Tessellations of Semi-Discrete Flow Matching

Emile Pierret[*]
DMA, ENS-PSL, MAP5
pierret@math.cnrs.fr

Johannes Hertrich[*]
University of Göttingen
johannes.hertrich@uni-goettingen.de

Samuel Hurault[*]
LIGM, Univ. Gustave Eiffel, CNRS
samuel.hurault@univ-eiffel.fr

Julie Delon
DMA, ENS-PSL
julie.delon@ens.fr

Abstract

We study Flow Matching in a semi-discrete setting where a Gaussian source is transported toward a discrete target supported on finitely many points. This semi-discrete regime is the theoretical setting behind the use of Flow Matching for generative modeling, where the target distribution is represented by a finite dataset. In this semi-discrete regime, the exact Flow Matching velocity field is available in closed form, which makes it possible to analyze the geometry induced by the terminal flow map independently of optimization and approximation effects. We investigate the terminal assignment regions, namely the preimages of the target atoms under the terminal flow. We show that these regions are open, simply connected and, under an additional assumption, homeomorphic to the unit ball. At the same time, a planar four-point example shows that these cells can differ sharply from Laguerre cells arising in semi-discrete optimal transport: they may be non-convex, have curved boundaries, and exhibit different boundedness and adjacency patterns. These results clarify the geometry intrinsically induced by the exact semi-discrete Flow Matching objective before neural approximation enters the picture.

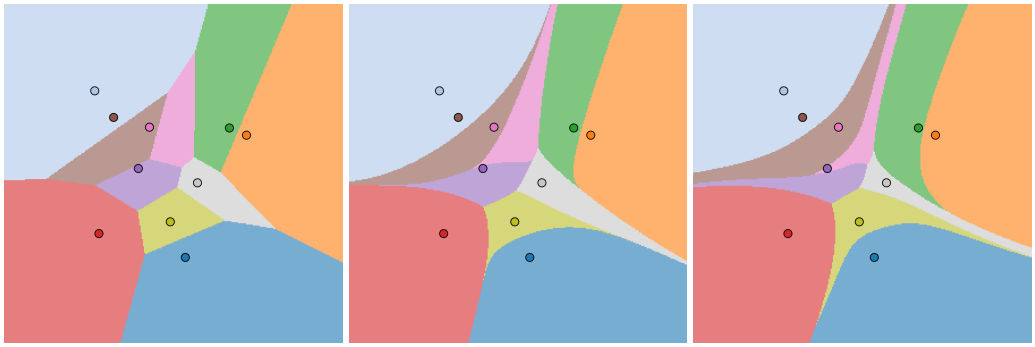
1 Introduction

Flow Matching (FM) and closely related regression-based formulations of generative transport appeared almost simultaneously in 2022 in three works: Albergo and Vanden-Eijnden [2023], Lipman et al. [2023], Liu et al. [2023]. These papers share the core idea of learning a continuous-time transport by regressing a velocity field along a prescribed interpolation path, thereby avoiding backpropagation through an ODE solver during training. Since then, FM has emerged as a powerful framework for training continuous normalizing flows and as a useful conceptual bridge between diffusion-type models and deterministic generative flows Lipman et al. [2024]. In particular, several recent works have emphasized the role of rectified flows and their relation to transport efficiency and straightness of trajectories Liu et al. [2023], Bansal et al. [2025], Tsimpos et al. [2026], while the precise connection between Flow Matching and optimal transport (OT) has only started to be clarified more recently Hertrich et al. [2026].

Among the probability paths that can be used in FM, OT-inspired paths play a particularly important role. They often lead to faster sampling and improved training behavior Lipman et al. [2023], Liu [2022], Liu et al. [2023], Zhang et al. [2025], Pooladian et al. [2023], Tong et al. [2024], Boïté et al. [2026], Chambolle and Hertrich [2026]. At the same time, recent results show that the relation between FM or rectified flows and OT is more subtle than early analogies might suggest, and

[*]Equal contribution.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).



(a) OT (b) FM with exact velocity (c) FM with learned velocity

Figure 1: Semi-discrete assignment cells Γ_k for 10 points in dimension 2, for (a) Optimal Transport, (b) Flow Matching with closed-form velocity and (c) Flow Matching with velocity approximated by a neural network. The colored regions partition the Gaussian source space into assignment cells, while the dots mark the target atoms a_k. Matching colors indicate the corresponding assignments. In (b), these regions are the FM cells Γ_k. In (c), each grid point is assigned the color of the target point closest to its image under the terminal flow.

that apparent similarities can break down outside rather restrictive regimes Hertrich et al. [2026], Lavenant and Santambrogio [2022], Chambolle and Hertrich [2026]. This raises a basic question: what geometry is intrinsically induced by the exact Flow Matching vector field itself, before any neural approximation is introduced?

We address this question in a semi-discrete setting where the source is Gaussian and the target is a finitely supported measure. This regime is appealing for two reasons. First, semi-discrete OT provides a classical reference geometry through Laguerre cells, whose geometry underlies both the theory and algorithms of semi-discrete transport Aurenhammer et al. [1998], Mérigot [2011], Kitagawa et al. [2019], Peyré and Cuturi [2019]. Second, when one starts from the independent coupling, the exact velocity field of semi-discrete FM is available in closed form Bertrand et al. [2026], making the model analytically tractable while retaining genuinely nontrivial geometric behavior.

Our objects of study are the terminal assignment regions

$$\Gamma_k := \{x \in \mathbb{R}^d : \gamma(x) = a_k\}, \tag{1}$$

where γ is the terminal flow map and $(a_k)_{1\le k\le n}$ are the support points of the target measure. These sets are the exact FM analogue of semi-discrete transport cells. A natural first guess would be that they inherit at least part of the usual OT geometry, for instance convexity, affine boundaries, or adjacency patterns comparable to those of Laguerre diagrams. Figure 1 already suggests that this intuition is misleading. While FM cells appear topologically simple (connected and without holes), they can be visibly non-convex, separated by curved boundaries and organized in a way that does not completely match the boundedness or neighborhood structure of the corresponding Laguerre cells.

Contributions In this paper, we make the observations from Figure 1 formal. First, we present some simplification results regarding the dynamics of semi-discrete FM. Second, as our first main contribution, we establish global topological properties of the FM cells, showing that the sets Γ_k are open and simply connected, and under an additional assumption we prove that they are homeomorphic to the unit ball. Third, for our second main contribution, we exhibit a simple planar four-point example showing that while FM cells may initially look like Laguerre cells, they can differ qualitatively from them, inducing a genuinely different geometry. Specifically, we show that they may be non-convex, have curved boundaries, and display different boundedness and adjacency patterns.

This perspective is relevant for generative modeling even though practical neural FM models only learn an approximation of the exact field studied here. In discrete or nearly discrete regimes these approximations may smooth, relax, or altogether blur the hard terminal assignments present in the exact dynamics. Our point is therefore not that trained networks should literally reproduce the cells Γ_k, although the right part of Figure 1 suggests that these cells are still visible in the trained velocity field, at least in small dimension. Rather, these cells reveal the canonical geometric bias encoded in

the FM objective before parameterization, optimization, and approximation effects enter the picture. Understanding this reference geometry is a first step toward disentangling what is intrinsic to the objective itself from what is induced by approximation, architecture, or optimization.

2 Background on Flow Matching and Semi-discrete Optimal Transport

Basics of Flow Matching We briefly recall the Flow Matching (FM) principle before specializing to the discrete target setting considered in this paper. Flow Matching connects a source distribution μ_0 and a target distribution μ_1, using a continuous sequence of probability measures $(\mu_t)_{t\in[0,1]}$. FM proposes a time-dependent velocity field v_t such that the associated continuity equation transports μ_0 into μ_1, or equivalently such that the flow map generated by the ODE

$$\dot{\gamma}_t(x) = v_t(\gamma_t(x)), \qquad \gamma_0(x) = x, \tag{2}$$

satisfies $(\gamma_t)_{\#}\mu_0 = \mu_t$ (provided that it exists and is unique).

A key feature of FM is that one does not need to integrate the ODE during training. Instead, one regresses the velocity field against a target conditional velocity induced by a prescribed family of conditional paths Albergo and Vanden-Eijnden [2023], Lipman et al. [2023, 2024], Liu et al. [2023]. More precisely, if (X_0, X_1) is a coupling between μ_0 and μ_1 and if $X_t = (1-t)X_0 + tX_1$, then the target FM velocity is given by the conditional expectation

$$v_t^\star(x) = \mathbb{E}\left[X_1 - X_0 \mid X_t = x\right].$$

This L^2 conditional expectation is naturally characterized as the minimizer of a least-squares regression problem. In other words, one can train a parametric field v_t^θ by regressing it against the conditional path derivative $X_1 - X_0$. Observe that in the case where $\mu_0 \sim \mathcal{N}(0, I)$ and where $(X_0, X_1) \sim \mu_0 \otimes \mu_1$ is the independent coupling, existence and uniqueness of the flow map for the velocity $v_t^\star$ are easy to show (see Appendix A). Further, the flow map γ_t is a bi-Lipschitz diffeomorphism for $0 \le t < 1$, even though it is not invertible for $t = 1$.

Link to Optimal Transport. For the quadratic cost, Monge optimal transport is classically formulated as the optimization problem

$$\inf_{T_{\#}\mu_0 = \mu_1} \int_{\mathbb{R}^d} \|x - T(x)\|^2 \, d\mu_0(x).$$

When an optimal map T^* exists (which holds as soon as μ_0 has a density for instance), it induces a canonical path $\mu_t := ((1-t)\mathrm{Id} + tT^*)\,\#\mu_0$ between μ_0 and μ_1. This is the basic intuition behind the link with Flow Matching: both frameworks describe transport through time-dependent probability paths and velocity fields, but they optimize different objects. In particular, Flow Matching looks for a velocity field associated with a prescribed interpolation $(\mu_t)_{t\in[0,1]}$, whereas OT selects directly the map minimizing the transport cost. The two viewpoints coincide in special cases, for instance in one dimension or for suitable Gaussian settings (for independent X_0 and X_1 with commuting covariances), but not in general Hertrich et al. [2026], Liu [2022], Pierret et al. [2026].

Semi-discrete case. We focus on the specific case where (X_0, X_1) is an independent coupling with $X_0 \sim \mu_0 := \mathcal{N}(0, \mathrm{Id})$ and $X_1 \sim \mu_1 := \frac{1}{n}\sum_{k=1}^n \delta_{a_k}$, with pairwise distinct support points $a_k \neq a_l$ for $k \neq l$. In this situation, both the optimal transport problem and the flow matching algorithm admit more tractable formulations.

The optimal transport map (with quadratic cost) between μ_0 and μ_1 is characterized by so-called Laguerre cells: there exists a vector of weights $\psi = (\psi_k)_{1\le k\le n}$ such that it holds $T(x) = a_k$ for $x \in \mathcal{L}_k(\psi)$ almost everywhere, where

$$\mathcal{L}_k = \mathcal{L}_k(\psi) := \left\{ x \in \mathbb{R}^d : \ \frac{1}{2}\|x - a_k\|^2 - \psi_k \le \frac{1}{2}\|x - a_j\|^2 - \psi_j, \ \forall j \right\}. \tag{3}$$

These cells form a polyhedral partition of the source space. When all weights are equal, for instance when $\psi_k = 0$ for every k, they are also known as Voronoi cells Aurenhammer et al. [1998], Mérigot [2011], Kitagawa et al. [2019], Peyré and Cuturi [2019].

For flow matching, the semi-discrete setting allows the derivation of a closed formula for the velocity field given by

$$v_t(x) := \sum_{k=1}^{n} \alpha_k(t,x) \frac{a_k - x}{1-t}, \quad \alpha(t,x) := \mathrm{softmax}\left(-\frac{1}{2(1-t)^2} \|x - t a_i\|^2 \right)_{1 \le i \le n}.$$

The velocity is a convex combination of the elementary drifts $(a_k - x)/(1-t)$ pointing toward the support points. The coefficients $\alpha_k(t,x)$ can be interpreted as posterior assignment weights: at time t and position x, they measure how likely the endpoint a_k is under the conditional interpolation model. We denote by $\gamma_t \colon \mathbb{R}^d \to \mathbb{R}^d$ the flow map generated by this field, defined by (2), and **we write $\gamma = \gamma_1$ for the terminal map.** Our work studies the topology and geometry induced by this terminal flow map γ. Since $\gamma_{\#}\mu_0 = \mu_1$, the definition of the push-forward measure gives $\mu_0(\Gamma_k) = \mu_0(\gamma^{-1}(\{a_k\})) = \mu_1(\{a_k\}) = 1/n$ for every k. Thus each cell carries exactly the same Gaussian probability mass, and $\mu_0(\mathbb{R}^d \setminus \bigcup_{k=1}^n \Gamma_k) = 0$. Since the Gaussian density is strictly positive everywhere, this complement also has Lebesgue measure zero.

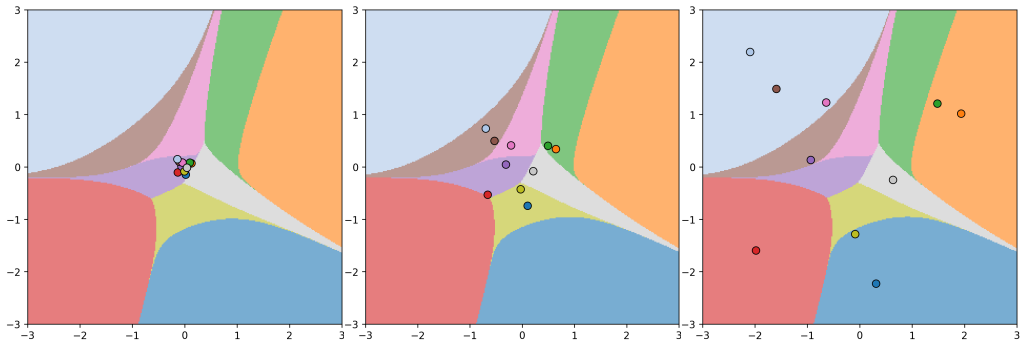
3 Some Simplification Results

Before presenting our main results, we first explain how several related problems can be reduced to the framework studied in this paper. Specifically, we establish three auxiliary statements. First, we observe that scaling, translating, or rotating the cells under consideration leaves our problem unchanged. Second, we prove that it suffices to restrict our analysis to the affine hull of the points a_k, and that the dynamics outside this hull are trivial. Third, we demonstrate that discrete FM arises as a limiting case of Gaussian mixture targets when the variances of the modes tend to zero.

Invariance under Scalings, Shifts and Rotations We note that the cells are invariant under scalings and shifts and equivariant under orthogonal transformations. More precisely, we have the following corollary, which directly follows from [Hertrich et al., 2026, Thm 2].

Corollary 1. *Let $A \in \mathbb{R}^{d\times d}$ be an orthogonal matrix, $b \in \mathbb{R}^d$ and $c > 0$. Let γ be the flow map for terminal points $a_1, ..., a_n \in \mathbb{R}^d$ and $\tilde{\gamma}$ be the flow map for terminal points $\tilde{a}_1, ..., \tilde{a}_n$ defined by $\tilde{a}_k = cAa_k + b$. Then it holds $\tilde{\gamma}(x) = cA\gamma(A^{-1}x) + b$ and $\tilde{\Gamma}_k = \{Ax : x \in \Gamma_k\}$, where Γ_k and $\tilde{\Gamma}_k$ are the assignment regions (1) corresponding to γ and $\tilde{\gamma}$.*

This corollary is illustrated in Figure 2. In particular, as in the case of Laguerre cells in optimal transport, a point may fail to belong to its associated cell.



(a) $c = 0.1$ (b) $c = 0.5$ (c) $c = 1.5$

Figure 2: Illustration of Corollary 1. For $n = 10$ points $(a_k)_{1\le k\le n}$, we display the associated semi-discrete flow matching cells of $(ca_k)_{1\le k\le n}$ for $c = 0.1, 0.5$, and 1.5, which coincide. This also shows that a point may or may not belong to its associated cell.

Reduction to the Affine Hull Next, we show that the FM dynamics can be completely reduced to the dynamics on the affine hull of the points a_k. More precisely, let

$$E := \mathrm{Aff}(a_k)_{1\le k\le n} = a_1 + \mathrm{span}\{a_k - a_1 : k = 2, ..., n\},$$

and decompose $x = x^{\parallel} + x^{\perp} \in \mathbb{R}^d$ into the orthogonal projection $x^{\parallel} \in E$ and its complement $x^{\perp} = x - x^{\parallel}$. Now the next proposition shows that the dynamics $\gamma_t(x)$ can be decomposed into the dynamics of $x^{\parallel}$ and a linear shift towards the affine space E. The proof is given in Section B.1.

Proposition 2. *For every $t < 1$ and $x = x^{\parallel} + x^{\perp} \in \mathbb{R}^d$, one has $\alpha_k(t, x) = \alpha_k(t, x^{\parallel})$ for all k and $\gamma_t(x) = \gamma_t(x^{\parallel}) + (1-t)x^{\perp}$.*

This reduction shows in particular that genuinely multidimensional phenomena can only arise when the support points span a space of dimension at least two. In the collinear case, the problem reduces to one dimension, which explains why the comparison with semi-discrete OT is especially natural in that regime Hertrich et al. [2026], Pierret et al. [2026].

Semi-discrete Flow Matching as Limit of Gaussian Mixtures Next, we show that semi-discrete FM can be understood as the limit case of FM with Gaussian mixture target for vanishing variance. More precisely, let $\mu_1^{\varepsilon} = \frac{1}{n}\sum_i \mathcal{N}(a_i, \varepsilon^2 I)$ and denote the corresponding velocity field from the FM dynamics by

$$v_t^{\varepsilon}(x) := \sum_{k=1}^{n} \alpha_k^{\varepsilon}(t,x) \frac{(1-t)(a_k - x) + t\varepsilon^2 x}{(1-t)^2 + t^2\varepsilon^2}, \quad \alpha^{\varepsilon}(t,x) := \operatorname{softmax}\left(-\frac{\|x - t a_i\|^2}{2\big((1-t)^2 + t^2\varepsilon^2\big)} \right)_{1 \le i \le n}.$$

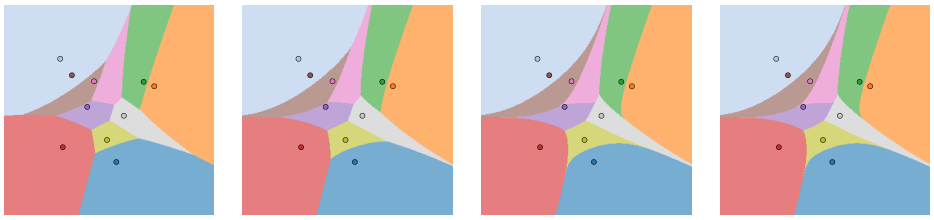
We denote by $\gamma_t^{\varepsilon} \colon \mathbb{R}^d \to \mathbb{R}^d$ the flow map defined by $\gamma_0^{\varepsilon}(x) = x$ and $\dot{\gamma}_t^{\varepsilon}(x) = v_t^{\varepsilon}(\gamma_t^{\varepsilon}(x))$ and use the notation $\gamma^{\varepsilon} = \gamma_1^{\varepsilon}$. For $\varepsilon = 0$, we obtain the velocity $v_t = v_t^0$ and flow map $\gamma_t = \gamma_t^0$ of our semi-discrete case. We then have the following pointwise convergence result for γ^{ε} toward γ as $\varepsilon \to 0$, valid inside the limiting cells. The proof is given in Section B.2.

Proposition 3. *Let $x \in \cup_k \Gamma_k$. For all $\delta > 0$, there is $\varepsilon_0 > 0$ and $T \in (0,1)$ such that for all $\varepsilon < \varepsilon_0$ and $t \in [T, 1]$ it holds $\|\gamma_t^{\varepsilon}(x) - \gamma_1(x)\| \le \delta$. In particular, for $t = 1$, $\gamma^{\varepsilon}(x) \to \gamma(x)$ as $\varepsilon \to 0$.*

We illustrate in Figure 3 the convergence from the proposition by considering the generalized assignment regions

$$\Gamma_k^{\varepsilon} = \{x \in \mathbb{R}^d : \|\gamma^{\varepsilon}(x) - a_k\| < \|\gamma^{\varepsilon}(x) - a_l\| \text{ for } l \neq k\}. \tag{4}$$

For $\varepsilon = 0$, these regions coincide with the assignment regions Γ_k from (1) and the above Proposition implies, for every $x \in \bigcup_{j=1}^n \Gamma_j$, $\mathbf{1}_{\Gamma_k^{\varepsilon}}(x) \to \mathbf{1}_{\Gamma_k}(x)$ (see Corollary 15 in section B.2). For $\varepsilon > 0$, we find that γ^{ε} is a bi-Lipschitz diffeomorphism such that the Γ_k^{ε} and the corresponding neighborhood relations are topologically equivalent to the Voronoi cells.



(a) $\varepsilon = 0.75$ (b) $\varepsilon = 0.5$ (c) $\varepsilon = 0.2$ (d) $\varepsilon = 0$

Figure 3: We plot the generalized assignment regions (4) for different values of ϵ. According to Proposition 3 these regions approximate the Γ_k from (1) as $\epsilon \to 0$.

4 Topology of the Flow Matching Cells

We now turn our attention to the topology of the assignment regions Γ_k. Informally, these regions specify which point in the source distribution is associated with which point in the target distribution. In this sense, the regions Γ_k play a role analogous to Laguerre cells, except that the optimal transport coupling is replaced by the coupling induced by flow matching. For this reason, we will also refer to the Γ_k as *flow matching cells*.

Although we have already observed numerically in the introduction (Figure 1) that the cells Γ_k do not share the neighborhood relations and boundedness properties of the Laguerre cells, we demonstrate

in this section that, from a topological perspective, they are nevertheless well-behaved: each Γ_k is an open, connected set without holes. Under an additional assumption, we will further prove that these cells are in fact topologically equivalent (homeomorphic) to the unit ball. The key ingredient for establishing these results is a late-time capture and stability result which we describe in the next subsection.

4.1 Late-Time Capture and Stability

Given that the flow map γ_t is a bi-Lipschitz diffeomorphism for $0 \leq t < 1$ but not even invertible for time $t = 1$, it will be important to understand in which sense $\gamma = \gamma_1$ is approximated by the maps γ_t for $0 \leq t < 1$. While it is directly clear from the existence of the flow map (see also Appendix A) that γ_t converges to γ almost everywhere, the next proposition shows a stronger convergence result. More precisely, we show that any trajectory $t \mapsto \gamma_t(x)$ which enters a ball around an atom a_k late enough converges to the atom.

Proposition 4. *Define $r < \frac{1}{2} \min_{k \neq l} \|a_k - a_l\|$. Then, there exists $\underline{t}_r \in (0, 1)$ such that for $x \in \mathbb{R}^d$, if there exists $\underline{t}_r < t < 1$ such that $\gamma_t(x) \in \bar{B}_r(a_k)$ then $\gamma_1(x) = a_k$.*

The proof is deferred to Section C.1 and relies on Lemma 17, which states that close enough to time 1, the velocity points inward on a small sphere around each support point. We also note that a similar result in the context of the probability flow of diffusion models was proven by Baptista et al. [2025].

4.2 Cell connectedness

Next, we prove that the flow matching cells Γ_k are open and do not admit any holes. We formalize this property by showing in the next theorem that the cells Γ_k are simply connected. Recall that $A \subseteq \mathbb{R}^d$ (viewed as a topological space inheriting the topology of $(\mathbb{R}^d, \|\cdot\|)$) is simply connected if for any loop $f\colon S^1 \to A$ we can find a continuous mapping (homotopy) $H\colon [0,1] \times S^1 \to A$ such that $H(0, x) = f(x)$ and $H(1, \cdot)$ is constant.

Theorem 5. *For $1 \leq k \leq n$, the set Γ_k is open and simply connected.*

The detailed proof is provided in Section C.2. We outline the key arguments as follows: The openness of Γ_k is an immediate consequence of the fact that it can be written as a union of open sets of the form $\gamma_t^{-1}(B_r(a_k))$, where r comes from Proposition 4 and t is sufficiently large. Using this open cover of Γ_k, we further show that for every compact subset $A \subseteq \Gamma_k$ there exists a time t_0 such that $\gamma_t(A) \subseteq B_r(a_k)$ for all $t \geq t_0$. Applying this to $f(S^1)$ yields the desired homotopy H: we first map by γ_{t_0}, then contract inside the ball $B_r(a_k)$, and finally apply the inverse of γ_{t_0}. By construction H is continuous, and by Proposition 4 it additionally satisfies $H(t, x) \in \Gamma_k$. Altogether, this shows that Γ_k is simply connected.

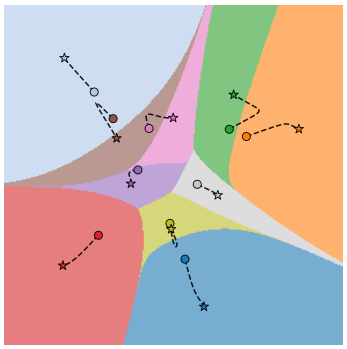


Figure 4: Illustration of the *centers* of the cells. For $n = 10$ points $(a_k)_{1 \leq k \leq n}$, we depict the associated semi-discrete flow matching cells and the corresponding *centers* $\lim_{t \to 1} \gamma_t^{-1}(a_k)$, represented by star symbols. The dotted lines indicate the curves $(\gamma_t^{-1}(a_k))_{0 \leq t < 1}$.

4.3 *Center* of the cells and contractibility

We have seen in the previous subsection that the cells Γ_k are simply connected, meaning that every loop in a cell can be continuously contracted to a point within that cell. In the usual topological terminology, this rules out "one-dimensional" holes, where the dimension refers to the loop detecting the hole. In order to prove the absence of higher dimensional holes, we use the notion of contractibility. Recall that $A \subseteq \mathbb{R}^d$ is contractible if there exists a continuous mapping (homotopy) $H\colon [0,1] \times A \to A$ such that $H(0, x) = x$ and $H(1, \cdot)$ is constant. Naturally, this property implies simply connectedness, but not vice versa. For example, the sphere $\{x \in \mathbb{R}^3 : \|x\| = 1\}$ is simply connected but not contractible.

To construct such a homotopy for the sets Γ_k, we again make use of Proposition 4. We define $H(t, \cdot)$ by first applying γ_t, then (for sufficiently large time) contracting inside the ball $B_r(a_k)$ towards a_k,

and finally applying γ_t^{-1}. Intuitively, as $t \to 1$, this converges to the point $\lim_{t\to 1} \gamma_t^{-1}(a_k)$. It is not immediately obvious that this limit exists and lies inside the cell Γ_k; this is established in the next proposition. The proof is quite technical and deferred to Section C.3. Since this limit always lies in (the closure of) the cell and we contract the cell onto it, we refer to it as the *center* of the cell.

Proposition 6. *The curve $t \mapsto \gamma_t^{-1}(a_k)$ has finite length for $t \in (0, 1)$. In particular, the limit $\lim_{t\to 1} \gamma_t^{-1}(a_k)$ exists and is located in $\overline{\Gamma_k}$.*

We plot the curves $t \mapsto \gamma_t^{-1}(a_k)$ in Figure 4. By definition, they start in a_k for $t = 0$ and end at $\lim_{t\to 1} \gamma_t^{-1}(a_k)$, which is always inside $\overline{\Gamma_k}$ even when $a_k \notin \overline{\Gamma_k}$.

Assuming that the limit $\lim_{t\to 1} \gamma_t^{-1}(a_k)$ lies strictly in the interior of the cell (and not on its boundary), we can construct a homotopy as outlined above, which shows that Γ_k is indeed contractible. In fact, under this assumption the cells turn out to be homeomorphic to $B_1(0)$, i.e., they are topologically equivalent to a ball. We summarize this in the following theorem. The explicit definition of the homotopy (together with a proof of its continuity) is given in Section C.3.

Theorem 7. *If $\lim_{t\to 1} \gamma_t^{-1}(a_k) \in \Gamma_k$, then Γ_k is contractible and homeomorphic to the open ball centered at 0 with radius 1, $B_1(0)$ (and hence also to $\mathbb{R}^d$).*

5 Geometry of the Flow Matching Cells

Next, we are interested in geometric properties of the flow matching cells Γ_k. In particular, we want to examine geometric properties, like non-convexity and boundedness, and the neighborhood relations between the cells. Our results in this section will primarily be negative. In particular, we will demonstrate on a simple example with four atoms in $\mathbb{R}^2$ that the cells are generally non-convex and that their boundedness properties and neighborhood relations sharply differ from those of the Laguerre cells.

Afterwards, we are interested in whether the flow map generated by the semi-discrete flow matching is monotone, which is of particular interest in the context of reflow Liu et al. [2023]. More precisely, Bansal et al. [2025], Tsimpos et al. [2026] present monotonicity as a sufficient criterion for the second reflow to be a straight-line flow. Also in this context we will find a negative result and show that monotonicity is violated for our four-point example as well as for a smoothed version with Gaussian mixtures.

5.1 A Four-Point Counterexample for Geometric Comparison with Semi-Discrete OT

In the following, we provide an example in which we can explicitly prove that the geometric structure of the flow matching cells differs from the Laguerre cells obtained from OT. More specifically, we study the setting of $n = 4$ points in $\mathbb{R}^2$, where the points a_k are defined by $a_k = (\cos(\frac{2k\pi}{3}), \sin(\frac{2k\pi}{3}))$ for $k = 1, 2, 3$ and $a_4 = (0, 0)$. The corresponding Laguerre cells and flow matching cells are illustrated in Figure 5 (a) and (b).

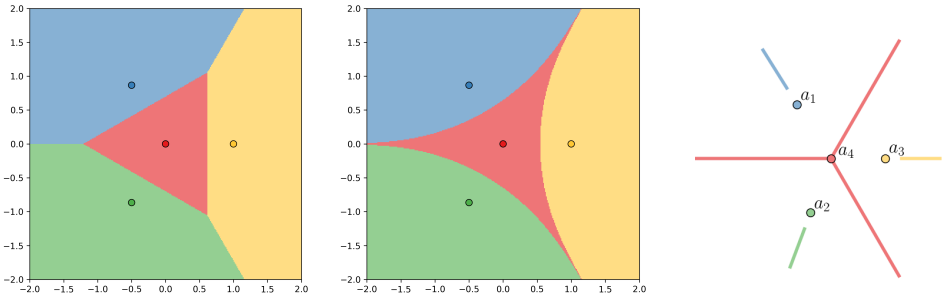
In particular, we observe the structural differences listed in the theorem below. We will formalize these differences in the remainder of this subsection.

Theorem 8. *Denote by $\Gamma_1, ..., \Gamma_n$ the flow matching cells and by $\mathcal{L}_1, ..., \mathcal{L}_n$ the Laguerre cells (3) corresponding to the atoms $a_1, ..., a_n$. Then, in our example:*

(i) There exists k such that the flow matching cell Γ_k is unbounded, but the corresponding Laguerre cell $\mathcal{L}_k$ is bounded.
(ii) There exist k and l such that the Laguerre cells $\mathcal{L}_k$ and $\mathcal{L}_l$ are neighbored (i.e., $\overline{\mathcal{L}_k} \cap \overline{\mathcal{L}_l} \neq \emptyset$), but Γ_k and Γ_l are not neighbored (i.e., $\overline{\Gamma_k} \cap \overline{\Gamma_l} = \emptyset$).
(iii) There exists k such that Γ_k is non-convex.
(iv) There exist k and l such that the boundary $\overline{\Gamma_k} \cap \overline{\Gamma_l}$ is not affine.

In order to formally show the theorem, we prove two lemmas, which determine certain subsets of the FM cells Γ_k. These subsets are illustrated in Figure 5 (c). The proofs are deferred to Section D.1. The first lemma shows that the flow map γ maps all points on the red rays in Figure 5 (c) to the central point $a_4 = (0, 0)$.

Lemma 9. *In our example, it holds $\{(-c\cos(\frac{2k\pi}{3}), -c\sin(\frac{2k\pi}{3})) : c \geq 0\} \subseteq \Gamma_4$ for $k = 1, 2, 3$.*



(a) OT (b) FM (c) Illustration of Lemmas 9 and 10

Figure 5: Illustration of the counterexample analyzed in Section 5.1. With this particular configuration of points a_k, we show the numerical computation of (a) the Optimal Transport (OT) Laguerre cells and (b) the Flow Matching (FM) cells. (c) Visualization of the lines that are proved to be contained in the corresponding cells, with each line shown in the color of its cell (see Lemma 9 and 10).

In particular, this implies that the red cell Γ_4 is unbounded for Flow Matching, whereas the corresponding Laguerre cell is bounded. It also follows that, unlike in the OT case, the cells Γ_1 (blue), Γ_2 (green), and Γ_3 (yellow) do not all share a common boundary. To further identify the geometry of the cells, we next show that for each $k = 1, 2, 3$ (blue, green and yellow cells), the half-line $\{c\, a_k : c \geq 0\}$ eventually lies entirely inside Γ_k.

Lemma 10. *In our example, there exists for $k = 1, 2, 3$ some $c_0 > 0$ such that $\{ca_k : c \geq c_0\} \subseteq \Gamma_k$.*

Combining the two lemmas above, it follows that the red cell Γ_4 is not convex and that the boundaries between Γ_4 and Γ_k, for $k = 1, 2, 3$, cannot be straight lines, see Section D.1 for a detailed derivation. Applying both lemmas we obtain the statements from Theorem 8.

5.2 Non-Monotone Flow Maps

An interesting consequence of the non-affine boundaries of the cells Γ_k is that the flow map is not monotone. Monotonicity of flow maps plays an important role in the context of the reflow algorithm Liu et al. [2023]. Reflow is an iterative procedure that repeatedly applies flow matching, using at each step the coupling $(\mu_0, \gamma_{\#}\mu_0)$ as the initial coupling, where γ denotes the flow map learned by the previous flow matching iteration. More precisely, Bansal et al. [2025] prove that reflow yields straight trajectories within the second iteration whenever the flow map in the first iteration is monotone. A similar condition was derived in Tsimpos et al. [2026].

More specifically, Bansal et al. [2025] assume that the flow map $\gamma = \gamma_1$ from the flow matching algorithm with independent coupling is differentiable and fulfills

$$\inf_x \lambda_{min} \left(J_\gamma(x) + J_\gamma(x)^\top \right) \geq 0, \tag{5}$$

or equivalently, that γ is differentiable and monotone. In the four-point example from the previous subsection, we can now observe that γ is non-monotone by the next lemma. The proof is deferred to Appendix D.2.

Lemma 11. *In the semi-discrete example of Figure 5, there are $c > 0$ and $x_1, x_2 \in \mathbb{R}^d$ such that $\langle \gamma(x_2) - \gamma(x_1), x_2 - x_1 \rangle < -c$.*

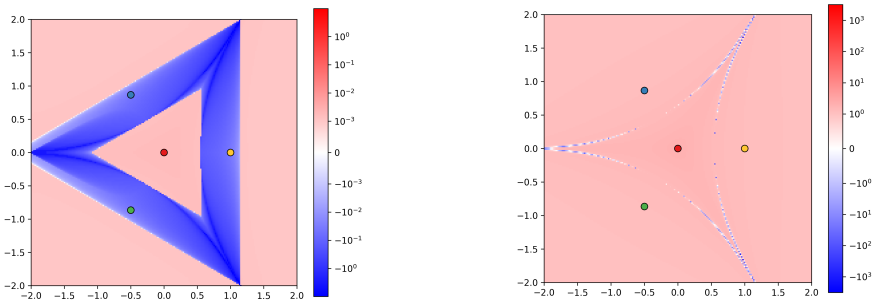
Of course our example is formulated in the semi-discrete regime, where μ_1 is not absolutely continuous (such that our setting differs from the one in Bansal et al. [2025]). However, by applying the smoothing from Proposition 3, we can derive the same result for the Gaussian mixture case. As in Section 3 we denote by γ^ϵ the flow map of the flow matching algorithm in the Gaussian mixture case, where each mode has standard deviation ϵ. Then, we obtain that the monotonicity assumption of Bansal et al. [2025] is violated. The proof can be found in Appendix D.2.

Lemma 12. *There are $\varepsilon > 0$ and $c > 0$ such that $\inf_x \lambda_{min} \left(J_{\gamma_1^\varepsilon}(x) + J_{\gamma_1^\varepsilon}(x)^\top \right) < -c$.*

More generally, one can see that in the semi-discrete setting (and thus also for the Gaussian mixture setting with sufficiently small standard deviations ε) the flow map γ can only be monotone if the

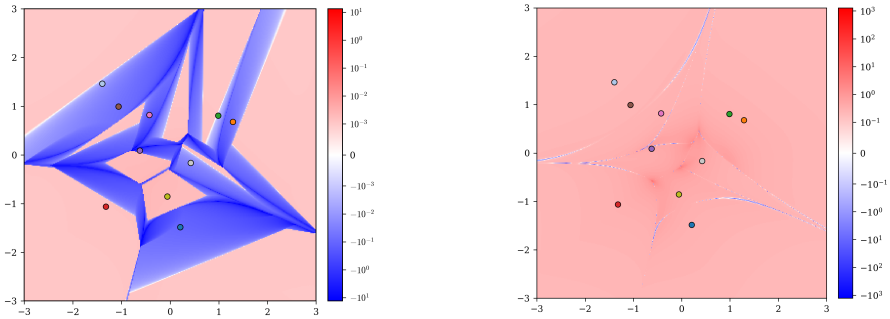
borders between the cells Γ_k are affine. While we prove non-monotonicity only for our specific four-point configuration, numerical experiments with other configurations suggest that monotonicity occurs only in very particular cases. Indeed, in Figures 6 and 7, we test condition (5) (and more generally monotonicity) for our four-point example and the ten-point configuration of Figure 1, respectively, and observe that it exactly fails around the boundaries of the cells.

Remark 13. *Although several works (Bansal et al., 2025, Tsimpos et al., 2026, Reu et al., 2026) employ monotone couplings to construct straight flows, the same authors consistently remark that monotonicity is not a necessary criterion. In fact, in our four-point example, the second reflow appears to be straight, since the regions are "star-shaped" with centers a_k: for every point in the cell $x \in \Gamma_k$, the whole segment $[x, \gamma(x) = a_k]$ remains entirely within the cell. This guarantees that the second reflow iteration has straight trajectories.*



(a) $\min_y \langle \gamma_1(y) - \gamma_1(x), y - x \rangle$ (b) $\lambda_{min}\left(J_{\gamma_1^\varepsilon}(x) + J_{\gamma_1^\varepsilon}(x)^\top\right)$

Figure 6: Illustration of the non-monotonicity of the flow map, in the four-point configuration of Section 5.1. Figure (a) illustrates non-monotonicity in the semi-discrete setting (Lemma 11), while Figure (b) illustrates Lemma 12 after Gaussian smoothing of the target atoms. In both cases, the loss of monotonicity is concentrated around the cell boundaries, suggesting that it is a geometric consequence of the non-affine FM boundaries.



(a) $\min_y \langle \gamma_1(y) - \gamma_1(x), y - x \rangle$ (b) $\lambda_{min}\left(J_{\gamma_1^\varepsilon}(x) + J_{\gamma_1^\varepsilon}(x)^\top\right)$

Figure 7: Illustration of the non-monotonicity of the flow map, in the ten-point configuration of Figure 1. Figure (a) shows non-monotonicity in the semi-discrete setting, while Figure (b) illustrates the non-monotonicity after Gaussian smoothing of the target atoms. In both cases, the loss of monotonicity is concentrated around the cell boundaries, suggesting that it is a geometric consequence of the non-affine FM boundaries.

6 Conclusion and Discussion

In this paper, we studied the geometry induced by the exact semi-discrete Flow Matching vector field, taking the viewpoint that this field is a canonical object in its own right and a useful theoretical proxy for the objectives used in generative modeling. Our main contribution is to show that the associated terminal assignment regions combine strong global topological structure with genuinely non-OT

geometry. This gives a first qualitative picture of semi-discrete FM tessellations and clarifies in which sense FM can and cannot be viewed as a surrogate for semi-discrete OT.

Limitations and future directions. One limitation of our analysis is that it concerns the *exact* semi-discrete FM vector field, whereas practical generative models only learn an approximation of it. Our results therefore do not directly characterize the geometry of the *trained* velocity field or of the transport induced by a finite neural network after optimization. The exact cells should instead be viewed as a reference geometry: they identify which structural biases are already present in the objective before approximation, architecture, and optimization enter the picture. An important next step is to understand to what extent these qualitative properties survive training, and whether the learned velocity field inherits a meaningful analogue of the exact assignment cells in discrete or nearly discrete regimes. As a first numerical indication, Appendix E reports a MNIST experiment in which a simple neural FM model produces assignment cells that closely resemble the exact closed-form FM cells on a two-dimensional PCA slice of the latent space. Nevertheless, a systematic understanding of when and how these qualitative properties survive training remains open.

Another limitation is that our results do not settle whether FM should be regarded as a useful proxy for OT from a geometric viewpoint. Our analysis shows that exact semi-discrete FM departs from Laguerre geometry outside effectively one-dimensional configurations. At the same time, this leaves open the possibility that trained or regularized FM models produce cells that are in some sense "nicer" than the exact ones, for instance smoother, more stable, or closer to OT cells after approximation effects are taken into account. Clarifying this point would help distinguish which geometric phenomena are intrinsic to FM and which are artifacts of the exact semi-discrete limit.

More broadly, several natural questions remain open. A first direction is to understand how the cells depend on the interpolation itself: how do they change for non-affine interpolations, or after adding noise to the target and considering smoothed variants such as Gaussian mixtures? A second direction is dynamical. One may study the geometry not only of the terminal cells Γ_k, but also of the intermediate cells defined along the flow, as well as their dependence on the locations of the support points $(a_k)_k$. A third direction concerns reflow and straightness: our counterexample rules out a general monotonicity principle in the present setting, but it remains unclear under which assumptions a second reflow becomes straight. For the four-point configuration of Section 5.1, we also conjecture that far from the origin the central cell Γ_4 becomes asymptotically close to the union of the three half-lines from Lemma 9, but a formal proof remains open. Additionally, we conjecture that the assumption of Theorem 7 is always fulfilled. While this was true in numerical simulations, it is important to find a formal proof. Finally, it would be interesting to extend the present picture beyond the equal-weight distinct-atom setting considered here, both to general discrete weights and to asymptotic regimes such as a mean-field limit where the number of support points grows.

Acknowledgments

This research was funded, in part, by the Agence nationale de la recherche (ANR), through the PEPR PDE-AI project (ANR-23-PEIA-0004). JH acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) within project no 530824055 and 572825596. This work was performed using HPC resources from GENCI–IDRIS (Grant 2025-AD011015483R1).

References

Michael Samuel Albergo and Eric Vanden-Eijnden. Building normalizing flows with stochastic interpolants. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=li7qeBbCR1t.

Franz Aurenhammer, Friedrich Hoffmann, and Boris Aronov. Minkowski-type theorems and least-squares clustering. *Algorithmica*, 20(1):61–76, 1998. doi: 10.1007/PL00009191.

Vansh Bansal, Saptarshi Roy, Purnamrita Sarkar, and Alessandro Rinaldo. On the Wasserstein convergence and straightness of rectified flow, 2025. URL https://arxiv.org/abs/2410.14949.

Ricardo Baptista, Agnimitra Dasgupta, Nikola B Kovachki, Assad Oberai, and Andrew M Stuart. Memorization and regularization in generative diffusion models. *arXiv:2501.15785*, 2025.

Quentin Bertrand, Anne Gagneux, Mathurin Massias, and Rémi Emonet. On the closed-form of flow matching: Generalization does not arise from target stochasticity. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2026. URL https://openreview.net/forum?id=kVz9uvqUna.

Samuel Boïté, Julie Delon, and Kimia Nadjahi. Expected batch optimal transport plans and consequences for flow matching. *arXiv preprint arXiv:2605.12174*, 2026.

Matthew G. Brin and T. L. Thickstun. 3-manifolds which are end 1-movable. *Memoirs of the American Mathematical Society*, 81(411):viii+73, 1989. ISSN 0065-9266,1947-6221. doi: 10.1090/memo/0411. URL https://doi.org/10.1090/memo/0411.

Antonin Chambolle and Johannes Hertrich. Limit points of reflow with minibatch optimal transport. *arXiv preprint arXiv:2608.07042*, 2026.

Michael H. Freedman and Frank Quinn. *Topology of 4-manifolds*, volume 39 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1990. ISBN 0-691-08577-3.

Johannes Hertrich, Antonin Chambolle, and Julie Delon. On the relation between rectified flows and optimal transport. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2026. URL https://openreview.net/forum?id=v04csnvCfd.

Jun Kitagawa, Quentin Mérigot, and Boris Thibert. Convergence of a newton algorithm for semi-discrete optimal transport. *Journal of the European Mathematical Society*, 21(9):2603–2651, 2019. doi: 10.4171/JEMS/889.

Hugo Lavenant and Filippo Santambrogio. The flow map of the Fokker–Planck equation does not provide optimal transport. *Applied Mathematics Letters*, 133:108225, 2022. ISSN 0893-9659. doi: https://doi.org/10.1016/j.aml.2022.108225. URL https://www.sciencedirect.com/science/article/pii/S089396592200180X.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matthew Le. Flow matching for generative modeling. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=PqvMRDCJT9t.

Yaron Lipman, Marton Havasi, Peter Holderrieth, Neta Shaul, Matt Le, Brian Karrer, Ricky T. Q. Chen, David Lopez-Paz, Heli Ben-Hamu, and Itai Gat. Flow matching guide and code, 2024. URL https://arxiv.org/abs/2412.06264.

Qiang Liu. Rectified flow: A marginal preserving approach to optimal transport. *arXiv preprint arXiv:2209.14577*, 2022.

Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=XVjTT1nw5z.

Gonzalo Mena, Arun Kumar Kuchibhotla, and Larry Wasserman. Statistical properties of rectified flow. *arXiv preprint arXiv:2511.03193*, 2025.

Quentin Mérigot. A multiscale approach to optimal transport. In *Computer Graphics Forum*, volume 30, pages 1583–1592, 2011. doi: 10.1111/j.1467-8659.2011.02032.x.

Gabriel Peyré and Marco Cuturi. Computational optimal transport. *Foundations and Trends in Machine Learning*, 11(5-6):355–607, 2019. doi: 10.1561/2200000073. URL https://arxiv.org/abs/1803.00567.

Emile Pierret, Valentine Tosel, Julie Delon, and Alasdair Newson. Flow matching for applied mathematicians. 2026. URL https://hal.science/hal-05538982/.

Aram-Alexandre Pooladian, Heli Ben-Hamu, Carles Domingo-Enrich, Brandon Amos, Yaron Lipman, and Ricky T Chen. Multisample flow matching: Straightening flows with minibatch couplings. *ICML 2023*, 2023.

Teodora Reu, Sixtine Dromigny, Michael M. Bronstein, and Francisco Vargas. Gradient variance reveals failure modes in flow-based generative models. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2026. URL https://openreview.net/forum?id=pVaqdFlUAO.

John Stallings. The piecewise-linear structure of Euclidean space. *Proceedings of the Cambridge Philosophical Society*, 58:481–488, 1962. ISSN 0008-1981.

Alexander Tong, Kilian Fatras, Nikolay Malkin, Guillaume Huguet, Yanlei Zhang, Jarrid Rector-Brooks, Guy Wolf, and Yoshua Bengio. Improving and generalizing flow-based generative models with minibatch optimal transport. *Transactions on Machine Learning Research*, 2024. ISSN 2835-8856. URL https://openreview.net/forum?id=CD9Snc73AW. Expert Certification.

Panos Tsimpos, Daniel Sharp, and Youssef Marzouk. One-shot generative flows: Existence and obstructions. *arXiv preprint arXiv:2604.15439*, 2026.

Stephen Zhang, Alireza Mousavi-Hosseini, Michal Klein, and Marco Cuturi. On fitting flow models with large Sinkhorn couplings, 2025. URL https://arxiv.org/abs/2506.05526.

A Existence and Uniqueness of the flow map in the Semi-discrete case

Several papers (including Bansal et al. [2025], Mena et al. [2025]) studied sufficient conditions for the existence and uniqueness of the flow map γ_t, while Chambolle and Hertrich [2026] studies weak solutions in the case where the flow map does not exist. However, in this paper we only study the case where $\mu_0 \sim \mathcal{N}(0, I)$ and where $(X_0, X_1) \sim \mu_0 \otimes \mu_1$ is the independent coupling. In this case, the result is directly clear (for μ_0-almost every x) without further assumptions. Due to the smoothness of the Gaussian, one finds that v_t^* is globally Lipschitz on $[0, 1-\epsilon] \times \mathbb{R}^d$ for any $\epsilon \in (0, 1)$. In particular, we get that the flow map exists and is unique on $[0, 1) \times \mathbb{R}^d$. Because of

$$\begin{aligned}\left(\int\int_0^1 \|\dot\gamma_t(x)\| dt d\mu_0(x)\right)^2 &\le \int\int_0^1 \|\dot\gamma_t(x)\|^2 dt d\mu_0(x) = \int\int_0^1 \|v_t^*(\gamma_t(x))\|^2 dt d\mu_0(x)\\ &= \int_0^1 \int \|v_t^*(x)\|^2 d\mu_t(x) dt \le \mathbb{E}\left[\|X_1 - X_0\|^2\right] < \infty,\end{aligned}$$

we also obtain that the trajectories $t \mapsto \gamma_t(x)$ have finite length (a.e.) and can uniquely be extended to time $t = 1$.

B Proofs of the Simplifications

B.1 Proof of Proposition 2

By Corollary 1 it suffices to consider the case $a_1 = 0$ in which case $E = \operatorname{span}(a_k)_{1\le k\le n}$. Then, for $1 \le k \le n$, $x \in \mathbb{R}^d$, $t \in [0, 1)$, we have by the Pythagorean theorem that

$$\|x - ta_k\|^2 = \|x^{\|} - ta_k\|^2 + \|x^{\perp}\|^2.$$

Consequently,

$$\alpha_k(t,x) = \frac{e^{-\frac{1}{2(1-t)^2}(\|x^{\|}-ta_k\|^2+\|x^{\perp}\|^2)}}{\sum_{i=1}^n e^{-\frac{1}{2(1-t)^2}(\|x^{\|}-ta_k\|^2+\|x^{\perp}\|^2)}} = \frac{e^{-\frac{1}{2(1-t)^2}\|x^{\|}-ta_k\|^2}}{\sum_{i=1}^n e^{-\frac{1}{2(1-t)^2}\|x^{\|}-ta_k\|^2}} = \alpha_k(t, x^{\|}).$$

Finally,

$$\frac{dx_t}{dt} = \frac{dx_t^{\|}}{dt} + \frac{dx_t^{\perp}}{dt} = \underbrace{\sum_{k=1}^n \alpha_k(t, x_t^{\|})\frac{a_k - x_t^{\|}}{1-t}}_{\in \operatorname{span}(a_k)_{1\le k\le n}} + \underbrace{-\frac{x_t^{\perp}}{1-t}}_{\in \operatorname{span}(a_k)_{1\le k\le n}^{\perp}}$$

This completes the proof. □

B.2 Proof of Proposition 3

We start with the following intermediate result which uniformly controls the distance between γ_t^ε and γ_t for $t \leq T < 1$.

Lemma 14. *For every $R > 0$ and $T < 1$,*

$$\sup_{\|x\| \leq R} \sup_{0 \leq t \leq T} \|\gamma_t^\varepsilon(x) - \gamma_t(x)\| \to 0$$

Proof. Let $e_t^\varepsilon(x) = \gamma_t^\varepsilon(x) - \gamma_t(x)$. Using $\dot{e}_t^\varepsilon(x) = v_t^\varepsilon(\gamma_t^\varepsilon(x)) - v_t(\gamma_t(x))$ and $e_0^\varepsilon(x) = 0$, we have:

$$\begin{aligned}\|e_t(x)\| &\leq \int_0^t \|v_s^\varepsilon(\gamma_s^\varepsilon(x)) - v_s(\gamma_s(x))\|\, ds \\ &\leq \int_0^t \|v_s^\varepsilon(\gamma_s^\varepsilon(x)) - v_s^\varepsilon(\gamma_s(x))\|\, ds + \int_0^t \|v_s^\varepsilon(\gamma_s(x)) - v_s(\gamma_s(x))\|\, ds \end{aligned} \tag{6}$$

We now prove that for bounded initialization x and $t \leq T$, the iterates $\gamma_t(x)$ and $\gamma_t^\varepsilon(x)$ remain uniformly bounded. Indeed, we can bound

$$\|v_t^\varepsilon(x)\| \leq \sup_{a_k} \left\| \frac{(1-t)(a_k - x) + t\varepsilon^2 x}{(1-t)^2 + t^2\varepsilon^2} \right\| \leq C_T(1 + \|x\|)$$

where C_T depends on T, not on ϵ. In the same way $\|v_t(x)\| \leq C_T(1 + \|x\|)$ and thus integrating, we get that there is $M > 0$ such that

$$\sup_{\|x\| \leq R} \sup_{0 \leq t \leq T} \|\gamma_t^\varepsilon(x)\| \leq M, \qquad \sup_{\|x\| \leq R} \sup_{0 \leq t \leq T} \|\gamma_t(x)\| \leq M$$

In order to handle the first term of (6), note that v_t^ε is uniformly Lipschitz on $K_T := [0, T] \times B(0, M)$: there is $L_T > 0$, such that for all $t \in [0, T]$ and $x, y \in B(0, R)$,

$$\|v_s^\varepsilon(\gamma_s^\varepsilon(x)) - v_s^\varepsilon(\gamma_s(x))\| \leq L_T \, \|e_s(x)\| .$$

Therefore, we get

$$\|e_t(x)\| \leq L_T \int_0^t \|e_s(x)\|\, ds + t \sup_{0 \leq s \leq T, \|x\| \leq M} \|v_s^\varepsilon(x) - v_s(x)\| .$$

Applying Gronwall's lemma gives:

$$\sup_{t \leq T, \|x\| \leq R} \|e_t(x)\| \leq T e^{TL_T} \sup_{0 \leq s \leq T, \|x\| \leq M} \|v_s^\varepsilon(x) - v_s(x)\| .$$

Finally, using that v_t^ε converges uniformly on K_T to its $\varepsilon = 0$ counterpart, the right-hand side tends to 0. □

Now, we can prove Proposition 3.

Proof of Proposition 3. Let $k \in \{1, \ldots, n\}$ be such that $\gamma_t(x) \to a_k$. Let

$$d_k := \min_{i \neq k} \|a_i - a_k\| ,$$

and choose $0 < \rho < d_k/4$. There exists $T_0 < 1$ such that

$$\gamma_t(x) \in B(a_k, \rho/8) \qquad \text{for all } t \in [T_0, 1).$$

By Lemma 14, for this fixed T_0, there exists $\varepsilon_0 > 0$ such that

$$\left\|\gamma_{T_0}^\varepsilon(x) - \gamma_{T_0}(x)\right\| \leq \rho/8 \qquad \text{for all } 0 < \varepsilon < \varepsilon_0.$$

Hence

$$\left\|\gamma_{T_0}^\varepsilon(x) - a_k\right\| \leq \rho/4 \qquad \text{for all } 0 < \varepsilon < \varepsilon_0. \tag{7}$$

We now define τ_ε as the time where the trajectory $\gamma_t^\varepsilon(x)$ leaves $B(a_k, \rho)$:

$$\tau_\varepsilon := \inf\left\{t \in [T_0, 1] : \ \gamma_t^\varepsilon(x) \notin B(a_k, \rho)\right\},$$

with the convention $\tau_\varepsilon = 1$ if the set is empty. Then

$$\gamma_t^\varepsilon(x) \in B(a_k, \rho) \qquad \text{for all } t \in [T_0, \tau_\varepsilon].$$

Set

$$D_\varepsilon(t) := (1-t)^2 + t^2\varepsilon^2, \qquad r_t^\varepsilon(x) := \gamma_t^\varepsilon(x) - a_k, \qquad m_t^\varepsilon(z) := \sum_{i=1}^n \alpha_i^\varepsilon(t,z)\, a_i.$$

The flow map $\gamma_t^\varepsilon(x)$ satisfies the ODE

$$\frac{d}{dt}\gamma_t^\varepsilon(x) = \frac{(1-t)\big(m_t^\varepsilon(\gamma_t^\varepsilon(x)) - \gamma_t^\varepsilon(x)\big) + t\varepsilon^2\gamma_t^\varepsilon(x)}{D_\varepsilon(t)},$$

which can be rewritten for $r_t^\varepsilon(x)$ as

$$\frac{d}{dt}r_t^\varepsilon(x) = -\frac{1-t-t\varepsilon^2}{D_\varepsilon(t)}\, r_t^\varepsilon(x) + \frac{1-t}{D_\varepsilon(t)}\big(m_t^\varepsilon(\gamma_t^\varepsilon(x)) - a_k\big) + \frac{t\varepsilon^2}{D_\varepsilon(t)}\, a_k.$$

We now introduce the rescaled quantity

$$q_t^\varepsilon(x) := \frac{r_t^\varepsilon(x)}{\sqrt{D_\varepsilon(t)}}$$

in order to factor out the scale $\sqrt{D_\varepsilon(t)}$ of r_t^ε near $t = 1$. It satisfies the ODE:

$$\frac{d}{dt}q_t^\varepsilon(x) = F_\varepsilon(t), \tag{8}$$

where

$$F_\varepsilon(t) := \frac{1-t}{D_\varepsilon(t)^{3/2}}\big(m_t^\varepsilon(\gamma_t^\varepsilon(x)) - a_k\big) + \frac{t\varepsilon^2}{D_\varepsilon(t)^{3/2}}\, a_k.$$

In order to study F_ϵ, we next estimate the norm $\|m_t^\varepsilon(z) - a_k\|$ for any $z \in B(a_k, \rho)$. We have

$$\|m_t^\varepsilon(z) - a_k\| = \left\|\sum_{i\neq k} \alpha_i^\varepsilon(t,z)(a_i - a_k)\right\| \leq \sum_{i\neq k} \frac{\alpha_i^\varepsilon(t,z)}{\alpha_k^\varepsilon(t,z)}\, \|a_i - a_k\|$$

For $i \neq k$, define

$$\begin{aligned}\Delta_i(t,z) &:= \|z - ta_i\|^2 - \|z - ta_k\|^2 \\ &= t\big(\|z-a_i\|^2 - \|z-a_k\|^2\big) - t(1-t)\big(\,\|a_i\|^2 - \|a_k\|^2\,\big),\end{aligned}$$

We have that $z \in B(a_k, \rho)$ implies $\|z-a_i\|^2 - \|z-a_k\|^2 \geq c_0$ for some $c_0 > 0$. Thus, after increasing T_0 if necessary, for all $t \in [T_0, 1]$ and $z \in B(a_k, \rho)$, for any $i \neq k$, $\Delta_i(t,z) \geq c_0/2$ and

$$\frac{\alpha_i^\varepsilon(t,z)}{\alpha_k^\varepsilon(t,z)} = \exp\left(-\frac{\Delta_i(t,z)}{2((1-t)^2 + t^2\varepsilon^2)}\right) \leq \exp\left(-\frac{c_0}{4((1-t)^2 + t^2\varepsilon^2)}\right).$$

Summing over $i \neq k$, we obtain constants $c, C > 0$ such that

$$\|m_t^\varepsilon(z) - a_k\| \leq C\exp\left(-\frac{c}{(1-t)^2 + \varepsilon^2}\right) \qquad \text{for all } t \in [T_0, 1],\ z \in B(a_k, \rho),\ 0 < \varepsilon < \varepsilon_0. \tag{9}$$

Hence, on $[T_0, \tau_\varepsilon]$, estimate (9) applies: the first term of $F_\varepsilon(t)$ is then bounded by

$$C\frac{1-t}{((1-t)^2 + \varepsilon^2)^{3/2}}\exp\left(-\frac{c}{(1-t)^2 + \varepsilon^2}\right),$$

whose integral on $[T_0, 1]$ is uniformly bounded in ε. Note also that the last term of F_ϵ is uniformly bounded in $L^1([T_0, 1])$. Indeed:

$$I_{T_0} := \int_{T_0}^1 \frac{t\varepsilon^2}{D_\varepsilon(t)^{3/2}}dt \leq \int_{T_0}^1 \frac{\varepsilon^2}{((1-t)^2 + T_0^2\varepsilon^2)^{3/2}}dt = \frac{1}{T_0^2}\int_{T_0}^1 \frac{1}{(\frac{(1-t)^2}{T_0^2\varepsilon^2} + 1)^{3/2}}\frac{1}{T_0\varepsilon}dt$$

and a change of variable $u = \frac{1-t}{T_0\varepsilon}$ gives $I_{T_0} \leq \frac{1}{T_0^2}$.

Thus, integrating (8) on $[T_0, \tau_\varepsilon]$, we obtain

$$\|q_t^\varepsilon(x)\| \leq \left\|q_{T_0}^\varepsilon(x)\right\| + C_1 \qquad \text{for all } t \in [T_0, \tau_\varepsilon].$$

Multiplying by $\sqrt{D_\varepsilon(t)}$, this gives

$$\|r_t^\varepsilon(x)\| \leq \frac{\sqrt{D_\varepsilon(t)}}{\sqrt{D_\varepsilon(T_0)}} \left\|r_{T_0}^\varepsilon(x)\right\| + C_1\sqrt{D_\varepsilon(t)}.$$

Since $D_\varepsilon = (1-t)^2 + t^2\varepsilon^2$ is convex in t, its maximum on $[T_0, 1]$ is attained at one of the endpoints. Now, for $\varepsilon < \varepsilon_0 < \frac{1-T_0}{1+T_0}$, we get $D_\varepsilon(1) = \varepsilon^2 \leq D_\varepsilon(T_0)$, for all $t \in [T_0, 1]$, $D_\varepsilon(1) \leq D_\varepsilon(T_0)$. Thus:

$$\|r_t^\varepsilon(x)\| \leq \left\|r_{T_0}^\varepsilon(x)\right\| + C_1\sqrt{D_\varepsilon(t)} \qquad \text{for all } t \in [T_0, \tau_\varepsilon].$$

Using (7) and $\sqrt{D_\varepsilon(t)} \leq (1-t) + \varepsilon$, we get

$$\|\gamma_t^\varepsilon(x) - a_k\| \leq \frac{\rho}{4} + C_1\big((1-t) + \varepsilon\big) \qquad \text{for all } t \in [T_0, \tau_\varepsilon]. \tag{10}$$

Now choose T_0 closer to 1, and decrease ε_0 if necessary, so that

$$\frac{\rho}{4} + C_1\big((1-T_0) + \varepsilon_0\big) < \rho.$$

Then (10) implies

$$\gamma_t^\varepsilon(x) \in B(a_k, \rho) \qquad \text{for all } t \in [T_0, \tau_\varepsilon].$$

By the definition of τ_ε, this forces $\tau_\varepsilon = 1$.

We have therefore proved that there are $T_0 > 0$ and $\varepsilon_0 > 0$ such that for all $t \in [T_0, 1]$ and all $0 < \varepsilon < \varepsilon_0$,

$$\|\gamma_t^\varepsilon(x) - a_k\| \leq \rho$$

□

Corollary 15. *Let* $\Gamma_k^\varepsilon = \left\{x \in \mathbb{R}^d : \|\gamma^\varepsilon(x) - a_k\| < \|\gamma^\varepsilon(x) - a_l\| \text{ for all } l \neq k\right\}$. *Then, for every* $x \in \bigcup_{j=1}^n \Gamma_j$ *and every* $k \in \{1, \ldots, n\}$,

$$\mathbf{1}_{\Gamma_k^\varepsilon}(x) \longrightarrow \mathbf{1}_{\Gamma_k}(x) \qquad \text{as } \varepsilon \to 0.$$

Proof. Let $x \in \bigcup_{j=1}^n \Gamma_j$. Then there exists a unique index $k \in \{1, \ldots, n\}$ such that $x \in \Gamma_k$. By definition of Γ_k, this means that $\gamma_1(x) = a_k$. By Proposition 3, we have $\gamma^\varepsilon(x) \to a_k$ as $\varepsilon \to 0$.

Since the atoms $a_1, \ldots, a_n$ are pairwise distinct, set $r_k := \frac{1}{2}\min_{j\neq k}\|a_k - a_j\| > 0$.

Using Proposition 3, for sufficiently small $\varepsilon > 0$, we have

$$\|\gamma^\varepsilon(x) - a_k\| < r_k$$

and for every $j \neq k$,

$$\|\gamma^\varepsilon(x) - a_j\| \geq \|a_k - a_j\| - \|\gamma^\varepsilon(x) - a_k\| > 2r_k - r_k = r_k.$$

Therefore, $\|\gamma^\varepsilon(x) - a_k\| < \|\gamma^\varepsilon(x) - a_j\|$ for all $j \neq k$ and thus $x \in \Gamma_k^\varepsilon$. In particular, $x \notin \Gamma_j^\varepsilon$ for all $j \neq k$. Consequently, for all sufficiently small $\varepsilon > 0$, $\mathbf{1}_{\Gamma_k^\varepsilon}(x) = 1 = \mathbf{1}_{\Gamma_k}(x)$, while for every $j \neq k$, $\mathbf{1}_{\Gamma_j^\varepsilon}(x) = 0 = \mathbf{1}_{\Gamma_j}(x)$. □

C Proofs for Topology of the FM Cells

C.1 Proof of Proposition 4

We start by proving that α_k converges to one in a ball around a_k for late enough time. More precisely, we have the following lemma.

Lemma 16. *Let $a_k \neq a_l$ for $k \neq l \in \{1, ..., n\}$ and define $r < \frac{1}{2}\min_{k\neq l} \|a_k - a_l\|$. Then, there exists a time $t_0 \in (0,1)$ and $\delta > 0$ such that for all $t > t_0$ it holds*

$$\alpha_k(t,x) > \frac{\exp\left(\frac{\delta}{2(1-t)^2}\right)}{\exp\left(\frac{\delta}{2(1-t)^2}\right) + n - 1} \to 1 \quad \text{as } t \to 1.$$

Proof. There exists $\delta > 0$ such that for all $x \in \bar{B}_r(a_k)$ and $l \neq k$ it holds

$$\|x - a_k\|^2 < \|x - a_l\|^2 - 2\delta.$$

Thus there exists $t_0 \in (0,1)$ such that for all $t > t_0$ it holds

$$\|x - ta_k\|^2 < \|x - ta_l\|^2 - \delta.$$

Furthermore, we have

$$\tilde{\alpha}_{k,l}(t,x) \coloneqq \frac{\exp\left(-\frac{\|x-ta_k\|^2}{2(1-t)^2}\right)}{\exp\left(-\frac{\|x-ta_l\|^2}{2(1-t)^2}\right)} > \exp\left(\frac{\delta}{2(1-t)^2}\right).$$

This implies (for such t, x)

$$\alpha_k(t,x) = \frac{1}{1 + \sum_{k\neq l} \frac{1}{\tilde{\alpha}_{k,l}(t,x)}} > \frac{1}{1 + (n-1)\exp\left(-\frac{\delta}{2(1-t)^2}\right)} = \frac{\exp\left(\frac{\delta}{2(1-t)^2}\right)}{\exp\left(\frac{\delta}{2(1-t)^2}\right) + n - 1}$$

□

This allows us now to prove that the velocity $v_t(x)$ on $\partial B_r(a_k)$ points into the ball for large enough time. We formalize this observation in the following lemma.

Lemma 17. *Let $a_k \neq a_l$ for $k \neq l \in \{1, ..., n\}$ and define $r < \frac{1}{2}\min_{k\neq l} \|a_k - a_l\|$. Then there exists $\underline{t}_r \in (0,1)$ such that for all $t > \underline{t}_r$ and $x \in \partial B_r(a_k)$ we have that $\langle a_k - x, v_t(x)\rangle > 0$.*

Proof. We have for $M \coloneqq \max_l \|a_k - a_l\|$

$$\begin{aligned}
(1-t)\langle a_k - x, v_t(x)\rangle &= \alpha_k(t,x)\|a_k - x\|^2 + \sum_{l\neq k} \alpha_l(t,x)\langle a_k - x, a_l - x\rangle \\
&\geq \alpha_k(t,x)r^2 - \sum_{l\neq k} \alpha_l(t,x)\|a_k - x\|\|a_l - x\| \\
&\geq \alpha_k(t,x)r^2 - r(M+r)\sum_{l\neq k} \alpha_l(t,x) \\
&= \alpha_k(t,x)r^2 - r(M+r)(1 - \alpha_k(t,x))
\end{aligned}$$

which converges to r^2 for $\alpha_k(t,x) \to 1$. Since $\alpha_k(t,x)$ converges to 1 uniformly on $\bar{B}_r(a_k)$ by Lemma 16, we obtain that there exists some time $\underline{t}_r$ such that the above expression is positive for all $t > \underline{t}_r$. □

This lemma directly implies that a trajectory will never leave the ball $\bar{B}_r(a_k)$ once it has entered it at a late enough time as formalized in the following corollary.

Corollary 18. *Let $x \in \mathbb{R}^d$. If there exists $\underline{t}_r < t_0 < 1$ such that $\gamma_{t_0}(x) \in \bar{B}_r(a_k)$ then $\forall t \geq t_0, \gamma_t(x) \in \bar{B}_r(a_k)$.*

Finally, we can prove the desired proposition, by establishing convergence rates using the "inward pointing lemma" Lemma 17.

Proof of Proposition 4. By contradiction, we suppose that $\gamma_t(x) \in \bar{B}_r(a_k)$ for $\underline{t}_r < t < 1$ and $\gamma_t(x)$ does not converge to a_k when $t \to 1$. With this assumption, there exists $0 < \varepsilon_0 < r$ such that for all $\underline{t}_{\varepsilon_0} < T < 1$ there exists $t > T$ such that $\|\gamma_t(x) - a_k\| > 2\varepsilon_0$. It implies that there exists a sequence (t_n) such that $t_n \to 1$, $t_n > \underline{t}_{\varepsilon_0}$ and $\|\gamma_{t_n}(x) - a_k\| > 2\varepsilon_0$. By Corollary 18, it implies that $\forall t \geq t_1$, $\|\gamma_t(x) - a_k\| > \varepsilon_0$. Consequently, for all $t \geq t_1$ and with $M \coloneqq \max_l \|a_k - a_l\|$,

$$
\begin{aligned}
&(1-t)\langle a_k - \gamma_t(x), v_t(\gamma_t(x))\rangle \\
&= \alpha_k(t,\gamma_t(x))\|a_k - \gamma_t(x)\|^2 + \sum_{l\neq k} \alpha_l(t,\gamma_t(x))\langle a_k - \gamma_t(x), a_l - \gamma_t(x)\rangle \\
&\geq \alpha_k(t,\gamma_t(x))\varepsilon_0^2 - \sum_{l\neq k} \alpha_l(t,\gamma_t(x))\|a_k - \gamma_t(x)\|\|a_l - \gamma_t(x)\| \\
&\geq \alpha_k(t,\gamma_t(x))\varepsilon_0^2 - r(M+r)\sum_{l\neq k} \alpha_l(t,\gamma_t(x)) \\
&= \alpha_k(t,\gamma_t(x))\varepsilon_0^2 - r(M+r)(1-\alpha_k(t,\gamma_t(x))) \\
&= \alpha_k(t,\gamma_t(x))(\varepsilon_0^2 + r^2 + rM) - r(M+r) \\
&> \frac{\exp\left(\frac{\delta}{2(1-t)^2}\right)}{\exp\left(\frac{\delta}{2(1-t)^2}\right) + n - 1}(\varepsilon_0^2 + r^2 + rM) - r(M+r) \\
&= \frac{\exp\left(\frac{\delta}{2(1-t)^2}\right)(\varepsilon_0^2 + r^2 + rM - r(M+r)) - r(M+r)(n-1)}{\exp\left(\frac{\delta}{2(1-t)^2}\right) + n - 1} \\
&= \frac{\exp\left(\frac{\delta}{2(1-t)^2}\right)\varepsilon_0^2 - r(M+r)(n-1)}{\exp\left(\frac{\delta}{2(1-t)^2}\right) + n - 1} \\
&= \varepsilon_0^2 - \frac{(\varepsilon_0^2 + r(M+r))(n-1)}{\exp\left(\frac{\delta}{2(1-t)^2}\right) + n - 1}
\end{aligned}
$$

Let us consider $F(t) = \frac{1}{2}\|\gamma_t(x) - a_k\|^2$, for all $t \geq t_1$,

$$
\begin{aligned}
\dot{F}(t) &= \langle \gamma_t(x) - a_k, \dot{\gamma}_t(x)\rangle = -\langle a_k - \gamma_t(x), v_t(\gamma_t(x))\rangle \\
&\leq \frac{1}{1-t}\frac{(\varepsilon_0^2 + r(M+r))(n-1)}{\exp\left(\frac{\delta}{2(1-t)^2}\right) + n - 1} - \frac{\varepsilon_0^2}{1-t}
\end{aligned}
$$

And, by integrating between t_1 and $t \geq t_1$,

$$
\begin{aligned}
F(t) &\leq F(t_1) + \int_{t_1}^{t} \frac{1}{1-s}\frac{(\varepsilon_0^2 + r(M+r))(n-1)}{\exp\left(\frac{\delta}{2(1-s)^2}\right) + n - 1} ds - \int_{t_1}^{t} \frac{\varepsilon_0^2}{1-s} ds \\
&= F(t_1) + (\varepsilon_0^2 + r(M+r))(n-1) \underbrace{\int_{t_1}^{t} \frac{1}{1-s}\frac{1}{\exp\left(\frac{\delta}{2(1-s)^2}\right) + n - 1} ds}_{<+\infty} + \varepsilon_0^2[\log(1-s)]_{t_1}^{t} \\
&= C_{t_1,r,\varepsilon_0} + \varepsilon_0^2 \log(1-t) \to -\infty \quad \text{as } t \to 1.
\end{aligned}
$$

Since we know that $F(t) \geq 0$, this is a contradiction and we are done. □

C.2 Proofs of the Topological Consequences

Proposition 19. *For $1 \leq k \leq n$, Γ_k is an open set.*

Proof. For $1 \leq k \leq n$, let us show that $\Gamma_k = \bigcup_{\underline{t}_r < t < 1} \gamma_t^{-1}(B_r(a_k))$ which is a union of open sets because γ_t^{-1} is well-defined and continuous for $t < 1$.

First, $\Gamma_k \subset \bigcup_{\underline{t}_r < t < 1} \gamma_t^{-1}(B_r(a_k))$. Indeed, if $\lim_{t \to 1^-} \gamma_t(x) = a_k$, there exists a time t_0 such that $\forall t \geq t_0, \gamma_t(x) \in B_r(a_k)$.

Second, by Proposition 4 and Corollary 18, $\bigcup_{\underline{t}_r < t < 1} \gamma_t^{-1}(B_r(a_k)) \subset \Gamma_k$. Indeed, if we suppose that $x \in \gamma_{t_0}^{-1}(B_r(a_k))$ for $\underline{t}_r < t_0 < 1$ then by Proposition 4, $x \in \Gamma_k$. □

Proposition 20. *Γ_k is path-connected.*

Proof. Let $x_0, x_1 \in \Gamma_k$ and take r and $\underline{t}_r$ from Lemma 17. Since the mappings $t \mapsto \gamma_t(x_0)$ and $t \mapsto \gamma_t(x_1)$ are continuous, there exist time points $t_0, t_1 \in (\underline{t}_r, 1)$ such that $\gamma_{t_0}(x_0), \gamma_{t_1}(x_1) \in \bar{B}_r(a_k)$. Given Corollary 18, it implies that for all $t \geq t_0, \gamma_t(x_0) \in \bar{B}(a_k, r)$ and for all $t \geq t_1, \gamma_t(x_1) \in \bar{B}(a_k, r)$. Without loss of generality we can assume that $t_0 > t_1$. Now, we define $x_c = \gamma_{t_0}^{-1}((1-c)\gamma_{t_0}(x_0) + c\gamma_{t_0}(x_1))$ and note that the path $c \mapsto x_c$ is continuous and connects x_0 and x_1. It remains to show that $x_c \in \Gamma_k$. This is provided by Proposition 4. Let $c \in [0, 1]$. Then $\gamma_{t_0}(x_c) = (1-c)\gamma_{t_0}(x_0) + c\gamma_{t_0}(x_1) \in \bar{B}_r(a_k)$ (by convexity of the ball). Because $t_0 \geq \underline{t}_r$, by Proposition 4, $x_c \in \Gamma_k$. □

To show that Γ_k is actually simply connected, we first prove that for any compact subset $A \subseteq \Gamma_k$ we have that $\gamma_t(A) \subseteq \bar{B}_r(a_k)$ for r small enough and t large enough. More precisely, we have the following lemma.

Lemma 21. *Let $A \subset \Gamma_k$ be compact and $r < \frac{1}{2} \min_{l \neq k} \|a_k - a_l\|$. Then, there exists some time $t^* \in (0, 1)$ such that for all $t > t^*$ it holds $\gamma_t(A) \subseteq \bar{B}_r(a_k)$*

Proof. Let $\underline{t}_r$ be from Proposition 4. Then, we denote $S_k(t) \coloneqq \gamma_t^{-1}(B_r(a_k))$ and observe that $\Gamma_k = \cup_{t \geq \underline{t}_r} S_k(t)$ with increasing $S_k(t)$. Since A is compact and the $S_k(t)$ for $t \geq \underline{t}_r$ are open and cover A, there exists a finite subcover $S_k(t_1), ..., S_k(t_T)$ such that $A \subset \cup_{i=1}^T S_k(t_i) = S_k(t^*)$ for $t^* = \max(t_1, ..., t_T)$. In other words, $\gamma_{t^*}(A) \subseteq \bar{B}_r(a_k)$ which implies by Proposition 4 and corollary 18 also that $\gamma_t(A) \subseteq \bar{B}_r(a_k)$ for all $t \geq t^*$. □

Applying this lemma to an arbitrary loop yields the following proposition.

Proposition 22. *Γ_k is simply connected.*

Proof. Take a continuous loop $f : S^1 \to \Gamma_k$. We have to build a homotopy $H : S^1 \times [0, 1] \to \Gamma_k$ s.t. $H(., 0) = f$ and $H(., 1)$ is a constant loop. By Lemma 21, we know that there exists some t^* such that $f(S^1) \subseteq \gamma_{t^*}^{-1}(\bar{B}_r(a_k)) \subseteq \Gamma_k$. Since $\gamma_{t^*}^{-1}(\bar{B}_r(a_k))$ is diffeomorphic to a ball and thus simply connected, there exists a continuous $H_{t^*} : S^1 \times [0, 1] \to \gamma_{t^*}^{-1}(\bar{B}_r(a_k))$ such that $H_{t^*}(., 0) = f$ and $H_{t^*}(., 1)$ is a constant loop. □

C.3 Existence of the limit of $\gamma_t^{-1}(a_k)$ for $t \to 1$ and Contractibility

First, we bound the gradient norm of v_t globally and locally around a_k.

Lemma 23. *There exists some $c > 0$ such that $\|\nabla v_t(x)\| \leq \frac{c}{(1-t)^3}$. Further, for any $\epsilon > 0$ and $r < \frac{1}{2} \min_{l \neq k} \|a_k - a_l\|$ there exists $\underline{t} \in (0, 1)$ such that for any $t > \underline{t}$ we have that for all $x \in B_r(a_k)$ it holds $\|\nabla v_t(x)\| < \frac{1+\epsilon}{1-t}$.*

Proof. Recall that $v_t(x) = -\frac{x}{1-t} + \frac{1}{1-t} \sum_{l=1}^n s_l(-\frac{\|x - t a_j\|^2}{2(1-t)^2}) a_l$, where s_l is the softmax function. Using the chain rule and the derivative $\partial_j s_l(x) = s_l(x)(\delta_{jl} - s_j(x))$ this implies that

$$\nabla v_t(x) = -\frac{I}{1-t} - \frac{1}{(1-t)^3} \sum_{l=1}^n \alpha_l(t, x) \Big(a_l - \sum_{j=1}^n \alpha_j(t, x) a_j \Big) (x - t a_l)^T$$

Noting that

$$\sum_{l=1}^n \alpha_l(t,x)\Big(a_l - \sum_{j=1}^n \alpha_j(t,x)a_j\Big)x^T = \sum_{l=1}^n \alpha_l(t,x)a_l x^T - \underbrace{\Big(\sum_{l=1}^n \alpha_l(t,x)\Big)}_{=1}\sum_{j=1}^n \alpha_j(t,x)a_j x^T = 0,$$

this simplifies to

$$\nabla v_t(x) = -\frac{I}{1-t} + \frac{t}{(1-t)^3}\underbrace{\sum_{l=1}^n \alpha_l(t,x)\Big(a_l - \sum_{j=1}^n \alpha_j(t,x)a_j\Big)a_l^T}_{A_t(x)}$$

This implies that $\|\nabla v_t(x)\| \leq \frac{1}{1-t} + \frac{\|A_t(x)\|}{(1-t)^3}$. Thus, we need to bound $\|A_t(x)\|$. We find that (for the convenience of notation, we omit the argument x)

$$A_t = \alpha_k(t)\Big((1-\alpha_k(t))a_k - \sum_{j\neq k}\alpha_j(t)a_j\Big)a_k^T + \sum_{l\neq k}\alpha_l(t)\Big(a_l - \sum_{j=1}^n \alpha_j(t)a_j\Big)a_l^T.$$

Using the notation $M = \max_{l=1,\dots,n}\|a_l\|$ and that $\alpha_j(t)$ sum to 1, we obtain

$$\begin{aligned}\|A_t\| &\leq \alpha_k(t)M\Big\|(1-\alpha_k(t))a_k + \sum_{j\neq k}\alpha_j(t)a_j\Big\| + 2M^2\sum_{l\neq k}\alpha_l(t)\\ &\leq 4M^2(1-\alpha_k(t))\end{aligned}$$

such that (writing the argument x again) we have

$$\|\nabla v_t(x)\| \leq \frac{1}{1-t} + \frac{4M^2(1-\alpha_k(t,x))}{(1-t)^3}$$

Noting that $1-\alpha_k(t,x) \leq 1$ shows the global bound. For the local bound, we find by Lemma 16 that there exist some $t_0 \in (0,1)$ and $\delta > 0$ such that for all $t \geq t_0$ and $x \in B_r(a_k)$ we have

$$1-\alpha_k(t,x) < \frac{n-1}{\exp(\frac{\delta}{2(1-t)^2}) + n - 1}.$$

Since the exponential function is growing faster than any polynomial, we find that for any $\epsilon > 0$ there exists some $1 > t_1 > t_0$ such that for all $t > t_1$ it holds $1-\alpha_k(t,x) < \epsilon\frac{(1-t)^2}{4M^2}$. In particular, for such t and x, we obtain $\|\nabla v_t(x)\| \leq \frac{1+\epsilon}{1-t}$. Choosing $\underline{t} = t_1$ yields the local bound. □

Next, we bound the gradient norm of γ_t^{-1}.

Lemma 24. *There exist constants $C_1^*, C_2^* > 0$ independent of t (but depending on $a_1, ..., a_n$) such that $\|\nabla[\gamma_t^{-1}](x)\| \leq \frac{C_1^*}{(1-t)^2}$ for all $t \in (0,1)$ and $x \in B_{C_2^*(1-t)}(a_k)$.*

Proof. For $w_s = v_{t-s}$ we consider the ODE $\dot\eta_s(x) = -w_s(\eta_s(x))$ with initial condition $\eta_0(x) = x$. We note that $\eta_t = \gamma_t^{-1}$.

Step 1: First, we bound the velocity $w_s(x)$ locally around a_k for small s. To this end, we note that $v_t(x) = \frac{1}{1-t}\sum_l \alpha_l(t,x)(a_l - x) = \frac{1}{1-t}\Big(\sum_l \alpha_l(t,x)a_l - a_k\Big) + \frac{a_k - x}{1-t}$. In particular, we obtain with $M = \max_l\|a_k - a_l\|$ that

$$\|v_t(x)\| \leq \frac{M}{1-t}\sum_{l\neq k}\alpha_l(t,x) + \frac{\|a_k - x\|}{1-t} = \frac{M(1-\alpha_k(t,x))}{1-t} + \frac{\|a_k - x\|}{1-t},$$

or equivalently

$$\|w_s(x)\| \leq \frac{M(1-\alpha_k(t-s,x))}{1-t+s} + \frac{\|a_k - x\|}{1-t+s}.$$

Now let $r \leq \frac{1}{2}\min_{l\neq k}\|a_k - a_l\|$. Then, we know by Lemma 16 that there exist some $t_0 \in (0,1)$ and $\delta > 0$ such that for all $t \geq t_0$ and $x \in B_r(a_k)$ we have $1-\alpha_k(t,x) < \frac{n-1}{\exp(\frac{\delta}{2(1-t)^2})+n-1}$. In particular, this implies that there exists some $1 > t_1 \geq t_0$ and $C > 0$ such that for all $t \geq t_1$ it holds $1-\alpha_k(t,x) < C\exp(-\frac{\delta}{2(1-t)^2})$. Inserting this in the estimation above, we obtain for $s \leq t - t_1$ that

$$\|w_s(x)\| \leq \frac{CM\exp(-\frac{\delta}{2(1-t+s)^2})}{1-t+s} + \frac{\|a_k - x\|}{1-t+s}.$$

Step 2: Next, we show that $\eta_s(x)$ remains close to a_k for small s whenever x is close to a_k. To this end let $r_t = \frac{r(1-t)}{2}$ and $x \in B_{r_t}(a_k)$. Then, we define the auxiliary function $F_x(s) = \|\eta_s(x) - a_k\|$. To show that $F(s) < r$ for small s, we define $s_0 = \min\{s \geq 0 : F(s) \geq r\}$ (note that the minimum exists, since F is continuous and fulfills $F(0) \leq r_t < r$). Then, it holds for all $s \leq \min(s_0, t - t_1)$ that

$$\dot{F}_x(s) \leq \|w_s(\eta_s(x))\| \leq \frac{CM\exp(-\frac{\delta}{2(1-t+s)^2})}{1-t+s} + \frac{F_x(s)}{1-t+s}.$$

In particular, we get by the Gronwall lemma that for $s \leq \min(s_0, t-t_1)$ it holds

$$\begin{aligned} F_x(s) &\leq CM(1-t+s)\int_0^s \frac{\exp(-\frac{\delta}{2(1-t+a)^2})}{(1-t+a)^2}da + \frac{1-t+s}{1-t}F(0) \\ &\leq \frac{2CMs}{\delta} + \frac{r_t}{1-t} = \frac{2CMs}{\delta} + \frac{r}{2}. \end{aligned}$$

Inserting $s^* = \min(s_0, t-t_1)$ this implies that $r = F(s^*) \leq \frac{2CMs^*}{\delta} + \frac{r}{2}$ and therefore $s_0 \geq s^* \geq \frac{\delta r}{4CM}$. By the definition of s_0, we have shown that for $s < \min(\frac{\delta r}{4CM}, t-t_1)$ it holds that $F(s) < r$ and therefore $\eta_s(a_k) \in B_r(a_k)$.

Step 3: Finally, we show the actual result on the Jacobian of $\eta_t = \gamma_t^{-1}$. For $J_s(x) = \nabla\eta_s(x)$ we know that it fulfills the ODE $\dot{J}_s(x) = -\nabla w_s(\eta_s(x))J_s(x)$ with initial condition $J_0(x) = I$. For the norm, we obtain

$$\frac{d}{ds}\|J_s(x)\| \leq \|\nabla w_s(\eta_s(x))\|\|J_s(x)\|$$

such that the Gronwall lemma implies

$$\|\nabla[\gamma_t^{-1}](x)\| = \|J_t(x)\| \leq \exp\Big(\int_0^t \|\nabla w_s(\eta_s(x))\|ds\Big) = \exp\Big(\int_0^t \|\nabla v_{t-s}(\eta_s(x))\|ds\Big).$$

To bound this, let $\underline{t}$ be from Lemma 23 for $\epsilon = 1$ and the r from above and set $t_2 := \max(\underline{t}, t_1, 1 - \frac{\delta r}{4CM})$. If $t \leq t_2$, then we get by the global bound of Lemma 23 for all $x \in \mathbb{R}^d$ that

$$\begin{aligned} \|\nabla[\gamma_t^{-1}](x)\| &\leq \exp\Big(\int_0^t \frac{c}{(1-t+s)^3}ds\Big) = \exp\Big(\int_0^t \frac{c}{(1-s)^3}ds\Big) \\ &= \exp\Big(\frac{c}{2(1-t)^2} - \frac{c}{2}\Big) \leq \exp\Big(\frac{c}{2(1-t_2)^2} - \frac{c}{2}\Big). \end{aligned}$$

For $t \geq t_2$, we obtain by Step 2 for $x \in B_{r_t}(a_k)$ that $\eta_s(x) \in B_r(a_k)$ for $s \leq t - t_2$ such that we can additionally apply the local bound of Lemma 23. This yields

$$\begin{aligned} \|\nabla[\gamma_t^{-1}](x)\| &\leq \exp\Big(\int_0^{t-t_2} \|\nabla v_{t-s}(\eta_s(x))\|ds + \int_{t-t_2}^t \frac{c}{(1-t+s)^3}ds\Big) \\ &\leq \exp\Big(\int_0^{t-t_2} \frac{2}{1-t+s}ds + \int_0^{t_2}\frac{c}{(1-s)^3}ds\Big) \\ &= \exp\Big(\int_{t_2}^t \frac{2}{1-s}ds + \frac{c}{2(1-t_2)^2} - \frac{c}{2}\Big) \\ &= \exp\Big(2\log(1-t_2) - 2\log(1-t) + \frac{c}{2(1-t_2)^2} - \frac{c}{2}\Big) \\ &\leq \frac{\exp\Big(\frac{c}{2(1-t_2)^2} - \frac{c}{2}\Big)}{(1-t)^2} \end{aligned}$$

Combining both estimates and setting $C_1^* \coloneqq \exp\left(\frac{c}{(1-t_2)^2} - \frac{c}{2}\right)$ and $C_2^* = \frac{r}{2}$, we obtain that $\|\nabla[\gamma_t^{-1}](x)\| \leq \frac{C_1^*}{(1-t)^2}$ for all $x \in B_{C_2^*(1-t)}(a_k)$. □

Using this lemma, we can finally show the intended result.

Proposition 25. *The curve $t \mapsto \gamma_t^{-1}(a_k)$ has finite length, i.e., $\int_0^1 \|\frac{d}{dt}\gamma_t^{-1}(a_k)\| < \infty$. In particular, the limit $\lim_{t\to 1}\gamma_t^{-1}(a_k)$ exists.*

Proof. By the chain rule and inverse function theorem we obtain

$$\frac{d}{dt}\gamma_t^{-1}(a_k) = -(\nabla\gamma_t(\gamma_t^{-1}(a_k)))^{-1}\dot{\gamma}_t(\gamma_t^{-1}(a_k)) = -(\nabla\gamma_t(\gamma_t^{-1}(a_k)))^{-1}v_t(a_k) = -\nabla[\gamma_t^{-1}](a_k)v_t(a_k).$$

Taking the norm and applying Lemma 24 gives

$$\|\frac{d}{dt}\gamma_t^{-1}(a_k)\| \leq \|\nabla[\gamma_t^{-1}](a_k)\|\|v_t(a_k)\| \leq \frac{C_1^*}{(1-t)^2}\|v_t(a_k)\|. \tag{11}$$

To bound $\|v_t(a_k)\|$ we note that $v_t(a_k) = \frac{1}{1-t}\sum_{l\neq k}\alpha_l(t,a_k)(a_l - a_k)$ which implies that $\|v_t(a_k)\| \leq (1-\alpha_k(t,a_k))\frac{M}{1-t}$ for $M = \max_{l\neq k}\|a_k - a_l\|$. Now, we know by Lemma 16 that there exist some $t_0 \in (0,1)$ and $\delta > 0$ such that for all $t \geq t_0$ it holds $1-\alpha_k(t,a_k) < \frac{n-1}{\exp(\frac{\delta}{2(1-t)^2})+n-1}$. In particular, this implies that there exists some $1 > t_1 \geq t_0$ and $C > 0$ such that for all $t \geq t_1$ it holds $1-\alpha_k(t,a_k) < (1-t)C\exp(-\frac{\delta}{2(1-t)^2})$ and therefore $\|v_t(a_k)\| \leq CM\exp(-\frac{\delta}{2(1-t)^2})$. For $t \leq t_1$ we still have the bound $\|v_t(a_k)\| \leq (1-\alpha_k(t,a_k))\frac{M}{1-t} \leq \frac{M}{1-t}$. Inserting this in (11), we get that

$$\|\frac{d}{dt}\gamma_t^{-1}(a_k)\| \leq \begin{cases} \frac{C_1^*M}{(1-t)^3} & \text{for } t \leq t_1, \\ \frac{C_1^*CM\exp(-\frac{\delta}{2(1-t)^2})}{(1-t)^2}. \end{cases}$$

Integrating over t gives

$$\int_0^1 \|\frac{d}{dt}\gamma_t^{-1}(a_k)\|dt \leq \int_0^{t_1}\frac{C_1^*M}{(1-t)^3}dt + \int_{t_1}^1 \frac{C_1^*CM\exp(-\frac{\delta}{2(1-t)^2})}{(1-t)^2}dt,$$

which is finite, since the first integrand is bounded from above by $\frac{C_1^*M}{(1-t_1)^3}$ and the second integrand is bounded from above by $\frac{2C_1^*CM}{\delta}$. □

We note that for $t \in (0,1)$ large enough it is directly clear from Proposition 4 that $\gamma_t^{-1}(a_k) \in \Gamma_k$. Due to the finite length of $t \mapsto \gamma_t^{-1}(a_k)$ this implies in the limit $\lim_{t\to 1}\gamma_t^{-1}(a_k) \in \overline{\Gamma_k}$.

Proposition 26. *Assume that $\lim_{t\to 1}\gamma_t^{-1}(a_k) \in \Gamma_k$ (and not only in $\overline{\Gamma_k}$). Then, the set Γ_k is contractible.*

Proof. We consider for $r < \frac{1}{2}\min_{l\neq k}\|a_k - a_l\|$ and the corresponding $\underline{t}$ from Proposition 4 the mapping

$$H(t,x) = \begin{cases} \gamma_t^{-1}(\phi(t,\gamma_t(x))\gamma_t(x) + (1-\phi(t,\gamma_t(x)))a_k), & \text{for } t < 1, \\ \lim_{t\to 1}\gamma_t^{-1}(a_k), & \text{for } t = 1, \end{cases}$$

where

$$\phi(t,x) = \begin{cases} 1, & \text{if } t < \underline{t} \text{ or } x \notin B_r(a_k), \\ (1-\psi(\|x-a_k\|)) + \psi(\|x-a_k\|)(\frac{1-t}{1-\underline{t}})^3, & \text{if } t \geq \underline{t} \text{ and } x \in B_r(a_k)\setminus B_{r/2}(a_k), \\ (\frac{1-t}{1-\underline{t}})^3, & \text{if } t \geq \underline{t} \text{ and } x \in B_{r/2}(a_k), \end{cases}$$

for a continuous decreasing function $\psi\colon [r/2, r] \to [0,1]$ with $\psi(r/2) = 1$ and $\psi(r) = 0$. We show that H is a homotopy between the identity on Γ_k and a constant map. It is directly clear from the definition that $H(0,x) = x$ and that $H(1,x) = \lim_{t\to 1}\gamma_t^{-1}(a_k)$ is constant. Next, we show that $H(t,x) \in \Gamma_k$ for all $(t,x) \in [0,1]\times\Gamma_k$: For $t = 1$, we know by assumption that

$H(t,x) = \lim_{t\to 1} \gamma_t^{-1}(a_k) \in \Gamma_k$. If $t < \underline{t}$ or $\|\gamma_t(x) - a_k\| \geq r$, we get that $H(t,x) = x \in \Gamma_k$. If $t \geq \underline{t}$ and $\gamma_t(x) \in B_r(a_k)$, we obtain by the definition of H and Proposition 4 that $H(t,x) \in \gamma_t^{-1}(B_r(a_k)) \subseteq \Gamma_k$. Thus, it remains to show that H is continuous on $[0,1] \times \Gamma_k$. By definition, H is on $[0,1) \times \Gamma_k$ the concatenation of continuous maps and therefore continuous. Thus, it remains to show that for $\hat{x} \in \Gamma_k$ and a sequence $(t_i, x_i)_i$ with $t_i \to 1$ and $x_i \to \hat{x}$ it holds $H(t_i, x_i) \to \lim_{t\to 1} \gamma_t^{-1}(a_k)$.

We first apply Lemma 21 to the compact set $\{x_i : i = 1, 2, ...\} \cup \{\hat{x}\} \subseteq \Gamma_k$. This yields that there exists some $t_0^* < 1$ such that $\gamma_t(x_i) \in B_{r/2}(a_k)$ for all i and $t \geq t_0^*$.

Next, we use Lemma 24 to show the convergence of $H(t_i, x_i)$. Since $t_i \to 1$, we find i_0 such that for all $i \geq i_0$ we have $t_i \geq t_0^*$. Then, we obtain that $\gamma_{t_i}(x_i) \in \overline{B_{r/2}(a_k)}$ such that $H(t_i, x_i) = \gamma_{t_i}^{-1}(a_k + (\frac{1-t_i}{1-\underline{t}})^3(\gamma_{t_i}(x_i) - a_k))$. Again since $t_i \to 1$ we can choose $i_1 \geq i_0$ such that for all $i \geq i_1$ it holds that $(\frac{1-t_i}{1-\underline{t}})^3 < \frac{2C_2^*}{r}(1-t_i)$ for C_2^* from Lemma 24. In particular, this yields that for $i \geq i_1$ it holds that $\|(\frac{1-t_i}{1-\underline{t}})^3(\gamma_{t_i}(x_i) - a_k)\| < C_2^*(1-t_i)$. Since by Lemma 24, $\gamma_{t_i}^{-1}$ is $\frac{C_1^*}{(1-t_i)^2}$-Lipschitz continuous on $B_{C_2^*(1-t_i)}(a_k)$, this implies that

$$\|H(t_i, x_i) - \gamma_{t_i}^{-1}(a_k)\| \leq \frac{C_1^*}{(1-t_i)^2}(\frac{1-t_i}{1-\underline{t}})^3\|\gamma_{t_i}(x_i) - a_k\| \leq \frac{C_1^* r(1-t_i)}{2(1-\underline{t})^3} \to 0$$

as $i \to \infty$. In particular, we have that

$$\|H(t_i, x_i) - \lim_{t\to 1} \gamma_t^{-1}(a_k)\| \leq \|H(t_i, x_i) - \gamma_{t_i}^{-1}(a_k)\| + \|\gamma_{t_i}^{-1}(a_k) - \lim_{t\to 1} \gamma_t^{-1}(a_k)\| \to 0$$

which shows that H is continuous and therefore a homotopy connecting the identity on Γ_k and a constant map. □

We would like to show that Γ_k is in fact homeomorphic to $B_1(0)$ (or equivalently to $\mathbb{R}^d$). In $d = 1$ and $d = 2$ this is true for any open and contractible set such that this directly follows from Theorem 7. However, for $d \geq 3$, we need an additional property given in the following lemma.

Lemma 27. *Let $d \geq 3$. Then, Γ_k is simply connected at infinity, i.e., for any compact set $K \subseteq \Gamma_k$ there exists another compact set $L \subset \Gamma_k$ with $K \subseteq L^\circ$ such that any loop in $\Gamma_k \setminus L$ is contractible in $\Gamma_k \setminus K$.*

Proof. Fix $r < \frac{1}{2}\min_{l\neq k} \|a_k - a_l\|$. By Lemma 21, we know that there exists t_0 such that $\gamma_t(K) \subseteq B_{r/2}(a_k)$ for all $t \geq t_0$. Moreover, we know by Proposition 4, that there exists $\underline{t}$ such that $B_r(a_k) \subset \gamma_t(\Gamma_k)$ for all $t \geq \underline{t}$. Now let $t^* = \max(t_0, \underline{t})$ and define $L = \gamma_{t^*}^{-1}(\overline{B_{r/2}(a_k)})$. Note that since γ_{t^*} is a bi-Lipschitz diffeomorphism, we have that L is compact and fulfills $L^\circ = \gamma_{t^*}^{-1}(B_{r/2}(a_k)) \supseteq K$.

Now let $f\colon S^1 \to \Gamma_k \setminus L$ be a loop, and define the map $H(x,t)\colon S^1 \times [0,1] \to \Gamma_k \setminus K$ for $t \leq \frac{1}{2}$ by

$$\begin{aligned} H(x,t) &= \gamma_{t^*}^{-1}\left(\left((1-2t) + \frac{rt}{\|\gamma_{t^*}(f(x))\|}\right)\gamma_{t^*}(f(x)) + 2ta_k\right) \\ &\in \gamma_{t^*}^{-1}\left(B_r(a_k) \setminus B_{r/2}(a_k)\right) \subseteq \Gamma_k \setminus K. \end{aligned}$$

Note that by definition, for $t = \frac{1}{2}$, we have that $\gamma_{t^*}(H(x,t)) = \frac{r\gamma_{t^*}(f(x))}{2\|\gamma_{t^*}(f(x))\|} \in \partial B_{r/2}(a_k)$. Further, we know since $d \geq 3$ that $\partial B_{r/2}(a_k)$ is simply connected. Thus, we know that there exists a homotopy $\tilde{H}\colon S^1 \times [0,1] \to \partial B_{r/2}$ such that $\tilde{H}(x,0) = \frac{r\gamma_{t^*}(f(x))}{2\|\gamma_{t^*}(f(x))\|}$ and such that $\tilde{H}(x,1)$ is constant. Thus, we can define for $t \geq \frac{1}{2}$

$$H(x,t) = \gamma_{t^*}^{-1}(\tilde{H}(x, 2t-1)) \in \gamma_{t^*}^{-1}(\partial B_{r/2}(a_k)) \subseteq \Gamma_k \setminus K.$$

In summary, it holds by definition that $H(x,t) \in \Gamma_k \setminus K$, that $H(x,1)$ is constant and that H is continuous as the concatenation of continuous maps. Consequently H is a homotopy and since K was chosen arbitrarily, we obtain that Γ_k is simply connected at infinity. □

Now we can conclude that Γ_k is topologically equivalent to a ball.

Corollary 28. *Assume that* $\lim_{t\to 1}\gamma_t^{-1}(a_k)\in\Gamma_k$. *Then,* Γ_k *is homeomorphic to* $\mathbb{R}^d$ *and to* $B_1(0)$.

Proof. For $d=1$ and $d=2$ any open and contractible subset of $\mathbb{R}^d$ is homeomorphic to $\mathbb{R}^d$ (and thus also to $B_1(0)$) such that the claim follows from Proposition 26. For $d\geq 3$, we additionally know from Proposition 26 and Lemma 27 that Γ_k is simply connected at infinity such that the claim follows from the fact that any set which is open, contractible and simply connected at infinity is homeomorphic to $\mathbb{R}^d$ (and to $B_1(0)$), see Brin and Thickstun [1989] for $d=3$, Freedman and Quinn [1990] for $d=4$ and Stallings [1962] for $d>4$. □

D Proofs for Geometry of the FM Cells

D.1 Proofs for the Four-Point Counterexample

Proof of Lemma 9. By symmetry it suffices to show the claim for $k=3$ for which the half-line is $L_3 \coloneqq \{(-c,0) : c\geq 0\}$ (see the red line in Figure 5 (c)).

First we show that for any initialization $x\in L_3$, its image under the flow map satisfies $\gamma_t(x)\in L_3$. First, note that by definition it holds for $x=(b,0)$ with $b\in\mathbb{R}$ that $\alpha_1(t,x)=\alpha_2(t,x)$. In particular, we get that $v_t(b,0)=(a,0)$ for some $a\in\mathbb{R}$. Moreover, we get for any t that $\alpha_1(t,(0,0))=\alpha_2(t,(0,0))=\alpha_3(t,(0,0))$ such that $v_t(0,0)=\alpha_1(t,(0,0))(a_1+a_2+a_3)=0$. From this, we can conclude for $b\geq 0$ that $\gamma_t(-b,0)\in L_3$.

Now we prove that for $x\in L_3$, $\gamma_t(x)$ actually converges to $a_4=(0,0)$ when $t\to 1$. Let $b_0\leq 0$ and define b_t such that $(b_t,0)=\gamma_t(b_0,0)$. We show that b_t converges to zero for $t\to 1$ in three steps.

Step 1: We show that for all $\epsilon>0$ there exists some time t_0 such that $b_t\geq -\frac{1}{2}-\epsilon$ for all $t\geq t_0$. To this end, we observe

$$\begin{aligned}(1-t)\dot{b}_t &= (\alpha_1(t,(b_t,0))+\alpha_2(t,(b_t,0)))(-\frac{1}{2}-b_t)\\ &\quad+\alpha_3(t,(b_t,0))(1-b_t)-\alpha_4(t,(b_t,0))b_t\\ &\geq \sum_{k=1}^4 \alpha_k(t,(b_t,0))(-\frac{1}{2}-b_t) = -\frac{1}{2}-b_t.\end{aligned}$$

Applying the Gronwall lemma on $\dot{b}_t\geq -\frac{\frac{1}{2}+b_t}{1-t}$ gives $b_t\geq (1-t)b_0-\frac{t}{2}\xrightarrow[t\to 1]{}-\frac{1}{2}$, which shows step 1.

Step 2: Next, we show that for $\epsilon=0.1$ there exists some t_1 such that $v_t(b,0)\geq\frac{c}{1-t}$ for some $c>0$ and all $b\in-\frac{1}{2}+[-\epsilon,\epsilon]$. Using the formula for v_t from the previous step, we get that $v_t(b,0)=(w_t(b),0)$ with

$$\begin{aligned}(1-t)w_t(b) &\geq (1-\alpha_4(t,(b,0)))(-\frac{1}{2}-b)-\alpha_4(t,(b,0))b\\ &= -\frac{1-\alpha_4(t,(b,0))}{2}-b\geq\frac{1}{2}-\epsilon-\frac{1-\alpha_4(t,(b,0))}{2}\end{aligned}$$

Observing that $\alpha_4(t,(b,0))$ converges to 1 uniformly on $\{(c,0):c\in-\frac{1}{2}+[-\epsilon,\epsilon]\}$ yields the claim of step 2.

Step 3: Set $t_2=\max(t_0,t_1,\underline{t}_r)$ with $\underline{t}_r$ from Proposition 4 and $r=\frac{1}{2}-\epsilon$. We show that there exists some $t\geq t_2$ such that $b_t\geq-\frac{1}{2}+\epsilon$. We get from step 1 that $b_{t_2}\geq-\frac{1}{2}-\epsilon$. To this end, we build a contradiction. Assume that $b_t<-\frac{1}{2}+\epsilon$ for all $t\in[t_2,1)$. Then, we get by step 1 that $b_t\in-\frac{1}{2}+[-\epsilon,\epsilon]$. By step 2 this implies that $\dot{b}_t\geq\frac{c}{1-t}$ for some $c>0$. Applying the Gronwall lemma on this leads for $t\in[t_2,1)$ to $b_t\geq b_{t_2}+c\log(\frac{1-t_2}{1-t})$. This converges to ∞ for $t\to 1$ which contradicts the assumption that $b_t<-\frac{1}{2}+\epsilon$.

Conclusion: By step 3, we know that there exists some $t\geq\underline{t}_r$ such that $(b_t,0)\in B_r(0,0)$. Applying Proposition 4 yields $b_1=0$ such that $(b_0,0)\in\Gamma_4$. □

Proof of Lemma 10. By symmetry it suffices to show the claim for $k = 3$ for which $a_k = (1, 0)$. For $b_0 \in \mathbb{R}$, we consider the curve $\gamma_t(b_0, 0) = (b_t, 0)$. We note that for any t we can estimate

$$\begin{aligned}(1-t)\dot{b}_t &= (\alpha_1(t,(b_t,0)) + \alpha_2(t,(b_t,0)))(-\frac{1}{2} - b_t) \\ &\quad + \alpha_3(t,(b_t,0))(1-b_t) - \alpha_4(t,(b_t,0))b_t \\ &\begin{cases} \geq -\frac{1}{2} - b_t \\ \leq 1 - b_t \end{cases}\end{aligned}$$

With the Gronwall lemma for both estimates this implies that

$$(1-t)b_0 - \frac{t}{2} \leq b_t \leq (1-t)b_0 + t. \tag{12}$$

Now we note that by Corollary 18 and Proposition 4, we deduce that for $r = 1/4$ there exists a time t_0 such that, for all $t \geq t_0$, the implication

$$1 - r \leq b_t \leq 1 + r \quad \Rightarrow \quad 1 - r \leq b_s \leq 1 + r \text{ for all } s \geq t$$

holds, and moreover $b_1 = 1$. Now choose $c_0 = \frac{1-r}{1-t_0} + \frac{t_0}{2(1-t_0)}$ such that $(1-t_0)c_0 - \frac{t_0}{2} = 1 - r$. Then it holds by (12) for $b_0 \geq c_0$ that $b_{t_0} \geq (1-t_0)b_0 - \frac{t_0}{2} \geq 1 - r$. In particular, we have that $b_t \geq 1 - r$ for all $t \geq t_0$. At the same time the other estimate in (12) gives that there exists $t \geq t_0$ such that $b_t \leq 1 + r$. Together, we obtain that there exists $t \geq t_0$ such that $1 - r \leq b_t \leq 1 + r$ such that $b_1 = 1$. □

Proof of Theorem 8 using Lemma 9 and Lemma 10. The first two parts are already explained in the main text.

Non-convexity: The convex hull of Γ_4 fulfills

$$\operatorname{conv}(\Gamma_4) \supseteq \operatorname{conv}(\{(-c\cos(\frac{2k\pi}{3}), -c\sin(\frac{2k\pi}{3})) : c \geq 0,\ k = 1,2,3\}) = \mathbb{R}^2.$$

Since Γ_k is non-empty for $k = 1, 2, 3$, we obtain that Γ_4 must be non-convex.

Non-affine boundary: Fix an arbitrary $k \in \{1, 2, 3\}$ and assume that the boundary between Γ_4 and Γ_k is affine. Since these cells are connected and open, an affine common boundary places them in opposite half-planes. Then, there exists some $0 \neq p \in \mathbb{R}^2$ and $q \in \mathbb{R}$ such that

$$\Gamma_k \subseteq \{x \in \mathbb{R}^2 : \langle p, x\rangle + q \leq 0\}, \qquad \Gamma_4 \subseteq \{x \in \mathbb{R}^2 : \langle p, x\rangle + q \geq 0\}.$$

By Lemma 9, we know that $-ca_l \in \Gamma_4$ for $l = 1, 2, 3$ which implies with $c \to \infty$ that $\langle p, a_l\rangle \leq 0$. Since $a_1 + a_2 + a_3 = 0$, we obtain that

$$0 = \langle p, a_1 + a_2 + a_3\rangle = \langle p, a_1\rangle + \langle p, a_2\rangle + \langle p, a_3\rangle.$$

Because all three terms are non-positive, this implies that $\langle p, a_1\rangle = \langle p, a_2\rangle = \langle p, a_3\rangle = 0$. Given that a_1, a_2 and a_3 span $\mathbb{R}^2$, this implies $p = 0$, which is a contradiction. □

D.2 Proof of the Non-Monotonicity

Proof of Lemma 11. Consider the two points $x_1 = c_1 a_3 = (c_1, 0)$ with $c_1 > c_0$ where c_0 is given by Lemma 10 and $x_2 = (-c_2\cos(-\frac{4\pi}{3}), -c_2\sin(\frac{4\pi}{3})) = \frac{c_2}{2}(1, \sqrt{3})$ such that $\frac{c_2}{2} - c_1 > 0$.

- By Lemma 10, $\gamma_t(x_1) \xrightarrow[t\to 1]{} a_3$ and by Corollary 18, there exists $t_0 > 0$ such that for $t \geq t_0$, $\gamma_t(x_1) = (b_t, 0)$ with $b_t \geq 1 - r$.
- By Lemma 9, $\gamma_t(x_2) \xrightarrow[t\to 1]{} a_4 = (0, 0)$ and by Corollary 18, there exists $t_1 > 0$ such that for $t \geq t_1$, $\gamma_t(x_2) = \frac{c_t}{2}(1, \sqrt{3})$ and $\|\gamma_t(x_2)\| = c_t \leq r$

We then get for $t \geq t_m := \max(t_0, t_1)$,

$$\begin{aligned}\langle \gamma_t(x_2) - \gamma_t(x_1), x_2 - x_1 \rangle &= \Big(\frac{c_t}{2} - b_t\Big)\Big(\frac{c_2}{2} - c_1\Big) + \frac{3c_t c_2}{4} \\ &\leq \Big(\frac{r}{2} - (1-r)\Big)\Big(\frac{c_2}{2} - c_1\Big) + \frac{3rc_2}{4} \\ &= -\frac{1}{2}\Big((1-3r)c_2 - (2-3r)c_1\Big).\end{aligned}$$

Then for any $c_2 = Kc_1$ with $K > \frac{2-3r}{1-3r}$, we get

$$\langle \gamma_t(x_2) - \gamma_t(x_1), x_2 - x_1 \rangle < -\frac{c_1}{2}\Big((1-3r)K - (2-3r)\Big).$$

□

Proof of Lemma 12. Let $x_1, x_2 \in \mathbb{R}^d$ and $c > 0$ be given by Lemma 11, so that

$$\langle \gamma(x_2) - \gamma(x_1), x_2 - x_1 \rangle < -c.$$

By Proposition 3, for each fixed $x_i \in \{x_1, x_2\}$, we have $\gamma^\varepsilon(x_i) \to \gamma(x_i)$ as $\varepsilon \to 0$. Hence, for all $\varepsilon > 0$ small enough,

$$\langle \gamma^\varepsilon(x_2) - \gamma^\varepsilon(x_1), x_2 - x_1 \rangle < -\frac{c}{2}.$$

Set $h := x_2 - x_1$. Since γ^ε is smooth, the fundamental theorem of calculus gives

$$\langle \gamma^\varepsilon(x_2) - \gamma^\varepsilon(x_1), h \rangle = \int_0^1 \langle J_{\gamma^\varepsilon}(x_1 + sh)h, h \rangle \, ds.$$

Equivalently,

$$\langle \gamma^\varepsilon(x_2) - \gamma^\varepsilon(x_1), h \rangle = \frac{1}{2}\int_0^1 h^\top \left(J_{\gamma^\varepsilon}(x_1 + sh) + J_{\gamma^\varepsilon}(x_1 + sh)^\top\right) h \, ds < -\frac{c}{2}.$$

Therefore there exists $s_0 \in [0,1]$ such that

$$h^\top \left(J_{\gamma^\varepsilon}(x_1 + s_0 h) + J_{\gamma^\varepsilon}(x_1 + s_0 h)^\top\right) h < -c.$$

It follows that

$$\lambda_{\min}\left(J_{\gamma^\varepsilon}(x_1 + s_0 h) + J_{\gamma^\varepsilon}(x_1 + s_0 h)^\top\right) < -\frac{c}{\|h\|^2}.$$

Thus, setting $c' := c/\|x_2 - x_1\|^2 > 0$, we obtain

$$\inf_x \lambda_{\min}\left(J_{\gamma_1^\varepsilon}(x) + J_{\gamma_1^\varepsilon}(x)^\top\right) < -c'.$$

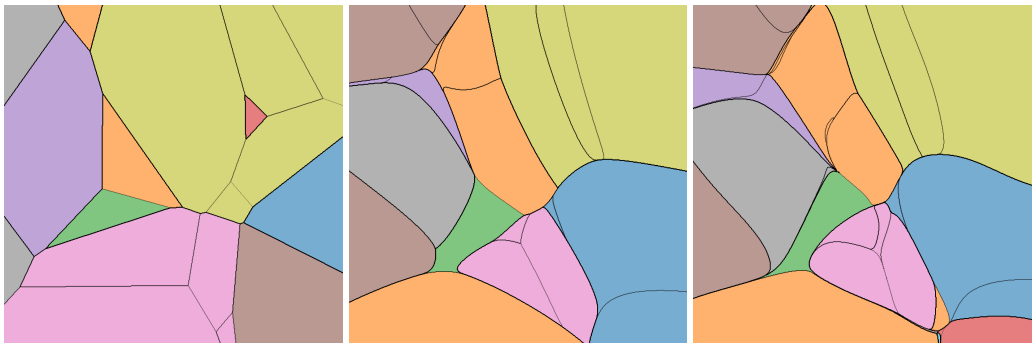
□

E Higher-Dimensional Training Example (MNIST)

We complement the two-dimensional examples from the main text with a training experiment on MNIST. We train a VAE encoder, parameterized by a two-layer MLP, to embed MNIST images from dimension 784 into a latent space of dimension 16. We then select 100 latent training points $(a_k)_{k=1}^{100}$, with 10 points from each digit class, and use them as the atoms of the semi-discrete target. On this dataset, we compare the semi-discrete OT Laguerre cells, the exact Flow Matching cells computed from the closed-form velocity, and the cells induced by a trained Flow Matching velocity field, parameterized by a 2-layer MLP with 256 hidden neurons.

To visualize these 16-dimensional cells, we restrict the comparison to a two-dimensional affine slice spanned by the first two PCA directions of the selected latent points. For the trained Flow Matching model, terminal points do not necessarily coincide exactly with the atoms a_k, reflecting the usual generalization and approximation effects of the learned velocity. We therefore assign each grid point to the atom closest to its terminal image under the learned flow.

Figure 8 shows that the qualitative differences between OT and FM persist in this higher-dimensional latent setting. Compared with Laguerre cells, the FM cells have curved boundaries, are often non-convex, and display different locations and neighborhood relations. Moreover, the trained FM cells closely resemble the exact FM cells, suggesting that even a simple learned velocity can recover the main geometric features of the exact semi-discrete Flow Matching tessellation.



(a) OT (b) FM with exact velocity (c) FM with learned velocity

Figure 8: Semi-discrete assignment cells Γ_k for 100 MNIST training points in a VAE latent space of dimension 16. Cells are visualized on a two-dimensional affine slice spanned by the first two PCA directions. Cells are colored according to the MNIST digit label of the corresponding target point. We compare (a) Optimal Transport Laguerre cells, (b) Flow Matching cells induced by the closed-form velocity, and (c) Flow Matching cells obtained from a neural-network approximation of the velocity. In panel (c), each grid point is assigned the color of the target point closest to its image under the terminal flow.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract describes the claims and our contributions.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: We include a limitations section in the conclusion.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.

- The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
- The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
- If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
- While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

   Justification: We include a proof for all theoretical claims.

   Guidelines:

   - The answer [N/A] means that the paper does not include theoretical results.
   - All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
   - All assumptions should be clearly stated or referenced in the statement of any theorems.
   - The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
   - Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
   - Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [N/A]

   Justification: We do not include major experiments (despite for illustrations). The only larger scale experiment is done in Appendix and is clearly described for reproduction.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.

- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [N/A]

   Justification: We do not include experiments (despite for illustrations).

   Guidelines:

   - The answer [N/A] means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (https://neurips.cc/public/guides/CodeSubmissionPolicy) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (https://neurips.cc/public/guides/CodeSubmissionPolicy) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [N/A]

   Justification: We do not include experiments (despite for illustrations).

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [N/A]

   Justification: We do not include experiments (despite for illustrations).

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.
   - It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
   - For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
   - If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [N/A]

   Justification: We do not include experiments (despite for illustrations).

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics https://neurips.cc/public/EthicsGuidelines?

   Answer: [Yes]

   Justification: We reviewed and followed the NeurIPS Code of Ethics.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.

- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [N/A]

    Justification: Our paper considers foundational research and we do not see any direct path to societal impacts.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
    - The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
    - The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
    - If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: Our paper does not pose such risks.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [N/A]

Justification: The paper does not use existing assets.

Guidelines:

- The answer [N/A] means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [N/A]

Justification: The paper does not release new assets.

Guidelines:

- The answer [N/A] means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [N/A]

Justification: The paper does not involve crowdsourcing nor research with human subjects.

Guidelines:

- The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [N/A]

Justification: The paper does not involve crowdsourcing nor research with human subjects.

Guidelines:

- The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

Answer: [N/A]

Justification: The core method development in this research does not involve LLMs as any important, original, or non-standard components.

Guidelines:

- The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.