Master 2 Research internship proposal

The Euclidean Random Assignment Problem: Theory and Application to Earth Observation

Matteo D'ACHILLE & Nicolas GASNIER & Jessie LEVILLAIN & Pascal MOYAL

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

December 11, 2025

The assignment between two sets of points is a recurring problem in satellite image processing. For example, 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.

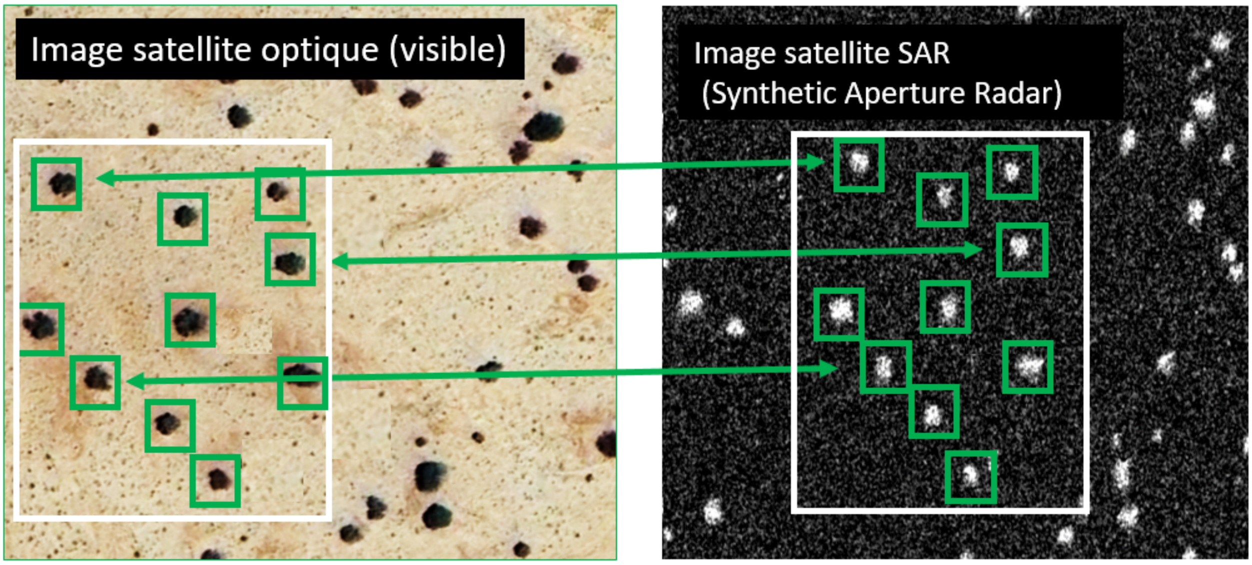

Figure 1: Optical (left) and Synthetic Aperture Radar image (right) satellite images. The shrubs (highlighted by green boxes) are indistinguishable but can be matched by minimizing the total distance among pairs in a well-chosen coordinate system

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 independent $\mathrm{Binom}(\nu, n)$ 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

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

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-Wasserstein distance to the power p between the empirical measures of $\mathcal{B}$ and $\mathcal{R}$. This minimization problem is called an Euclidean Random Assignment Problem (ERAP).

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

- *One dimensional toy-model:* In a previous internship on this topic [15], 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, including the cycle decomposition of the optimal permutation.

- *Theoretical aspects on first and second moments:* It is established in [1] that $\mathbb{E}[\mathcal{H}_{opt}] \sim \frac{1}{2\pi} \log n$ in dimension $d = 2$. We would like here to investigate further this asymptotic expansion, for example by comparing two domains Ω and Ω', via $\mathbb{E}[\mathcal{H}_{opt}^{\Omega} - \mathcal{H}_{opt}^{\Omega'}]$. Another possibility would be to prove upper and lower bounds on the variance of the ground state energy, $\mathrm{Var}[\mathcal{H}_{opt}]$.

- *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 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.

- *Application to satellite data acquisition:* The Hungarian method solves the ERAP in time complexity $\mathcal{O}(n^3)$ in the worst case [13]. 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 [10] (see e.g. [12]), or the classical online matching problem of [8], see e.g. [14], which could incorporate a temporal variable into the problem. The previous theoretical results will also be tested numerically on CNES data.

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; no prior knowledge of random geometry, hydrology, or EO is required.

What we offer: The project can last **between 4** and **6 months** and is hosted by the Institut Élie Cartan de Lorraine, with a flexible starting date in **April 2025**. It will be possible to visit the CNES headquarters in Toulouse, depending on the positive outcome of a background check. A pursuit as a PhD student on the subject can also be envisioned.

Application: We require a detailed CV (including a complete transcript of grades), a cover letter explaining your interest in the project, and 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 **January 31st 2026**.

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

Nicolas Gasnier is an image and signal processing engineer at CNES.

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

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

References

[1] L. Ambrosio, M. Goldman, D. Trevisan, On the quadratic random matching problem in two-dimensional domains, Electron. J. Probab. 27, 1-35, 2022.

[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] **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

[7] 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.

[8] 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.

[9] 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.

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

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

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

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

[14] 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.

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