |\Omega_j|}{2\pi} \log n + o(\log n) = \frac{1}{2\pi} \log n + o(\log n).$$

However, our experimental findings show:

- The limit constant f appears to be independent on model \curvearrowleft ;
- $f > \frac{1}{2\pi}!$

2. Nice formulas in one dimensional ERAPs

(following discussions with Nathanaël Enriquez, Sophie Laruelle and Andrea Sportiello, 2021-)

Take $\Omega = [0, 1]$, $\mathcal{D} = |\cdot|$ and $v_{\mathcal{B}} = v_{\mathcal{R}} := v = \mathbf{1}_{[0,1]}(x)$.

Using generalised Selberg integrals, Caracciolo *et al.* 2019 have shown, for integer ℓ ,

$$\mathbb{E}[|b_k - r_k|^\ell] = \frac{\Gamma^2(n+1)\Gamma(k + \frac{\ell}{2})\Gamma(n-k+1 + \frac{\ell}{2})\Gamma(1+\ell)}{\Gamma(k)\Gamma(n-k+1)\Gamma(n+1 + \frac{\ell}{2})\Gamma(n+1+\ell)\Gamma(1 + \frac{\ell}{2})}.$$

This implies the exact formula for $E_{p,1}|_{p \geq 1}(n)$ for $f = |\cdot|^p$.

Question*: Which $f = f(|\cdot|)$ give a “nice formula” for $E(n)$ after resummation?

*Question raised by N. Enriquez (Paris-Saclay (Orsay)) at this talk for ALEA Days 2021, online at CIRM Marseilles - Luminy, 18 March 2021.

3. Dyck matchings and the clustering function in the KPKVB model

(W.i.p. with Pim van der Hoorn, 2021-)

- Let $\gamma_\alpha(k)$ be the local clustering coefficient avg. over all vertices of degree k in the KPKVB model of random graphs in hyperbolic geometry of parameter α (see [here](#), Prop. 1.4);
- Let $E_{\text{Dyck}}^{(p)}(n)$ be exp. energy of Dyck matchings in ERAP with $p \in (0, 1)$ and (e.g.) uniform disorder (see Slide [21](#)).

Then

$$k(n)^2 \gamma_{\alpha(p)}(k(n)) \underset{k,n \rightarrow \infty}{\sim} C n^{-p} E_{\text{Dyck}}^{(p)}(n),$$

under

$$\begin{cases} k(n) = \sqrt{n} \\ \alpha(p) = \frac{1+p}{2} \end{cases} \quad \text{or} \quad \begin{cases} n(k) = k^2 \\ p(\alpha) = 2\alpha - 1 \end{cases}$$

where C depends only on α .

Question[†]: Why?

[†]After talks by the author and by P. van der Hoorn (TU Eindhoven) at Bézout-Eurandom meeting, Institut Henri Poincaré, Paris, 7 July 2021.

4. Schrödinger problem approach for asymptotic constants

(W.i.p. with Alberto Chiarini and Oliver Tse, 2021-)

At $d \geq 3$ and $p = 2$, and uniform disorder, the PDE technique by Ambrosio–Stra–Trevisan (PTRF 2019) recovers the exponent $\gamma_{2,d} = \frac{d-2}{d}$, but the leading constants $K_{2,d}$ remain inaccessible: \mathcal{B} and \mathcal{R} “remain gaussians” of arbitrary variance as $n \rightarrow \infty$ if $d \geq 3$.

A promising path to bound (or possibly get!) these $K_{2,d}$'s appears if one considers ERAP as a limit of the **Schrödinger problem** in the sense of Γ -convergence (Léonard 2012). We are currently investigating this route using **Evolutionary Variational Inequality** and the **dual Schrödinger problem** for the lower bound, and **path measure** techniques for the upper bound.

Thank you for your attention!