Optimal Transport, Latent Representations and Generative Models
PhD Project

This PhD project will be carried out within PR[AI]RIE-PSAI.

Host laboratory: DI ENS (DATA team), within PSL
Supervisor: Kimia Nadjahi (CNRS)

This PhD project studies **modern generative models** as **dynamical (optimal) transport**, with a particular focus on **latent low-dimensional representations**, **geodesics in the space of probability measures**, and **statistical analysis**. More specifically, it focuses on the theoretical analysis of **diffusion models** [7] and **flow matching** [1], which consist in learning a transport between a source distribution and the data distribution. The goal is to identify what makes this transport easier or harder to learn and to simulate, and to derive principled guidelines for the design of more accurate and efficient generative models.

1 Context

State-of-the-art generative models rely on learning a dynamics that progressively connects a source distribution, typically easy to sample from (often a standard Gaussian), to the data distribution. In diffusion models, this dynamics is described through a noising and denoising process, and training targets the score of intermediate distributions. In flow matching, the objective is to learn a *velocity field* associated with a family of probability paths connecting the source and target distributions. These approaches thus rely on the same strategy: they learn a dynamical transport in high dimension, then integrate it numerically in order to generate samples.

The difficulty of this problem depends on several modeling choices. These components are still mostly chosen based on empirical benchmarks, and their joint theoretical analysis is largely open.

Choice of data representation. To reduce computational cost, generative models are often trained not in the original data space, but in a lower-dimensional latent space learned by an autoencoder [5]. This strategy changes the geometry of the problem: it affects the geometry of intermediate distributions, the regularity of the score or velocity field, and the resulting approximation and discretization errors.

Choice of the path between source and target distributions. The learned dynamics follows a path in the space of probability measures between the source distribution (p_0) and the data distribution (p_1). This path is not unique. In particular, flow matching allows for infinitely many possible interpolations (*stochastic interpolants* [1]) through random variables of the form,

$$I_t = \alpha_t X_0 + \beta_t X_1 \,, \tag{1}$$

where $X_0 \sim p_0$, $X_1 \sim p_1$, and $(\alpha_t, \beta_t)_{t\in[0,1]}$ are deterministic functions satisfying suitable boundary conditions. The law of I_t then defines a family of intermediate distributions between p_0 and p_1. The choice of this path affects the regularity of the field to be learned, the stability of the dynamics, and the numerical cost of sampling [2, 8]. Yet, for practitioners, this choice is still largely heuristic.

Choice of parameterization. The architectures used to model the score or the velocity field define a class of functions and therefore induce a representation bias. In particular, the growing use of *Transformer* architectures in generative modeling raises the question of whether some classes of fields or some families of interpolants are better suited than others to such parametrizations.

2 Scientific objectives

The main objective of the thesis is to develop a theoretical framework to relate modeling choices (data representation, probability path, and parametrization) to two important performance criteria in generative modeling: the approximation error of the learned transport, and the numerical cost of integrating it. To this end, we will notably leverage the mathematical framework of **optimal transport**.

2.1 Axis 1: Latent representations and transport

A first objective is to analyse the effect of data representation, especially when it is defined by an encoder-decoder pair (g_θ, f_ω). A generative dynamics defined in data space induces, through the encoder g_θ, a dynamics in latent space whose geometric properties may be significantly different. The goal is to identify measurable quantities (e.g., involving the Jacobian of g_θ, geometric distortion, local regularity, or low-dimensional structure) that can be related to bounds on approximation and discretization errors. This aims to clarify when generative modeling in latent space is simpler, from both statistical and numerical considerations, than modeling in the original data space.

A key aspect of this axis is to view the autoencoder not merely as a preprocessing tool optimized for reconstruction and then frozen, but as a component that shapes the transport problem itself. Therefore, latent representations are connected to questions of compression, low-dimensional structure, and generalization.

2.2 Axis 2: Probability paths, optimal transport, and latent geometry

A second, important axis will study the influence of the path chosen between source and target distributions, and then the effect of changing representation on this analysis. In flow matching, $(\alpha_t, \beta_t)_{t\in[0,1]}$ (eq. (1)) determine a specific family of interpolants ; in diffusion models, analogous quantities characterize the *noise schedule*. The first goal will be to identify, for criteria such as field regularity or integration error, which families of coefficients are the most favourable.

Recent analyses show that the Lipschitz regularity of exact fields plays a central role in controling trajectory stability and discretization error in terms of the **Wasserstein distance** [2, 8]. Building on their results and underlying assumptions, the thesis will compare different concrete classes of paths and study how they affect the difficulty of learning and sampling.

These conclusions will also be compared to the **dynamic formulation of optimal transport** (OT) by Benamou-Brenier. Although OT is too costly to be used directly in high dimension, it provides a rich and relevant mathematical framework for comparing distributions. It has already yielded some insights into *rectified flow* dynamics in simplified settings [4]. Our goal is to extend this line of work by assessing to what extent the favorable trajectories in diffusion or flow matching (i.e., the coefficients $(\alpha_t, \beta_t)_{t\in[0,1]}$ yielding regular fields or minimal discretization error) depart from the OT paths, and with what consequences for approximation and numerical integration. This analysis will first be carried out in tractable settings, e.g., Gaussian data, and then extended to latent settings, for instance under **low-rank assumptions**.

2.3 Axis 3: Neural parametrization and implicit regularization

A third, more exploratory axis concerns the interaction between the chosen path, the parametrization of the score or velocity field, and the training procedure. One objective will be to investigate whether the regularity of the learned field can be explained as a form of implicit regularization induced by the optimization algorithm, the structure of the training objective, or the function class itself. Another objective will focus on specific parametrizations, in particular Transformer-based architectures. The goal will be to identify the constraints induced by such architectures on the class of representable fields, for instance when the velocity field is parametrized by a single attention layer, and to analyze how these constraints interact with the choice of path. To this end, we will build on the measure-theoretic viewpoint on attention mechanisms that has recently been introduced and studied [6, 3].

2.4 Timeline

This PhD project is structured over three years. **Year 1** will focus on the necessary background and on first theoretical results for latent representations in tractable settings (e.g., Gaussian and low-rank models), with initial numerical experiments. **Year 2** will be devoted to the analysis of favorable paths with comparisons with optimal transport and extensions to latent-space settings. **Year 3** will concentrate on neural parametrization and implicit regularization, while consolidating the theoretical and numerical results into a coherent framework. Overall, the thesis aims to deliver theoretical guidelines on data representation, probability paths, and parametrization in modern generative models, supported by numerical experiments and open-source code.

3 Application procedure

Candidate profile. The project requires a strong background in machine learning or applied mathematics, with good foundations in probability, optimization, and linear algebra. Prior exposure to optimal transport, deep learning, or generative modeling is a strong asset. Programming skills in Python and modern machine learning libraries (PyTorch, JAX) are expected. The project combines mathematical analysis with numerical experiments, so the candidate should be comfortable navigating between theory and implementation.

How to apply? Interested candidates should send the following application materials **by May 18, 2026** to kimia.nadjahi@ens.fr:

- a curriculum vitae;
- a one-page motivation letter describing their interest in the proposed topic and explaining why their background is relevant to the project;
- copies of diplomas and academic transcripts.

Results of the selection process will be communicated in **two phases** between **May 30–mid-June 2026**.

Non-discrimination, openness, and transparency. All PR[AI]RIE-PSAI partners are committed to supporting and promoting equality, diversity, and inclusion within their communities. We welcome applications from a wide range of backgrounds, and we are committed to ensuring an open and transparent recruitment process.

References

[1] Michael Albergo, Nicholas M. Boffi, and Eric Vanden-Eijnden. Stochastic interpolants: A unifying framework for flows and diffusions. *Journal of Machine Learning Research*, 26(209):1–80, 2025.

[2] Eliot Beyler and Francis Bach. Convergence of Deterministic and Stochastic Diffusion-Model Samplers: A Simple Analysis in Wasserstein Distance. 2025.

[3] Etienne Boursier and Claire Boyer. Softmax as linear attention in the large-prompt regime: a measure-based perspective, 2025.

[4] Hertrich, Chambolle, and Delon. On the relation between rectified flows and optimal transport. In *Proceedings of NeurIPS'25*, 2025.

[5] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVPR*, pages 10684–10695, June 2022.

[6] Michael E. Sander, Pierre Ablin, Mathieu Blondel, and Gabriel Peyré. Sinkformers: Transformers with doubly stochastic attention. In *Proceedings of AISTATS*, volume 151 of *Proceedings of Machine Learning Research*, 2022.

[7] Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *ICLR*, 2021.

[8] Arthur Stéphanovitch. Lipschitz regularity in flow matching and diffusion models: sharp sampling rates and functional inequalities, 2026.