Introduction to Rectified Flow

Xuda Ye

Purdue University

November 21, 2025

In this note, we introduce Rectified Flow for optimal transport and generative modeling. We first review the fundamentals of Flow Matching [1], and subsequently derive key convergence results for Rectified Flow in [2]. While Flow Matching provides a general framework for continuous normalizing flows, Rectified Flow distinguishes itself by enforcing straight flow trajectories. This geometric property significantly minimizes the transport cost and approximates the optimal transport map between data distributions. Consequently, Rectified Flow mitigates the discretization error inherent in numerical integration, enabling high-quality sampling with fewer time steps.

1 Problem Setup

Let μ and ν be probability distributions on $\mathbb{R}^d$ (accessible via empirical samples). The goal of Flow Matching is to determine a velocity field $u(x,t)$ defined on $\mathbb{R}^d \times [0,1]$ such that μ is transported to ν. Specifically, the probability density path $p_t(x)$ satisfying the continuity equation

$$\partial_t p_t(x) + \nabla \cdot \big(u(x,t)p_t(x)\big) = 0 \tag{1}$$

must adhere to the marginal constraints $p_0 = \mu$ and $p_1 = \nu$. This formulation is equivalent to finding an admissible solution $u(x,t)$ to a control problem.

Furthermore, Rectified Flow aims to identify an optimal solution $u(x,t)$ that minimizes the L^2 transport cost:

$$\min_u \int_0^1 \int_{\mathbb{R}^d} p_t(x)|u(x,t)|^2 \mathrm{d}x\mathrm{d}t \tag{2}$$

subject to the continuity equation (1) and the marginal constraints $p_0 = \mu$ and $p_1 = \nu$. As indicated by Brenier's theorem [3], this optimal control problem is equivalent to solving the optimal transport problem between μ and ν:

$$\min_{\pi \in \Gamma(\mu,\nu)} \mathbb{E}_{(X_0,X_1)\sim\pi}\big[|X_1 - X_0|^2\big]. \tag{3}$$

Specifically, the optimal values of (2) and (3) coincide and are equal to $\mathcal{W}_2^2(\mu,\nu)$, where $\mathcal{W}_2(\cdot,\cdot)$ denotes the Wasserstein-2 distance. This relationship constitutes the foundation for Rectified Flow.

2 Flow Matching

Let $\pi \in \Pi(\mu, \nu)$ denote an arbitrary coupling between the distributions μ and ν. The Flow Matching framework constructs an admissible solution to (1) by minimizing the following L^2 functional:

$$\mathcal{L}[u] = \int_0^1 \mathbb{E}_{(X_0,X_1)\sim\pi}\Big[\big|X_1 - X_0 - u(X_t, t)\big|^2\Big]\mathrm{d}t. \tag{4}$$

Here, the random variables (X_0, X_1) are sampled from the coupling π, and $X_t = tX_1 + (1-t)X_0$ represents the linear interpolation path. The minimizer of (4) is given by the conditional expectation

$$u(x, t) = \mathbb{E}_{(X_0,X_1)\sim\pi}\big[X_1 - X_0\big|X_t = x\big]. \tag{5}$$

(Some papers, they like to write $u(x,t) = \mathbb{E}[\dot{X}_t|X_t = x]$. I don't like it, okay? It's a DIASTER!)

We now provide a more explicit expression for $u(x, t)$. Let $p_t(x)$ denote the marginal distribution of X_t induced by the coupling $(X_0, X_1) \sim \pi$, defined formally as

$$p_t(x) = \mathbb{E}_{(X_0,X_1)\sim\pi}\big[\delta(X_t - x)\big]. \tag{6}$$

Consequently, the velocity field $u(x, t)$ in (5) can be equivalently expressed as

$$u(x, t) = \frac{1}{p_t(x)}\mathbb{E}_{(X_0,X_1)\sim\pi}\big[(X_1 - X_0)\delta(X_t - x)\big]. \tag{7}$$

We now state the following result regarding the continuity equation (1). This result confirms that the velocity field $u(x, t)$ in (5) generates the probability flow $\{p_t(x)\}_{0\leqslant t\leqslant 1}$, thereby transporting the distribution from $p_0 = \mu$ to $p_1 = \nu$.

Theorem 1 *Let μ and ν be probability distributions on $\mathbb{R}^d$, and let $\pi \in \Pi(\mu, \nu)$ be a coupling. Then, the velocity field $u(x, t)$ defined in (5) generates the probability flow $\{p_t(x)\}_{0\leqslant t\leqslant 1}$ in (6). Equivalently, the continuity equation holds:*

$$\partial_t p_t(x) + \nabla \cdot \big(u(x, t)p_t(x)\big) = 0,$$

subject to the marginal constraints $p_0 = \mu$ and $p_1 = \nu$.

Proof We verify the continuity equation (1) in the distributional sense. Let $\phi \in C_c^\infty(\mathbb{R}^d)$ be a smooth test function with compact support. The weak formulation of (1) requires

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} p_t(x)\phi(x)\mathrm{d}x = \int_{\mathbb{R}^d} p_t(x)u(x, t)\cdot\nabla\phi(x)\mathrm{d}x. \tag{8}$$

We evaluate the left-hand side (LHS) and right-hand side (RHS) of (8) separately.

For the LHS, utilizing the definition of $p_t(x)$ and the property of the Dirac delta, we obtain:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} p_t(x)\phi(x)\mathrm{d}x &= \mathbb{E}_{(X_0,X_1)\sim\pi}\left[\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d}\delta\big(x - tX_1 - (1-t)X_0\big)\phi(x)\mathrm{d}x\right] \\
&= \mathbb{E}_{(X_0,X_1)\sim\pi}\left[\frac{\mathrm{d}}{\mathrm{d}t}\phi\big(tX_1 + (1-t)X_0\big)\right] \\
&= \mathbb{E}_{(X_0,X_1)\sim\pi}\Big[(X_1 - X_0)\cdot\nabla\phi\big(tX_1 + (1-t)X_0\big)\Big].
\end{aligned}$$

For the RHS, we substitute the expression

$$p_t(x)u(x,t) = \mathbb{E}_{(X_0,X_1)\sim\pi}\Big[(X_1 - X_0)\delta\big(x - tX_1 - (1-t)X_0\big)\Big]$$

and apply Fubini's theorem:

$$\begin{aligned}\int_{\mathbb{R}^d} p_t(x)u(x,t)\cdot\nabla\phi(x)\mathrm{d}x &= \int_{\mathbb{R}^d} \mathbb{E}_{(X_0,X_1)\sim\pi}\Big[(X_1 - X_0)\delta\big(x - tX_1 - (1-t)X_0\big)\Big]\cdot\nabla\phi(x)\mathrm{d}x \\ &= \mathbb{E}_{(X_0,X_1)\sim\pi}\left[(X_1 - X_0)\cdot\int_{\mathbb{R}^d}\delta\big(x - tX_1 - (1-t)X_0\big)\nabla\phi(x)\mathrm{d}x\right] \\ &= \mathbb{E}_{(X_0,X_1)\sim\pi}\Big[(X_1 - X_0)\cdot\nabla\phi\big(tX_1 + (1-t)X_0\big)\Big].\end{aligned}$$

The equality of the LHS and RHS establishes the validity of the continuity equation (1). ■

Illustrative numerical experiments utilizing Flow Matching are detailed in the note *Handbook of Generative Models for Mathematicians*, available at https://xuda-ye.com/wp-content/uploads/2025/11/handbook-of-generative-models-1.pdf.

3 Rectified Flow

Let π be a coupling between the distributions μ and ν on $\mathbb{R}^d$. Once the velocity field $u(x,t)$ is established as in (5), we define a process $\{Z_t\}_{0\leqslant t\leqslant 1}$ such that $Z_0 \sim \mu$ and Z_t satisfies the ordinary differential equation (ODE):

$$\frac{\mathrm{d}}{\mathrm{d}t}Z_t = u(Z_t,t), \quad t\in[0,1]. \tag{9}$$

According to Theorem 1, the marginal distribution of Z_t is exactly $p_t(x)$, which coincides with the law of X_t. This induced dynamic $\{Z_t\}_{0\leqslant t\leqslant 1}$ is termed the **rectified flow** of the coupling π, effectively transporting the distribution from μ to ν. The pair of random variables (Z_0, Z_1) constitutes a new coupling between μ and ν, referred to as the **rectified coupling** of π.

For notational convenience, let π' denote the joint distribution of (Z_0, Z_1). We formally define the rectified flow and rectified coupling operations as:

$$\{Z_t\}_{0\leqslant t\leqslant 1} = \texttt{RectFlow}(\pi), \qquad \pi' = \texttt{RectCoup}(\pi) = \mathrm{Law}(Z_0, Z_1).$$

The fundamental distinction between the rectified flow $\{Z_t\}_{0\leqslant t\leqslant 1}$ and the original process $\{X_t\}_{0\leqslant t\leqslant 1}$ lies in the non-crossing property of the trajectories. Specifically, in the standard Flow Matching framework, the linear interpolation paths $X_t = tX_1 + (1-t)X_0$ may intersect for different realizations of the pairs (X_0, X_1). In contrast, the trajectories generated by the ODE solution $\{Z_t\}_{0\leqslant t\leqslant 1}$ in (9) are non-intersecting for distinct initial values Z_0, owing to the uniqueness of ODE solutions. Intuitively, this rectification "rewires" the transport to eliminate crossings, yielding trajectories that are straighter and geometrically more efficient.

The geometric notion of straightness admits a quantitative characterization. A fundamental property is the monotonicity of the transport cost; specifically, the rectified flow induces a transport cost no greater than that of the original coupling.

Theorem 2 *Let π be a coupling between the probability distributions μ and ν on $\mathbb{R}^d$, and let $\pi' =$ $\texttt{RectCoup}(\pi)$ denote the rectified coupling. Then, the following inequality holds:*

$$\mathbb{E}_{(Z_0,Z_1)\sim\pi'}\big[|Z_1 - Z_0|^2\big] \leqslant \mathbb{E}_{(X_0,X_1)\sim\pi}\big[|X_1 - X_0|^2\big].$$

Proof Invoking the definition of the ODE (9) and applying Cauchy's inequality, we obtain

$$\begin{aligned}
\mathbb{E}_{(Z_0,Z_1)\sim\pi'}\big[|Z_1 - Z_0|^2\big] &= \mathbb{E}_{Z_0\sim\mu,\dot{Z}_t=u(Z_t,t)}\big[|Z_1 - Z_0|^2\big] \\
&= \mathbb{E}_{Z_0\sim\mu,\dot{Z}_t=u(Z_t,t)}\left[\left|\int_0^1 u(Z_t,t)\mathrm{d}t\right|^2\right] \\
&\leqslant \int_0^1 \mathbb{E}_{Z_0\sim\mu,\dot{Z}_t=u(Z_t,t)}\big[|u(Z_t,t)|^2\big]\mathrm{d}t. \qquad (10)
\end{aligned}$$

According to Theorem 1, the stochastic process $\{Z_t\}_{0\leqslant t\leqslant 1}$, generated by $Z_0 \sim \mu$ and $\dot{Z}_t = u(Z_t, t)$, admits the same marginal distributions as $\{X_t\}_{0\leqslant t\leqslant 1}$, where $(X_0, X_1) \sim \pi$ and $X_t = tX_1 + (1-t)X_0$ denotes the linear interpolation. Consequently, from (10) we obtain the inequality:

$$\mathbb{E}_{(Z_0,Z_1)\sim\pi'}\big[|Z_1 - Z_0|^2\big] \leqslant \int_0^1 \mathbb{E}_{(X_0,X_1)\sim\pi}\big[|u(X_t,t)|^2\big]\mathrm{d}t. \qquad (11)$$

On the other hand, the velocity $u(x, t)$ given by the conditional expectation (7):

$$u(x,t) = \frac{1}{p_t(x)}\mathbb{E}_{(X_0,X_1)\sim\pi}\big[(X_1 - X_0)\delta(X_t - x)\big]$$

satisfies Cauchy's inequality:

$$\begin{aligned}
p_t(x)|u(x,t)|^2 &\leqslant \frac{1}{p_t(x)}\mathbb{E}_{(X_0,X_1)\sim\pi}\big[|X_1 - X_0|^2\delta(X_t - x)\big]\cdot\mathbb{E}_{(X_0,X_1)\sim\pi}\big[\delta(X_t - x)\big] \\
&= \mathbb{E}_{(X_0,X_1)\sim\pi}\big[|X_1 - X_0|^2\delta(X_t - x)\big].
\end{aligned}$$

Integrating over $x \in \mathbb{R}^d$ yields

$$\int_{\mathbb{R}^d} p_t(x)|u(x,t)|^2\mathrm{d}x \leqslant \mathbb{E}_{(X_0,X_1)\sim\pi}\big[|X_1 - X_0|^2\big]. \qquad (12)$$

Since $p_t(x)$ represents the marginal distribution of X_t, (12) immediately implies

$$\mathbb{E}_{(X_0,X_1)\sim\pi}\big[|u(X_t,t)|^2\big] \leqslant \mathbb{E}_{(X_0,X_1)\sim\pi}\big[|X_1 - X_0|^2\big], \quad \forall t \in [0,1]. \qquad (13)$$

Finally, integrating (13) over the interval $t \in [0, 1]$ and combining the result with (11), we obtain

$$\mathbb{E}_{(Z_0,Z_1)\sim\pi'}\big[|Z_1 - Z_0|^2\big] \leqslant \mathbb{E}_{(X_0,X_1)\sim\pi}\big[|X_1 - X_0|^2\big]. \qquad \blacksquare$$

For the ODE flow $\{Z_t\}_{0\leqslant t\leqslant 1}$ governed by (9), we introduce the following functional to quantify its *straightness*:

$$S\big(\{Z_t\}_{0\leqslant t\leqslant 1}\big) := \int_0^1 \mathbb{E}_{Z_0\sim\mu,\dot{Z}_t=u(Z_t,t)}\Big[\big|Z_1 - Z_0 - u(Z_t,t)\big|^2\Big]\mathrm{d}t. \qquad (14)$$

As $S\big(\{Z_t\}_{0\leqslant t\leqslant 1}\big)$ approaches zero, the coupling π' associated with (Z_0, Z_1) admits transport via an ODE with linear trajectories, thereby aligning with the properties of optimal transport.

Utilizing the straightness functional $S(\{Z_t\}_{0\leqslant t\leqslant 1})$, we derive a sharpened version of Theorem 2:

Corollary 3 *Let π be a coupling between the probability distributions μ and ν on $\mathbb{R}^d$, and let $\{Z_t\}_{0\leqslant t\leqslant 1} = \texttt{RectFlow}(\pi)$ denote the rectified flow. Then, the following inequality holds:*

$$\mathbb{E}_{Z_0\sim\mu,\dot{Z}_t=u(Z_t,t)}\left[|Z_1 - Z_0|^2\right] + S(\{Z_t\}_{0\leqslant t\leqslant 1}) \leqslant \mathbb{E}_{(X_0,X_1)\sim\pi}\left[|X_1 - X_0|^2\right].$$

Proof Invoking the fundamental identity

$$Z_1 - Z_0 = \int_0^1 u(Z_t, t)\mathrm{d}t,$$

we obtain the following algebraic decomposition:

$$\int_0^1 \left|Z_1 - Z_0 - u(Z_t,t)\right|^2\mathrm{d}t + \left|\int_0^1 u(Z_t,t)\mathrm{d}t\right|^2 = \int_0^1 |u(Z_t,t)|^2\mathrm{d}t.$$

Taking the expectation with respect to the rectified flow (where $Z_0 \sim \mu$ and $\dot{Z}_t = u(Z_t, t)$) yields

$$S(\{Z_t\}_{0\leqslant t\leqslant 1}) + \mathbb{E}_{Z_0\sim\mu,\dot{Z}_t=u(Z_t,t)}\left[\left|\int_0^1 u(Z_t,t)\mathrm{d}t\right|^2\right] = \int_0^1 \mathbb{E}_{Z_0\sim\mu,\dot{Z}_t=u(Z_t,t)}\left[|u(Z_t,t)|^2\right]\mathrm{d}t. \tag{15}$$

The equality (15) constitutes a strengthened version of the inequality derived in (10). Combining (15) with the bound established in (13) yields the desired conclusion. ■

Theorem 2 establishes that the transport cost is non-increasing after each rectification step. This monotonicity motivates the use of the iterative rectification procedure to approximate the optimal transport map:

Algorithm 1: Rectified Flow [2]

Input: Coupling π between the distributions μ and ν in $\mathbb{R}^d$.
Output: A sequence of rectified flows $\{Z_t^{(k)}\}_{0\leqslant t\leqslant 1}$ for $k = 1, 2, \ldots$.
Set the initial coupling $\pi^{(0)} = \pi$.
for $k = 1, 2, \ldots$ **do**
  Train the velocity field $u^{(k)}(x, t)$ on $\mathbb{R}^d \times [0, 1]$ by minimizing the functional:

$$\mathcal{L}_k[u] = \int_0^1 \mathbb{E}_{(X_0,X_1)\sim\pi^{(k-1)}}\left[\left|X_1 - X_0 - u(X_t,t)\right|^2\right]\mathrm{d}t.$$

  Generate the endpoint samples $Z_1^{(k)}$ by simulating the ODE from $Z_0^{(k)} \sim \mu$.
  Update the coupling $\pi^{(k)} = \mathrm{Law}(Z_0^{(k)}, Z_1^{(k)})$.

Theoretically, $Z_1^{(k)}$ strictly follows the distribution ν; consequently, the definition of the rectified coupling $\pi^{(k)}$ is independent of the direction of integration. In practice, one may choose to solve the ODE in either the forward or reverse direction.

Applying Corollary 3 to the Rectified Flow algorithm yields the recursive inequality:

$$\mathbb{E}_{(Z_0^{(k)},Z_1^{(k)})}\left[\left|Z_1^{(k)} - Z_0^{(k)}\right|^2\right] + S\left(\{Z_t^{(k)}\}_{0\leqslant t\leqslant 1}\right) \leqslant \mathbb{E}_{(Z_0^{(k-1)},Z_1^{(k-1)})}\left[\left|Z_1^{(k-1)} - Z_0^{(k-1)}\right|^2\right].$$

Summing these inequalities over $k = 1, \dots, K$, we obtain the global bound:

$$\mathbb{E}_{(Z_0^{(K)}, Z_1^{(K)}) \sim \pi^{(K)}} \left[\left| Z_1^{(K)} - Z_0^{(K)} \right|^2 \right] + \sum_{k=1}^{K} S\big(\{Z_t^{(k)}\}_{0 \leqslant t \leqslant 1}\big) \leqslant \mathbb{E}_{(X_0, X_1) \sim \pi} \left[|X_1 - X_0|^2 \right].$$

In particular, the boundedness of the sum implies

$$\lim_{K \to \infty} S\big(\{Z_t^{(K)}\}_{0 \leqslant t \leqslant 1}\big) = 0,$$

which indicates that the rectified flow trajectories asymptotically become straight as $K \to \infty$.

4 Numerical Tests

4.1 2D Gaussian mixture model

In this experiment, we examine the transport problem between a source distribution μ and a target distribution ν in $\mathbb{R}^2$. The source distribution μ is defined as a mixture of two isotropic Gaussian distributions, denoted as μ_1 (red) and μ_2 (blue). The probability density function is given by:

$$\mu(x) = \frac{1}{2}\mu_1(x) + \frac{1}{2}\mu_2(x) = \frac{1}{2}\mathcal{N}(x \,|\, m_1, \sigma_\mu^2 I_2) + \frac{1}{2}\mathcal{N}(x \,|\, m_2, \sigma_\mu^2 I_2),$$

where the centers are located at $m_1 = (-1, 1)^\top$ and $m_2 = (-1, -1)^\top$, with a standard deviation of $\sigma_\mu = 0.2$. The target distribution ν (lime green) is constructed as a noisy approximation of a vertical line segment. Specifically, samples from ν are generated by:

$$x = \begin{pmatrix} 1 \\ y \end{pmatrix} + \xi, \quad \text{where} \quad y \sim \mathcal{U}(-1, 1), \quad \xi \sim \mathcal{N}(0, \sigma_\nu^2 I_2),$$

with a noise level of $\sigma_\nu = 0.25$. Figure 1 illustrates the particle trajectories under the independent coupling $\pi_0 = \mu \otimes \nu$, where the source and target samples are matched randomly.

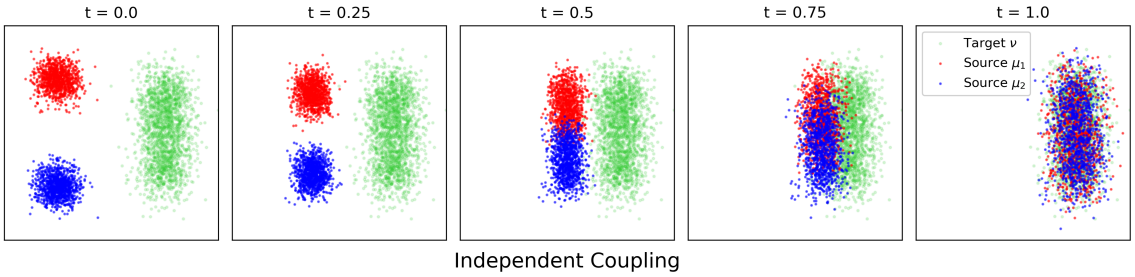

Figure 1: Trajectory visualization under the independent coupling scheme (linear interpolation between independent samples).

Subsequently, we apply the Rectified Flow (Algorithm 1) to iteratively straighten the transport paths. We perform $K = 3$ iterations of the rectification process. The resulting transport trajectories for the k-rectified flow (for $k = 1, 2, 3$) are presented in Figure 2.

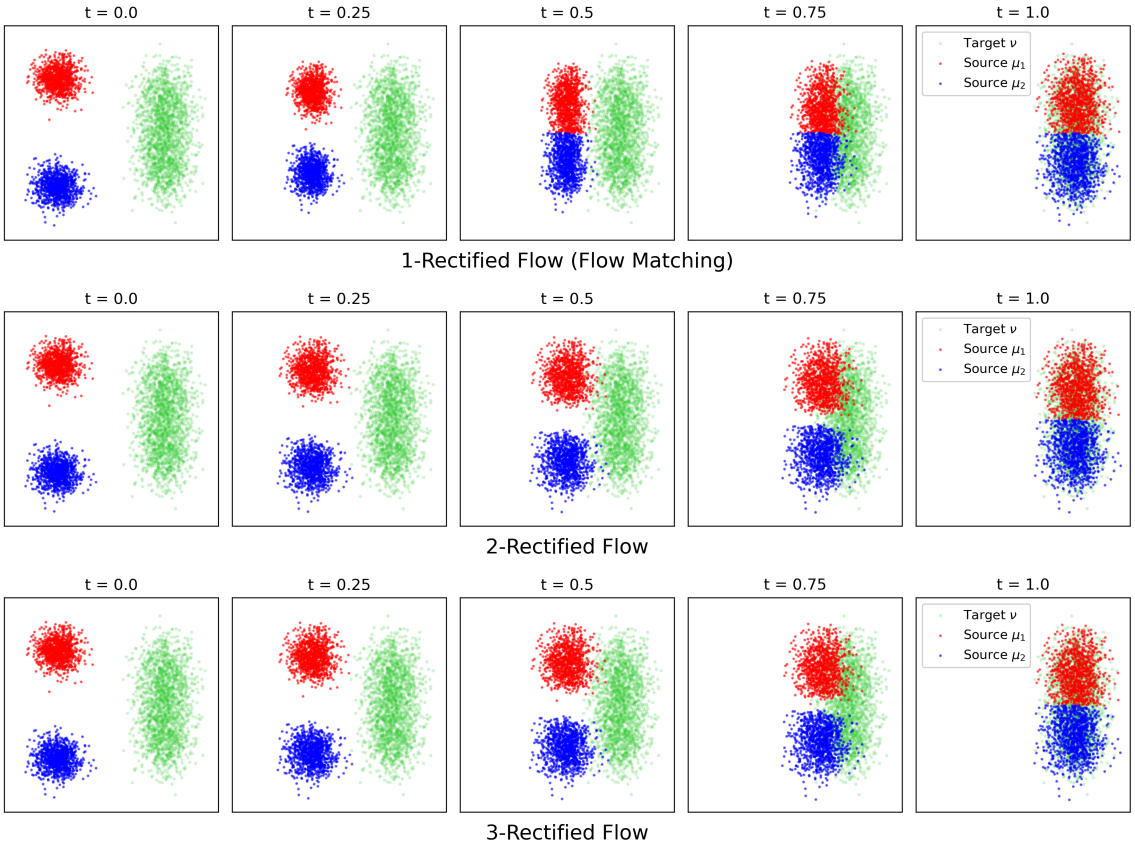

Figure 2: Trajectories of the k-Rectified Flow for iterations $k = 1, 2, 3$. The source modes (red/blue) are transported to the target distribution (green).

As shown in the first row of Figure 2, the 1-Rectified Flow (Flow Matching) successfully disentangles the transport map. Specifically, the algorithm preserves the vertical topology of the data: the upper Gaussian component (μ_1, red) is transported to the upper segment of the target distribution, while the lower component (μ_2, blue) maps to the lower segment.

Notably, the trajectories in the 1-Rectified Flow are already predominantly linear. Consequently, the results for the 2-Rectified and 3-Rectified flows (middle and bottom rows) appear virtually identical to the first iteration. This visual observation is corroborated by the numerical results in Table 1, where the straightness metric S drops by approximately four orders of magnitude after the first step ($4.12 \times 10^{-1} \to 5.08 \times 10^{-5}$), indicating that the flow has rapidly converged to the optimal transport solution.

Iteration (k)	Straightness (S)
1	4.12×10^{-1}
2	5.08×10^{-5}
3	4.81×10^{-5}

Table 1: Straightness metric S for the first three Rectified Flow iterations.

4.2 MNIST Digits Generation

In this section, we evaluate the performance of Rectified Flow on the generative modeling of digits from the MNIST dataset. [numerical setup] To analyze the flow properties, we visualize the generation trajectories for the digits 6 and 8 across two sequential rectification steps.

As illustrated in Figures 3 and 4, the 1-Rectified Flow (Flow Matching) yields high-quality samples with coherent structures. In contrast, the performance of the 2-Rectified Flow deteriorates significantly. This degradation is attributed to the distribution mismatch introduced during the reflow step, where the inferred noise distribution diverges from the prior used during inference. Consequently, for general image generation tasks, the 1-Rectified Flow is typically sufficient to achieve optimal results.

5 Conclusion

In this note, we introduced the Rectified Flow framework, a continuous normalizing flow approach that enforces straight flow trajectories to minimize transport costs between distributions. Theoretically, we demonstrated that the iterative rectification procedure is guaranteed to reduce the transport cost monotonically and that the straightness metric S converges to zero as the number of iterations $K \to \infty$.

Our numerical experiments yielded two distinct insights. On the low-dimensional 2D Gaussian mixture model, the algorithm demonstrated rapid convergence, where the 1-Rectified Flow achieved near-optimal straightness, rendering subsequent reflow steps numerically superior but visually similar. Conversely, on the high-dimensional MNIST dataset, while the 1-Rectified Flow produced high-quality, coherent samples, the 2-Rectified Flow exhibited significant performance degradation. This deterioration is attributed to the distribution mismatch between the standard Gaussian prior and the inferred noise distribution during the reflow step. Therefore, we conclude that for general

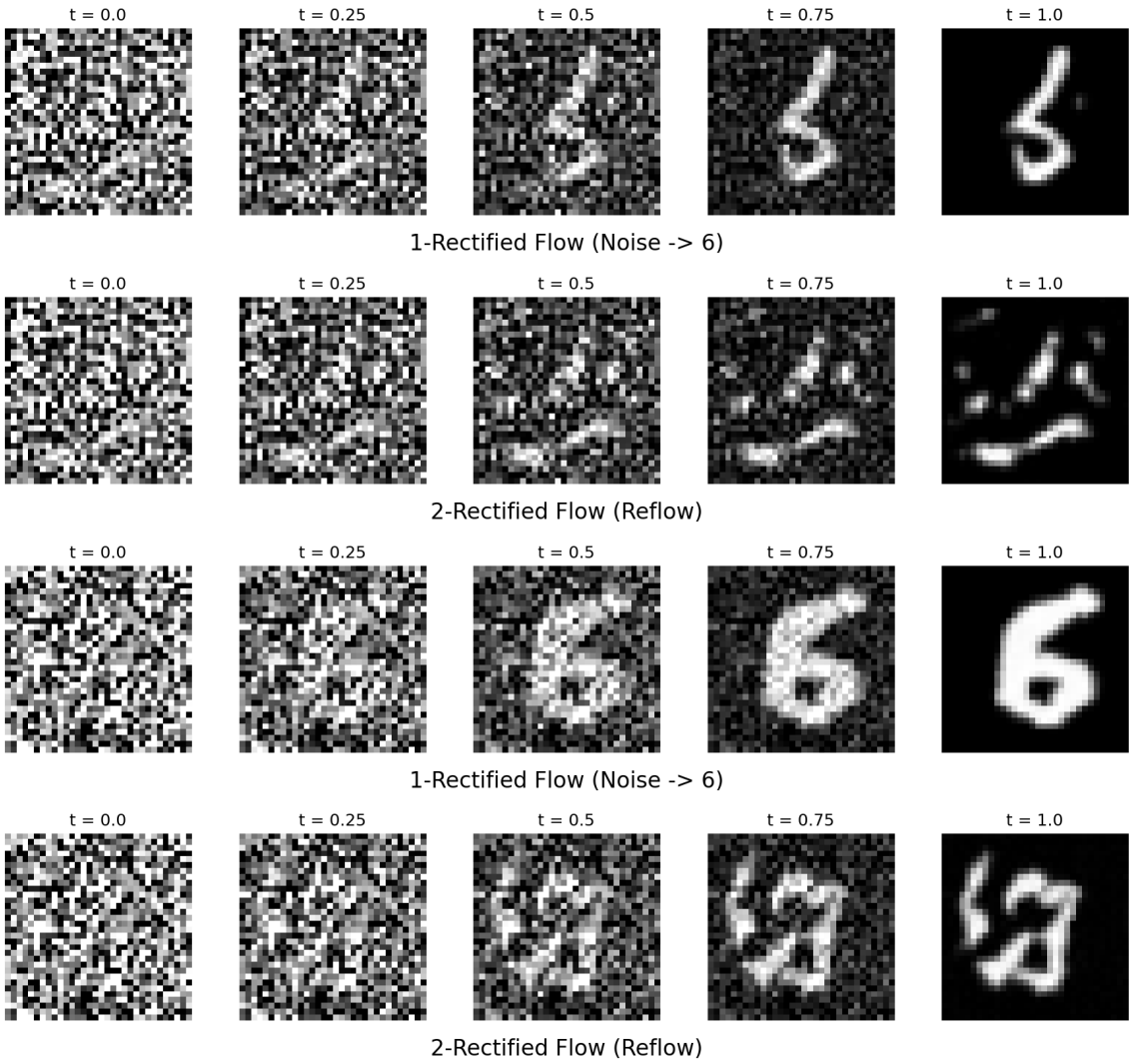

Figure 3: Trajectories of digit 6 across two Rectified Flow steps.

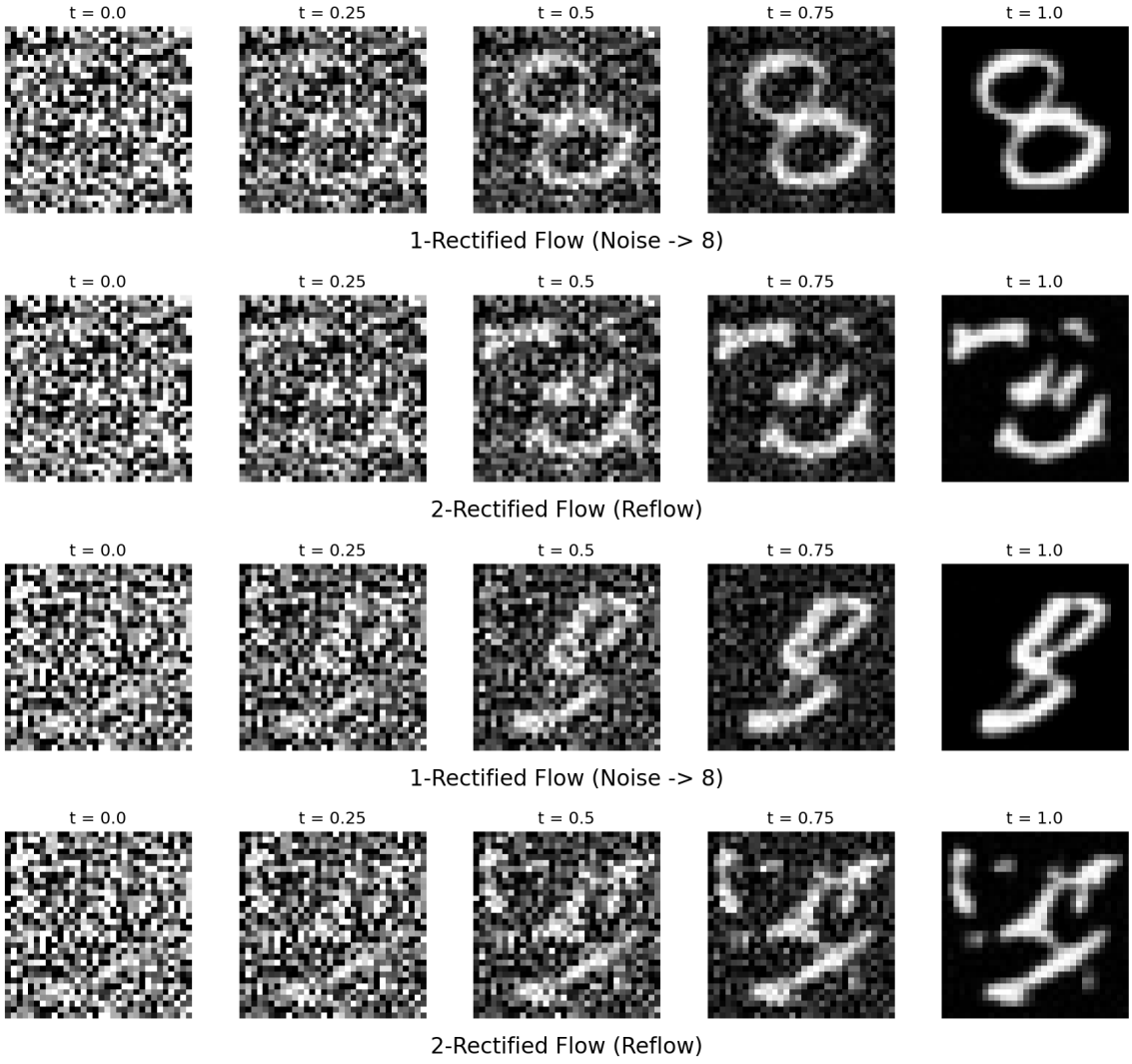

Figure 4: Trajectories of digit 8 across two Rectified Flow steps.

image generation tasks, the 1-Rectified Flow is typically sufficient to achieve optimal generative performance.

The source codes can be found at https://xuda-ye.wordpress.com/wp-content/uploads/2025/11/codes.zip.

References

[1] Yaron Lipman, Ricky TQ Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. *arXiv preprint arXiv:2210.02747*, 2022.

[2] Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. *arXiv preprint arXiv:2209.03003*, 2022.

[3] Cédric Villani et al. *Optimal transport: old and new*, volume 338. Springer, 2008.