PhD Thesis proposal

Low dimensional Euclidean Random Assignment Problems and applications

Matteo D'ACHILLE & Jessie LEVILLAIN & Pascal MOYAL

matteo.d-achille@univ-lorraine.fr jessie.levillain@cnes.fr pascal.moyal@univ-lorraine.fr

May 9, 2026

Introduction. The assignment between two sets of points is a recurring problem, for example in satellite image processing. There, matching corresponding points in two different satellite images is needed for georeferencing an image using a reference image or for 3D reconstruction through stereo-photogrammetry. To do this, most approaches use local descriptors [4] that are based on the similarity of the corresponding areas from one image to another. However, while robust to noise and perturbations, these approaches encounter limitations when dealing with repetitive patterns (e.g., rows of vineyards or trees) or when matching objects between different modalities (e.g., an optical image and a Synthetic Aperture Radar image (SAR)). In these cases, for example, for matching the points corresponding to the same trees in two SAR images, the problem boils down to a random assignment problem preceded by a co-registration.

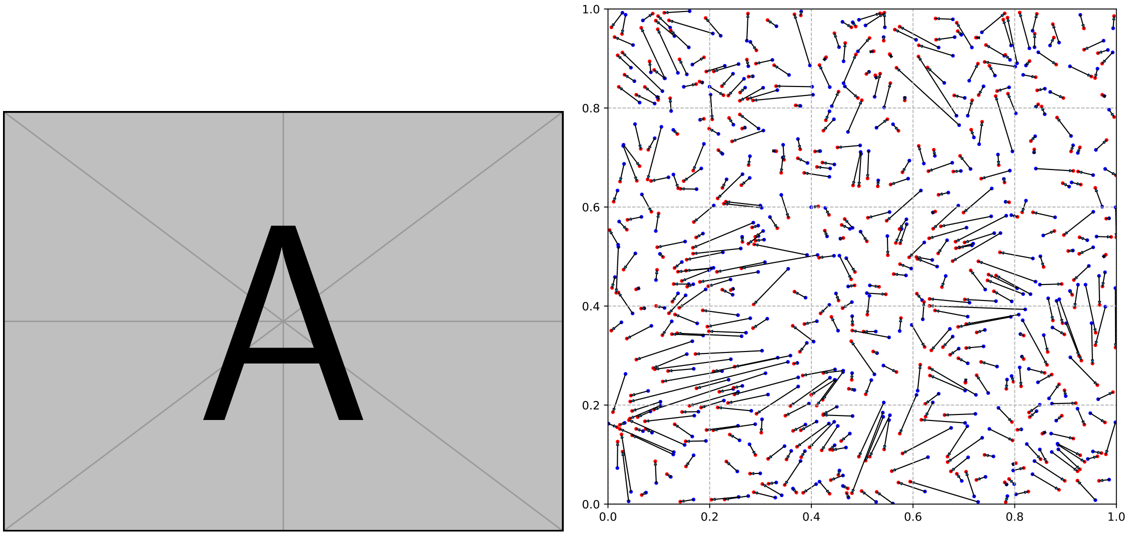

Figure 1: Left: . Right: portrait a solution on $\mathcal{M} = [0,1]^2$ with $p = 1$ and $n = 512$ blue and red points.

Model. Given a coordinate reference system, we wish to match the observation corresponding to of

$n \gg 1$ indistinguishable objects in two different satellite images. The n points are modeled in the first image I_1 by a family of blue points, $\mathcal{B} = \{b_i\}_{i=1}^n$. The same n objects seen in the second image I_2 are modeled by red points $\mathcal{R} = \{r_i\}_{i=1}^n$. $\mathcal{B}$ and $\mathcal{R}$ are two independent $\text{Binom}(\nu, n)$ point processes, where ν is a given intensity measure defined over a Polish metric space $(\mathcal{M}, \mathcal{D})$. The assignment of b_i to r_j is mapped by the following cost function:

$$c : \begin{cases} \mathcal{M} \times \mathcal{M} \to \mathbb{R} \\ (b_i,\, r_j) \mapsto c_{ij} = c(b_i,\, r_j). \end{cases}$$

A configuration of the whole system can then be encoded by a permutation $\pi \in \mathcal{S}_n$, with the associated permutation matrix P_π. The energy of that configuration is then denoted by

$$\mathcal{H}(\pi) = \sum_{i=1}^{n} c_{i\pi(i)} = \text{Tr}(P_\pi c).$$

An optimal assignment π_{opt} minimizes the energy $\mathcal{H}_{opt} := \mathcal{H}(\pi_{opt}) = \min_{\pi \in \mathcal{S}_n} \mathcal{H}(\pi)$, which is a random variable which we call *ground state energy.* When $p \geq 1$, the ground state energy $\mathcal{H}_{opt}$ is proportional to the $p-$th power of the p-Wasserstein distance between the empirical measures of $\mathcal{B}$ and $\mathcal{R}$. This random combinatorial minimization problem is called an **Euclidean Random Assignment Problem (ERAP)** [5].

Possible research directions. Depending on the candidate background and/or preferences, we have identified five possible research directions:

- *One dimensional toy-model:* In a previous internship on this topic [17], a one dimensional toy-model has been introduced aimed at providing an efficient criterion for quasi-one dimensional motion of identical particles depending on the satellite speed and the density of points. In the limit $n \to \infty$, this model displays a phase transition and a Poissonian description at the critical scale. We wish to provide an in-depth description of this Poissonian description at the critical scale depending on both the assumptions on the law of the points and of the perturbations. This limit result would provide a statistical estimator which could be useful in applications to satellite data analysis.

- *One dimensional ERAPs in the convex regime:* In one spatial dimension, as a consequence of the classical Hardy–Littlewood–Pólya rearrangement inequality, exact formulas for $\mathbb{E}[\mathcal{H}_{opt}]$ for any value of $n \geq 1$ are well-known when $\mathcal{M} = [0, 1]$ and $c(x, y) = |x - y|^p$, for $p \geq 1$. It would be interesting to investigate whether further exact formulas can be obtained for other choices of c, thus adressing the open question in [6, Exercise 4, Session 3].

- *Theoretical aspects on first and second moments:* It is established in [1] that $\mathbb{E}[\mathcal{H}_{opt}] \underset{n\to\infty}{\sim} \frac{1}{2\pi} \log n$ in dimension $d = 2$. We would like here to investigate further this asymptotic expansion, for example by comparing two domains Ω and Ω', via $\mathbb{E}[\mathcal{H}_{opt}^{\Omega} - \mathcal{H}_{opt}^{\Omega'}]$. Also, very little is known for the asymptotic behavior of the second moment of $\mathcal{H}_{opt}$. A possibility would be to prove upper and lower bounds on the variance of the ground state energy, $\text{Var}[\mathcal{H}_{opt}]$. This would be useful to settle the open problem of concentration of $\mathcal{H}_{opt}$ in dimension 2, which has been recently settled for $d \geq 3$ when $p \in [1, \frac{d^2}{2}[$ in [7].

- *Discrete Fourier analysis in 2D:* Considering now that $\mathcal{B}$ are nodes of a 2D grid, $\mathcal{R}$ still being a binomial process, we would like to use tools from discrete Fourier analysis to study the different contributions to $\mathbb{E}[\mathcal{H}_{opt}]$. This research direction, which has been partly started in [5, Chapter 3.3], can then be linked directly to the last point.

- *Applications to satellite data acquisition and to particle suspensions in 2D:* The Hungarian method solves the ERAP in time complexity $\mathcal{O}(n^3)$ in the worst case [15]. The goal is to investigate this algorithm, along with variants (e.g. primal-dual algorithms), apply them to the ERAP on satellite data, and to exploit the connections with other related matching problems to implement alternative algorithms, such as: Optimization algorithms for stochastic matching problems, as defined in [12] (see e.g. [14]), or the classical online matching problem of [10], see e.g. [16], which could incorporate a temporal variable into the problem. The previous theoretical results will also be tested numerically on CNES data. Also, Ici un paragraphe concernant le problème avec bleus, rouges et verts, et les données issus de la collaboration avec le Wisconsin.

Keywords: ERAPs, Monge-Kantorovich problem, Hungarian method, discrete Fourier analysis.

Candidate profile: You are a motivated Master 2 student in Mathematics with a solid background in probability theory and/or optimal transport. Basic coding skills (especially in Python) are expected.

PhD thesis supervisors

Matteo D'Achille is Associate Professor at IECL (Metz site), Université de Lorraine. He is specialized in random geometry and statistical mechanics, with a PhD on ERAPs.

Jessie Levillain is currently a postdoctoral research fellow at CNES Data Campus. She is specialized in mathematical modeling and numerical analysis.

Pascal Moyal is Full Professor at IECL (Nancy site), Université de Lorraine. He has recently worked on stochastic and online matching problems on graphs.

What we offer: The successful candidate will be based at the **Institut Élie Cartan de Lorraine (IECL)** with a starting date in **October 2026**. The candidate will be mostly based at the Metz site of the Institute, and will be able to apply there for a teaching position within the UFR MIM. He/she will be encouraged to spend time in Nancy for scientific collaborations and seminars.

Application: We require:

- A detailed CV including a complete transcript of grades;
- A cover letter explaining your interest in the project;
- A letter from the director of master studies assessing the candidate's ranking within their graduating class particularly in terms of the percentage of students who go on to pursue a doctoral dissertation;
- The names of one or two permanent professors who are willing to write a recommendation letter.

The application material should be sent via email to the three above emails before **May 29th 2026**.

References

[1] L. Ambrosio, F. Stra, D. Trevisan, *A PDE approach to a 2-dimensional matching problem*, Probab. Theory Relat. Fields, 173(1), 2019.

[2] R. Cargnello, *Hybridization of multi-sensor and multi-temporal data for monitoring water resources*, internship report, CNES, 2024.

[3] M. Chertkov, L. Kroc, F. Krzakala, M. Vergassola, and L. Zdeborová, *Inference in particle tracking experiments by passing messages between images*, Proc. Natl. Acad. Sci. 107 (17) 7663-7668, 2010.

[4] Scharstein, D. and Szeliski, R. and Zabih, R, *A taxonomy and evaluation of dense two-frame stereo correspondence algorithms*, Electron. IEEE SMBV, 2001.

[5] **M. D'Achille**, *Statistical Properties of the Euclidean Random Assignment Problem*, PhD thesis, Paris-Saclay University, 2020. https://hal.archives-ouvertes.fr/tel-03098672

[6] **M. D'Achille**, *An invitation to random assignment problems*, AESIM-CIMPA Summer School "Probability with / on random graphs: trees, networks, dimers, assignments", Nesin Mathematics Village, Sirinçe, Turkey, August 2024. https://matteodachille.github.io/teaching

[7] **M. D'Achille**, F. Mattesini, and D. Trevisan, *Concentration for random Euclidean combinatorial optimization*, arXiv/math.PR 2602.21851, 2026. Available at https://arxiv.org/abs/2602.21851

[8] N. Gasnier, *Use of multi-temporal and multi-sensor data for continental water body extraction in the context of the SWOT mission*, PhD thesis, Institut Polytechnique de Paris, 2022. https://theses.hal.science/tel-03578831

[9] S. Geman and D. Geman, *Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images*, IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-6, no. 6, pp. 721-741, 1984.

[10] R. M. Karp, U. V. Vazirani, and V. V. Vazirani, *An optimal algorithm for on-line bipartite matching*, in *Proceedings of the Twenty-Second Annual ACM Symposium on Theory of Computing*, pp. 352–358, 1990.

[11] J. D. Lafferty, A.McCallum, and F. C. N. Pereira, *Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data*, In: Proceedings of the Eighteenth International Conference on Machine Learning. ICML '01. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., pp. 282–289, 2001.

[12] J. Mairesse and **P. Moyal**, *Stability of the stochastic matching model*, *Journal of Applied Probability*, vol. 53, no. 4, pp. 1064–1077, 2016.

[13] M. Moharrami, C. Moore, J. Xu, *The planted matching problem: Phase transitions and exact results*, Ann. Appl. Probab. 31(6), 2663-2720, 2021.

[14] M. Nazari and A. L. Stolyar, *Reward maximization in general dynamic matching systems*, *Queueing Systems*, vol. 91, no. 1, pp. 143–170, 2019.

[15] J. Munkres, *Algorithms for the Assignment and Transportation Problems*. SIAM J. Appl. Math. 5, 32–38, 1957.

[16] C. Papadimitriou, T. Pollner, A. Saberi, and D. Wajc, *Online stochastic max-weight bipartite matching: Beyond prophet inequalities*, in *Proceedings of the 22nd ACM Conference on Economics and Computation*, pp. 763–764, 2021.

[17] T. Quarck, *Random assignment problems for satellite data acquisition*, M2 internship report, Sorbonne Université, 2025