Flow/Bridge Matching

Transport any base p_0 to any target p_1

Dario Shariatian, Giovanni Conforti

March 17, 2025

Introduction

Introduction - Flow/Bridge Matching

Flow/bridge matching generative models consist in

- Designing paths $\psi_{x_0,x_1}(t)$ from $x_0 \sim p_0$ conditioned on an endpoint $x_1|x_0 \sim p_{1|0}$. Usually, we find an equivalent ODE flow with a corresponding vector field denoted by $v_1(\cdot|x_1)$.

- Using a bridge mixture theorem to obtain $\psi_{x_0}(t)$ transporting p_0 to p_1. This is usually obtained from a vector field v_t computed as a *bridge mixture* of the conditional vector fields $\mathbb{E}[v_1(\cdot|x_1)]$. Note that the initial coupling might be lost in this process.

- Matching $\psi_{x_0}(t)$ (or v_t) with a neural network $\psi^{\theta}_{x_0}(t)$ (or $v_\theta(\cdot, t)$), using a *bridge matching loss* which, alike denoising score matching in diffusion, consists in some regression problem.

Introduction - Roadmap

- Flow Matching (learn the velocity field of an ODE) [Lip+23]
- Rectified Flows and ReFlow (straighten the velocity field learned by flow matching) [LGL22]
- Bridge Matching (learn the drift of an SDE) [Pel23]

Flow Matching

Flow Matching – Overview and Notations

- **Design a deterministic forward mapping** between a base distribution p_0 (e.g., Gaussian) to the target distribution p_1 (e.g. images).
- **Match the forward mapping** with the *bridge matching loss*. No forward/backward structure.
- **Not necessarily noise to data**. Can work with any coupling $\pi \in \mathcal{M}(p_0, p_1)$. Examples: straighter paths with mini-batch OT [Ton+24], or super-resolution:

$$\pi(x_0, x_1) = p_{0|1}(x_0|x_1)p_1(x_1)$$

  from blurred image to clean image (interesting use-case: CMB de-lensing?).
- **Equivalence with diffusion models** with independent noise to data coupling [Gao+24].

Flow Matching – Forward Process

- **Forward process:**

$$X_t = a_t X_0 + b_t X_1, \quad X_0 \sim p_0, \; X_1 \sim p_1.$$

- The sequence of marginal distributions of $(X_t)_{0 \leqslant t \leqslant 1}$ is denoted by $(p_t)_{0 \leqslant t \leqslant 1}$, and is called the probability path. When

$$a_0 = b_1 = 1, \quad a_1 = b_0 = 0 \,, \tag{1}$$

$(p_t)_{0 \leqslant t \leqslant 1}$ **interpolates** p_0 and p_1, and we consider these conditions satisfied thereafter.
- **Conditionally** on X_0 and X_1, the path is deterministic and differentiable in t.

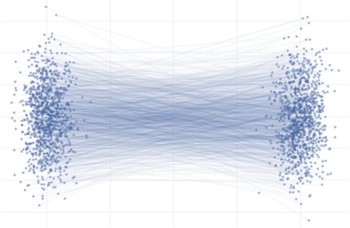

Figure: Conditional paths for two standard Gaussian as base and target distributions [FMD24]

Flow Matching – Forward Process

Depending on the choice of a_t, b_t, one can obtain

- Diffusion models trajectories
- Straight trajectories, by setting $a_t = 1 - t, b_t = t$. This is either referred to as OT (Optimal Transport) or rectified flow [LGL22].

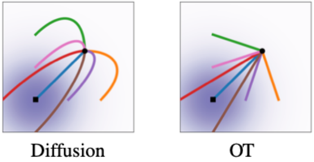

Figure: Diffusion and OT trajectories [Lip+23]

Flow Matching – From Forward Process to ODE (I)

- **Define the forward mapping** (conditioned on x_1):

$$\psi_t(x_0 \mid x_1) \;=\; a_t\, x_0 \;+\; b_t\, x_1 .$$

- **Find the corresponding vector field v_t**

$$v_t\left(\psi_t(x_0 \mid x_1)|x_1\right) = \frac{\mathrm{d}}{\mathrm{d}t}\,\psi_t(x_0 \mid x_1) \;=\; \psi_t'(x_0 \mid x_1)\ .$$

  This gives us a *conditional vector field* $v_t(\cdot \mid x_1)$.

- **Conditional Forward ODE**

$$\mathrm{d}X_t \;=\; v_t\big(X_t \mid X_1\big)\,\mathrm{d}t, \quad \text{with } X_0 \sim p_0,\; X_1 \sim p_1 .$$

- **Problem: conditioning on** x_1.
  - For each sample, x_1 is drawn, and we get a different ODE.
  - We want a single *unconditional* vector field to describe the same marginal path $(p_t)_{0\leqslant t\leqslant 1}$.

Flow Matching – From Forward Process to ODE (II)

- **Goal:** Construct v_t (no dependence on x_1) such that integrating

$$\mathrm{d}Z_t = v_t(Z_t)\,\mathrm{d}t \quad \text{with } Z_0 \sim p_0$$

produces the same marginal distributions p_t as X_t. By abuse of notations I will write Z_t as X_t.

- **Unconditional vector field ([Lip+23]):**

$$v_t(x_t) = \mathbb{E}\left[v_t(x_t \mid X_1)\,\frac{p_t(x_t \mid X_1)}{p_t(x_t)}\right]. \tag{2}$$

- This is the *marginal* (or unconditional) vector field.

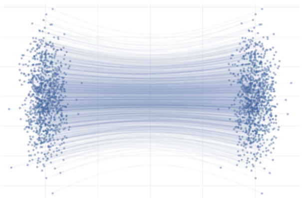

Figure: Marginal paths, as given by the marginal vector field, for two standard Gaussian as base and target distributions [FMD24]

Flow Matching – Unconditional Vector Field – Proof

Proof of 2.

Flow Matching – Backward Process

- **Two views of the forward process**
  - *Linear mixture view:* $X_t = a_t\, X_0 + b_t\, X_1$.
  - *ODE view:* X_t solves $\mathrm{d}X_t = v_t(X_t)\,\mathrm{d}t,\ X_0 \sim p_0$.

  Both describe the same $(p_t)_{0 \leqslant t \leqslant 1}$.

- **We cannot directly learn** v_t as it depends on the unknown p_t.

Flow Matching – Generative Process and Training Objective

- **Generative process**

$$\mathrm{d}X_t^\theta = v_\theta(X_t^\theta, t)\,\mathrm{d}t, \quad X_1^\theta \sim p_0. \tag{3}$$

- **Conditional Flow Matching Loss** We fit $v_\theta \approx v_t$ by minimizing a squared ℓ_2 loss:

$$\mathcal{L}_{\mathsf{CFM}} : \theta \mapsto \mathbb{E}\Big[\Big\| v_\theta(X_t, t) \;-\; v_t(X_t | X_1) \Big\|^2\Big] \,. \tag{4}$$

- **Simulation-free training:**
  - The forward process X_t conditioned on X_0 is explicitly characterizable (linear in a noise term X_1).
  - No need to simulate any ODE during training.

Flow Matching – CFM Loss – Proof

Proof of 11.

Rectified Flows and ReFlow

Rectified Flows and ReFlow

Setup for Rectified Flows Same procedure as Flow Matching with *OT transport plan*:

$$\mathrm{d}X_t = (X_1 - X_0)\mathrm{d}t, \qquad (X_0, X_1) \sim \pi_{0,1} \ . \tag{5}$$

References [LGL22], [Kim+24]

Reflow procedure

Iterative procedure to straighten paths Denote by $\pi^0_{0,1}$ the initial data/noise coupling.

- **(1-Reflow)** Train a neural network model θ_1 based on

$$(X_0, X_1) \sim \pi^0_{0,1} \; .$$

  This defines a new coupling $(X_0, X_1) \sim \pi^{\theta_1}_{0,1}$ defined with

$$X_0 \sim p_0 \; , \qquad \mathrm{d}X_t = v_{\theta_1}(X_t, t)\mathrm{d}t \; .$$

- **(2-Reflow)** Train a neural network model θ_2 based on

$$(X_0, X_1) \sim \pi^{\theta_1}_{0,1} \; .$$

  This defines a new coupling $(X_0, X_1) \sim \pi^{\theta_2}_{0,1}$ defined with

$$X_0 \sim p_0 \; , \qquad \mathrm{d}X_t = v_{\theta_2}(X_t, t)\mathrm{d}t \; .$$

- **(k-Reflow)** etc.

Bridge Matching

General Framework

- Denote by $\mathcal{P}(C([0,1],\mathbb{R}^d))$ the set of path measures. The subset of $\mathcal{P}$ consisting of SDEs of the form $\mathrm{d}X_t = v_t(X_t)\mathrm{d}t + \sigma_t \mathrm{d}B_t$ with σ, v locally Lipschitz is denoted $\mathcal{M}$

- For $P_0, P_1 \in \mathcal{P}(\mathbb{R}^d)$, we denote by $\mathcal{M}(P_0, P_1)$ their set of couplings, i.e., all $\pi_{0,1} \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ such that, for all measruable set $A \subset \mathbb{R}^d$, $\int \pi(A, \mathrm{d}y) = P_0(A)$ and $\int \pi(\mathrm{d}x, A) = P_1(A)$.

- We choose a *reference process* $R \in \mathcal{M}$. We denote $R_{|0,1}$ its diffusion bridge, i.e., the distribution of R conditioned on both endpoints. Similarly, we define R_t, $R_{s,t}$, $R_{s|t}$ etc.

- We denote the mixture of bridges as $\pi = R_{|0,1}\pi_{0,1}$, i.e., $\pi(\cdot) = \int R_{|0,1}(\cdot, x_0, x_1)\pi_{0,1}(\mathrm{d}x_0, \mathrm{d}x_1)$.

Bridge Matching

Based on the work of [Pel23], [Bor+23]. Also see Generator matching [Hol+25].
Choose a coupling $\pi \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$.

- **Reference process** we choose R as the distribution of

$$\mathrm{d}X_t = \sigma \mathrm{d}B_t \; . \tag{6}$$

- **Brownian bridge** $\pi = R_{|0,1}\pi_{0,1}$ is the distribution of the process

$$\mathrm{d}X_t = \frac{X_1 - X_t}{1-t}\mathrm{d}t + \sigma \mathrm{d}B_t \;, \quad (X_0, X_1) \sim \pi_{0,1} \; . \tag{7}$$

  We call $(p_{t|1}(\cdot|x_1))_{0 \leqslant t \leqslant 1}$ the conditional probability path. We call $(p_t)_{0 \leqslant t \leqslant 1}$ the marginal probability path, where

$$p_t(x) = \int p_{t|1}(x|x_1)p_1(x_1)\mathrm{d}x_1 \; . \tag{8}$$

- **Sampling** $X_t|X_0, X_1$ It is easy to sample from the Brownian bridge, since:

$$X_t \stackrel{d}{=} (1-t)X_0 + tX_1 + \sigma\sqrt{t(1-t)}Z \;, \qquad Z \sim \mathcal{N}(0, \mathrm{I}_d) \;, \quad (X_0, X_1) \sim \pi_{0,1} \; . \tag{9}$$

Bridge Matching – Proof for Bridge Distribution

Proof of (7) and of (9).

Bridge Matching – Markovian Projection and Loss Objective

- **Markovian projection** The Brownian bridge in (7) admits a Markovian projection which admits the same marginal probability path $(p_t)_{0 \leqslant t \leqslant 1}$:

$$\mathrm{d}X_t = \mathbb{E}\left[\frac{X_1 - X_t}{1-t} | X_t\right] \mathrm{d}t + \sigma \mathrm{d}B_t \, , \quad X_0 \sim p_0 \, . \tag{10}$$

This can be proved with Fokker-Planck (continuity equation with noise).

- **Loss function** With v_θ a neural network, the following two objective functions are equivalent:

$$\mathcal{L}_{\mathsf{FM}}(\theta) = \mathbb{E}\left[v_\theta(t, X_t) - v_t(X_t)\right] \tag{11}$$

$$\mathcal{L}_{\mathsf{CFM}}(\theta) = \mathbb{E}\left[v_\theta(t, X_t) - v_{t|1}(X_t|X_1)\right] \, , \tag{12}$$

where

$$v_{t|1}(x_t|x_1) = \frac{x_1 - x_t}{1-t} \, , \quad v_t(x_t) = \mathbb{E}\left[v_{t|1}(X_t|X_1)|X_t = x_t\right] \, . \tag{13}$$

Thus in practice, we regress against the conditional vector field.

Bridge Matching – Proof

Proof of (10) and of (11).

Go further

- Doob's h-transforms for non-generic Bridges
  `https://linbaba.wordpress.com/2010/06/02/doob-h-transforms/`
- ...

Thanks for listening

Reference I

[LGL22] Xingchao Liu, Chengyue Gong, and Qiang Liu. *Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow*. 2022. arXiv: 2209.03003 [cs.LG]. URL: https://arxiv.org/abs/2209.03003.

[Bor+23] Valentin De Bortoli et al. *Augmented Bridge Matching*. 2023. arXiv: 2311.06978 [cs.LG]. URL: https://arxiv.org/abs/2311.06978.

[Lip+23] Yaron Lipman et al. *Flow Matching for Generative Modeling*. 2023. arXiv: 2210.02747 [cs.LG]. URL: https://arxiv.org/abs/2210.02747.

[Pel23] Stefano Peluchetti. *Non-Denoising Forward-Time Diffusions*. 2023. arXiv: 2312.14589 [cs.LG]. URL: https://arxiv.org/abs/2312.14589.

[FMD24] Tor Fjelde, Emile Mathieu, and Vincent Dutordoir. *An Introduction to Flow Matching*. Jan. 2024. URL: https://mlg.eng.cam.ac.uk/blog/2024/01/20/flow-matching.html.

[Gao+24] Ruiqi Gao et al. "Diffusion Meets Flow Matching: Two Sides of the Same Coin". In: 2024. URL: https://diffusionflow.github.io/.

[Kim+24] Beomsu Kim et al. *Simple ReFlow: Improved Techniques for Fast Flow Models*. 2024. arXiv: 2410.07815 [cs.LG]. URL: https://arxiv.org/abs/2410.07815.

Reference II

[Ton+24] Alexander Tong et al. *Improving and generalizing flow-based generative models with minibatch optimal transport*. 2024. arXiv: 2302.00482 [cs.LG]. URL: https://arxiv.org/abs/2302.00482.

[Hol+25] Peter Holderrieth et al. *Generator Matching: Generative modeling with arbitrary Markov processes*. 2025. arXiv: 2410.20587 [cs.LG]. URL: https://arxiv.org/abs/2410.20587.