Transport rates for dependent point processes

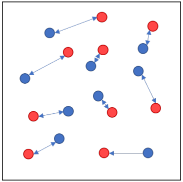
Raphaël Lachièze-Rey (Inria Paris and Université Paris Cité)
With Yogeshwaran D. (ISI Bangalore)

AKT Theorem

- Let $X_1, \dots, X_n, Y_1, \dots, Y_n$ i.i.d. uniform points in $\Lambda_n := [0, n^{1/d})^d$.
- Denote the minimal matching cost as

$$W_1 = \min_{\sigma \in \Sigma_n} \sum_{i=1}^{n} \|X_i - Y_{\sigma(i)}\|.$$

- What is the magnitude of W_1? (for instance the expectation)

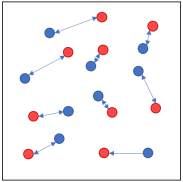


blue: X_i, red: Y_j, $\mathrm{Vol}\ (\Lambda_n) = n$.

AKT Theorem

- Let $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ i.i.d. uniform points in $\Lambda_n := [0, n^{1/d})^d$.
- Denote the minimal p-matching cost as

$$W_p^p = \min_{\sigma \in \Sigma_n} \sum_{i=1}^{n} \|X_i - Y_{\sigma(i)}\|^p, p \geqslant 1.$$



blue: X_i, red: Y_j

Theorem (Ajtaí-Komlós-Tusnadý '84)

With high probability, as $n \to \infty$,

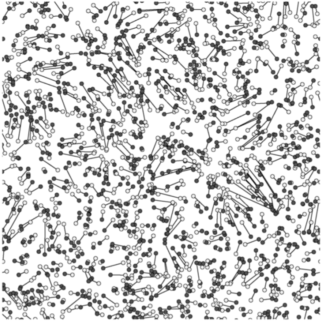
$$W_p^p \sim \begin{cases} n n^{p/2} & \textit{if } d = 1 \textit{ (edges measure typically } \sqrt{n}) \\ n \ln(n)^{p/2} & \textit{if } d = 2 \textit{ (edges measure typically } \sqrt{\ln(n)}) \\ n & \textit{if } d \geqslant 3 \textit{ (edges measure typically } 1) \end{cases}$$

Infinite version

[Holroyd, Pemantle, Peres, Schramm 2009, and Huessman and Sturm 2012]

Let μ, ν two i.i.d. homogeneous UI Poisson processes on $\mathbb{R}^d$, how to assess the transport cost $\mu \to \nu$? $W_p(\mu, \nu)$ is clearly infinite...

A **matching** between μ and ν is a translation invariant one-to-one map $T : \mu \to \nu$ (it exists for any stationary point process under very loose conditions) **[Holroyd, Pemantle, Peres, Schramm '09]**



Infinite analogue of AKT Theorem

- The typical distance of a matching T is a variable X_T with distribution function

$$\mathbf{P}(X_T \geqslant r) := \mathbf{E} \sum_{x \in \mu \cap [0,1]^d} \mathbf{1}_{\{\|x - T(x)\| > r\}}.$$

- (Equivalent to transport cost of 0 in Palm measure)

Theorem (HPPS)

- *In dimension* $d = 1, 2$, $\mathbf{E}(X_T^{d/2}) = \infty$ *for all* T *but it is possible to construct matching* T *such that* $\mathbf{E}\left(\frac{X_T^{d/2}}{\ln(X_T)^2}\right) < \infty$
- *In dimension* $d \geqslant 3$, *for some* $0 < c < C < \infty$, $\mathbf{E}(\exp(CX_T)) = \infty$ *for all* T *but there exists* T *with* $\mathbf{E}(\exp(cX_T)) < \infty$.

Very brief history

- **[AKT '84]**: Seminal paper, upper and lower bounds whp (7 pages)
- **[Talagrand '92]**: Generalisation to non-uniform law
- **[Fournier Guilin '15]**: Large deviations, dependent systems
- **[Ambrosio, Stra, Trevisan '19]**: Computation of the constant

$$\lim_{n\to\infty} \frac{n}{\ln(n)} \mathbb{E}\left[W_2^2 \left(\sum_{i=1}^n \delta_{X_i}, \sum_{j=1}^n \delta_{Y_j} \right) \right] = \frac{1}{2\pi}$$

- Applications in Image Processing, Machine Learning, ...

Do we have the same rates for **dependent** sequences of points?

- **[Fournier and Guillin '15]**: AKT-type rates for ρ-mixing stationary sequences $(X_i)_{i=1,\dots,n}$, i.e. such that

$$|\mathrm{Corr}(f(X_i), g(X_j))| \leqslant \rho_{|i-j|} \text{ and } \sum_n \rho_n < \infty.$$

  for all $f, g \in L^2(\mu)$.
- **[Dedecker and Merlevède '17] [Ammous '23]** (Dedecker, Duval): W_1-Rates for samples of real-valued stationary sequences, under α-dependance of β-dependance
- **[Borda '21]**,**[Bobkov, Ledoux '21]** Similar results for mixing processes
- **[Clozeau, Mattesini '24]** Processes generated via mixing Markov chains

Infinite dependent systems

[Hoffman, Holroyd, Peres '13] Their method can be generalised to a non-Poisson processes μ, the bound is

$$F(r) := \mathbf{P}(X_T \geqslant r) \leqslant c \frac{\sqrt{\mathrm{Var}\,(\mu(B(0,r)))}}{r^d}$$

$$\mu(B(0,r)) := |\mu \cap B(0,r)| \text{ (seen as random measure)}$$

- **Poisson process**, or other standard process: $\mathrm{Var}\,(\#\mu \cap B(0,r)) \sim r^d$, hence $F(r) \leqslant r^{-d/2}$

$$\mathbf{E}\left(X_T^{d/2}\right) \leqslant 1 + \int_1^\infty \mathbf{P}(X_T^{d/2} \geqslant r)dr = 1 + \int_1^\infty r^{-1}dr \leqslant \infty \qquad :($$

$$\mathbf{E}\left(\frac{X_T^{d/2}}{\ln(X_T)^2}\right) < \infty \qquad :)$$

- Linear rate in dimension $d \geqslant 3$ ($\mathbf{E}(X_T^{d/2}) < \infty$)

Reduced pair correlation measure (RPCM)

Let $\mu = \sum_i \delta_{x_i}$ be a stationary point process with unit intensity, i.e.

$$\mathbb{E}(\mu(dx)) = |A|^{-1}\mathbb{E}(\mu(A)) = 1,\ A \subset \mathbb{R}^d.$$

Measure 2d-order asymptotic independence through

$$\beta(x) = \frac{\mathbb{P}(dx \in \mu \mid 0 \in \mu)}{dx} - 1 \qquad (\text{ Poisson: } \beta \equiv 0)$$

More formally: for $\varphi \geqslant 0$ with compact support

$$\mathbb{E}(\sum_{x,y\in\mu} \varphi(x,y)) = \underbrace{\int \varphi(x,x)dx}_{\text{diagonal terms}} + \int \varphi(x,y)dx(dy + \beta(dy)).$$

Closely related to variance : if $|\beta|(\mathbb{R}^d) < \infty$ (i.e. β is integrable)

$$\mathrm{Var}\left(\mu(B(0,r))\right) = |B(0,r)|\left(1 + \beta(\mathbb{R}^d)\right) + o(r^d).$$

Proposition (Lr,Yogeshwaran 24')

Assume the RPCM β is integrable. Then there is a matching T with

- *In dimension $d = 1, 2$*

$$\mathbf{E}\left(\frac{X_T^{d/2}}{\ln(X_T)^2}\right) < \infty$$

- *In dimension $d \geqslant 3$, $\mathbf{E}(X_T^2) < \infty$.*

(Better for $d = 3$, equal for $d = 1, 2, 4$, less good in $d \geqslant 5$)

Can we achieve a linear rate in dimension 2?

- What random sequences of points can achieve a better average rate than $\sqrt{\ln(n)}$ in dimension 2 (or $\sqrt{n}$ in dimension 1)?

Can we achieve a linear rate in dimension 2?

- What random sequences of points can achieve a better average rate than $\sqrt{\ln(n)}$ in dimension 2 (or $\sqrt{n}$ in dimension 1)?
- Obviously, a **perturbed lattice** works:

$$\mu_n = \sum_{k \in \mathbb{Z}^d \cap \Lambda_n} \delta_{k+U_k}, \qquad \nu_n = \sum_{k \in \mathbb{Z}^d \cap \Lambda_n} \delta_{k+U'_k}$$

for nicely behaved U_k, U'_k. In this case, a good matching is

$$T(k+U_k) = k + U'_k$$

transport cost is linear (proportionnal to n):

$$p\text{−Transport cost } \sim \mathbf{E}\Big(\sum_k \|U_k - U'_k\|^p\Big) \sim \text{ Const. } n \ll n\sqrt{\ln(n)}$$

Can we achieve a linear rate in dimension 2?

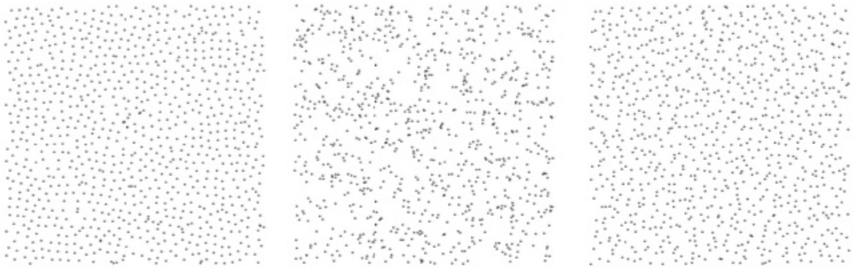
- What random sequences of points can achieve a better average rate than $\sqrt{\ln(n)}$ in dimension 2 (or $\sqrt{n}$ in dimension 1)?
- Obviously, a **perturbed lattice** works
- Is it possible to have a process with some "disordered behaviour"?

Hyperuniform processes

- Torquato, Stillinger, Lebowitz, Ghosh, etc...since the 90's
- A hyperuniform point process μ is such that

$$\mathrm{Var}(\mu(B(0,r))) = o(|B(0,r)|).$$

"Fluctuations are suppressed at large scales"



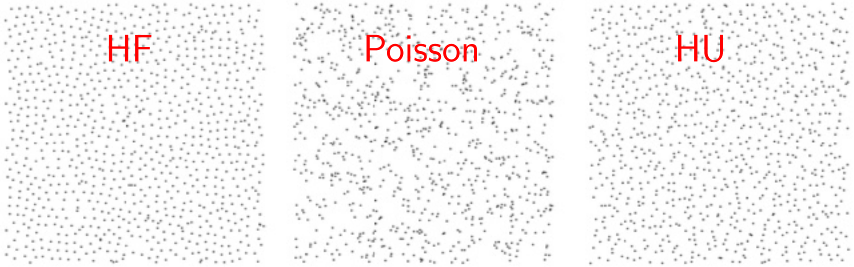
[Beliaev, Marinucci, Cammarota]

Hyperuniform processes

- Torquato, Stillinger, Lebowitz, Ghosh, etc...since the 90's
- A hyperuniform point process μ is such that

$$\mathrm{Var}(\mu(B(0,r))) = o(|B(0,r)|).$$

"Fluctuations are suppressed at large scales"



[Beliaev, Marinucci, Cammarota]

- A desirable property in physics is that μ is "amorphous" (mixing and isotropic?), i.e. $\neq$ from lattice.
  - Infinite Ginibre process - eigenvalues of random matrices/DPP/Gibbs measure
  - Projector DPPs
  - Some Coulomb gases
  - Zeros of the random planar Gaussian Analytic Function
  - ...
- Can be used to sample "efficiently" **[Bardenet & Hardy '20]**, **[Hawat, Bardenet, Lachieze-Rey '23]**

HU transport recent history - optimal rates in dimension $d = 2$ (large finite samples)

- **[Chafai, Hardy Maïda '18]**: Large deviations for N-samples of a 2D Coulomb gas with confining potential in d_{BL} - the Fortet-Mourier distance
- **[Jalowy '21]**: $\mathbb{E}(\sum_i \|X_i - Y_{\sigma(i)}\|) \sim n$ for the Ginibre process
- **[Prod'Homme '21]**: $\mathbb{E}(\sum_i \|X_i - Y_{\sigma(i)}\|^2) \sim n$ for Ginibre
- Bound generalised from **[HPPS '08]**:

$$F(r) \leqslant c \frac{\sqrt{\mathrm{Var}\left(\mu(B(0,r))\right)}}{r^d}$$

yields linear rate in dimension 2 if $\frac{\mathrm{Var}(\mu(B(0,r)))}{r^d} \leqslant c \ln(r)^{-2}$:

$$\mathbf{E}(X_T) < \infty$$

HU processes are L^2-perturbation of one another

- HU processes are believed to look like a lattice at large scales
  Global order and local disorder
- As such, they should have good transport properties to $\mathbb{Z}^d$!

Theorem (Lr, Yogeshwaran '24)

Let μ be a HU stationary point process.

- *If $d \geqslant 3$, there exists matching T such that $\mathbf{E}(X_T^2) < \infty$ if*

$$\int |\beta|(dx) < \infty$$

- *If $d = 2$, there exists matching T such that $\mathbf{E}(X_T^2) < \infty$ if*

$$\int \ln(1 + |x|)|\beta|(dx) < \infty$$

- (Very) recent / announced similar results from Butez, Dellaporta, Garcia-Zelada and Leblé, Huessman

Application to Determinantal point processes

- A stationary DPP is characterised by : for all $k \geqslant 1$

$$\mathbf{E} \sum_{x_1,\dots,x_k}^{\neq} \varphi(x_1,\dots,x_k) = \int |\underbrace{\det(K(x_i,x_j))}_{\sim \frac{\mathbf{P}(\{x_1,\dots,x_k\}\subset\mu)}{dx_1\dots dx_k}}| \varphi(x_1,\dots,x_k) dx_1 \dots dx_k$$

  where $K : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$ is Hermitian.
- **Poisson process:** $K(x,y) = \mathbf{1}[x=y]$
- **Infinite Ginibre ensemble:** $K(x,y) = e^{x\bar{y} - \frac{1}{2}|x|^2 - \frac{1}{2}|y|^2}; x,y \in \mathbb{C}$.

Proposition (d = 2)

μ is a L^2-perturbed lattice if

$$\int |K(0,z)|^2 dz = K(0,0)^2 \ (\mu \text{ is } HU) \ ,$$

$$\int \ln(|z|)|K(0,z)|^2 dz < \infty.$$

Optimal transport

- A matching T between two point samples $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_n\}$ is a particular instance of a **transport plan** between measure $\mu = \sum_i \delta_{x_i}, \nu = \sum_j \delta_{y_j}$: $M := \sum_i \delta_{(x_i, T(x_i))}$ satisfies the coupling relation

$$M(A, \mathbb{R}^d) = \mu(A),\ M(\mathbb{R}^d, B) = \nu(B),\ A, B \subset \mathbb{R}^d$$

- The transport distance between two finite measures μ, ν with same mass is often measured in terms of **Wasserstein distance**:

$$W_p^p(\mu, \nu) = \inf_{M \text{ coupling}(\mu,\nu)} \int \|x - y\|^p M(dx, dy), p > 0.$$

- Satisfies triangular inequality

Other formulations I

- **Finite samples:** Let μ stationary with integrable RPCM β,

$$\mu_n = \mu \mathbf{1}_{[0,n^{1/d}]^d}$$
$$\tilde{\mu}_n = \frac{n}{\mu([0,n^{1/d}]^d)} \mu_n \text{ (total mass } n) \ .$$

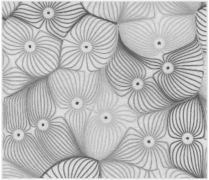
and $\mu', \mu'_n, \tilde{\mu}'_n$ similar with μ' other process with integrable RPCM.
Then in dimension $d = 2$

$$W_2^2(\tilde{\mu}_n, \tilde{\mu}'_n) \leqslant \alpha(n) := \begin{cases} cn\sqrt{\ln(n)} \text{ (general case, AKT rate)} \\ cn \text{ if } \mu \text{ is HU and } \ln(|x|)|\beta| \text{ integrable} \end{cases}$$

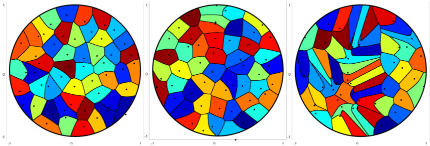
Other formulations II

- **Random allocation:** Let $\mathbf{Leb}_n^d$ be Lebesgue measure restricted to $[0, n^{1/d})^d$

$$W_2^2(\tilde{\mu}_n, \mathbf{Leb}_n^d) \leqslant \alpha(n)$$



Allocation of Poisson points
[Chatterjee, Peled, Peres, Romik , Annals of Maths '10]



[Jalowy '10]: *Left:* Ginibre, *Middle:* Eigenvalues of random matrix, *Right:* Poisson

Other formulations III

- **Perturbed lattices:** Let $\mathbf{Z}_n := \sum_{k\in\mathbb{Z}^d\cap[0,n^{1/d})^d} \delta_k$,

$$W_2^2(\mu_n', \mathbf{Z_n}) \leqslant W_2^2(\mu_n', \textbf{Leb}_n^d) + W_2^2(\textbf{Leb}_n^d, \mathbf{Z_n}) \leqslant 2\alpha(n)$$

"HU processes are lattices perturbed by a L^2 field":

$$\mu = \{k + T(k); k \in \mathbb{Z}^2\}$$

with T stationary and L^2. The $T(k)$ are **not** independent!

In conclusion, the variance reduction of HU entails a regular distribution of the points.

And futher...

- Toric distance
- General (stationary) random measures
- Unbalanced transport between μ_n and $\textbf{Leb}^d$ (or μ_n and $\mathbf{Z}^d$)

The Bobkov-Ledoux method

Relies on two arguments for a probability measure μ on $\Lambda_1 := [0, 2\pi)^d$:

- Fourier-Stieltjes transform for W_1

$$f_\mu(m) = \int_{\Lambda_1} e^{imx}\mu(dx), m \in \mathbb{Z}^d.$$

Then for μ with mass 1, for $T > 0$

$$\max\left[\tilde{W}_1^2(\mu, \textbf{Leb}_1^d), \tilde{W}_2^2(\mu, \textbf{Leb}_1^d)\right] \leqslant \sum_{1<|m|<\infty} \frac{1}{m^2}|f_\mu(m)|^2$$

- The $H^{-1} - W_2$ inequality: for μ proba with a density on Λ_1

$$W_2^2(\mu, \textbf{Leb}_1^d) \leqslant \int_{\Lambda_1} |\nabla\Delta^{-1}\left(\frac{d\mu}{dx} - 1\right)|^2 dx$$

Bobkov-Ledoux method (Cont'd)

- Smoothing by the density q_t of tH for some well-chosen random variable H with finite second moment

$$\tilde{\mu} = \mu \star q_t,\ |f_{\tilde{\mu}}(m)| = |\varphi_{tH}(m) f_\mu(m)|$$
$$\tilde{W}_p(\mu, \tilde{\mu}) \leqslant ct^{-2}$$

- Apply triangular inequality

$$W_p(\mu, \mathbf{Leb}_1^d) \leqslant W_p(\tilde{\mu}, \mathbf{Leb}_1^d) + W_p(\tilde{\mu}, \mu)$$

- Choose random variable H with second moment and spectrum support in Λ_1

$$\tilde{W}_p^2(\mu, \mathbf{Leb}_1^d) \leqslant \sum_{0<|m|<ct^{-1}} \frac{1}{m^2} |f_\mu(m)|^2 + ct^{-2}$$

- Finally, optimize in $t \Rightarrow\ t = n^{-1/d}$

BL method applied to stationary processes

- Hyperuniformity is more generally defined in terms of the "Structure factor" $S(dx)$, Fourier Transform of $\beta(dx)$

$$\mu \text{ Hyperuniform } \Leftrightarrow S(0) = 0$$

- BL method is expressed in terms of Fourier transform on finite samples, close to the "Scattering intensity"

$$S_n(k) = \frac{1}{N} \Big| \sum_{x \in \Lambda_n \cap \mu} \exp(ikx) \Big|^2$$

- Convergence $S_n(k) \to S(k)$ under some hypotheses

Thank you for your attention!