# GRADUAL FINE-TUNING FOR FLOW MATCHING MODELS

**Gudrun Thorkelsdottir & Arindam Banerjee**
University of Illiois Urbana-Champaign
{gudrunt2, arindamb}@illinois.edu

## ABSTRACT

Fine-tuning flow matching models is a central challenge in settings with limited data, evolving distributions, or strict efficiency demands, where unconstrained fine-tuning can erode the accuracy and efficiency gains learned during pretraining. Prior work has produced theoretical guarantees and empirical advances for reward-based fine-tuning formulations, but these methods often impose restrictions on permissible drift structure or training techniques. In this work, we propose Gradual Fine-Tuning (GFT), a principled framework for fine-tuning flow-based generative models when samples from the target distribution are available. For stochastic flows, GFT defines a temperature-controlled sequence of intermediate objectives that smoothly interpolate between the pretrained and target drifts, approaching the true target as the temperature approaches zero. We prove optimization results for both marginal and conditional GFT objectives, enabling the use of suitable (e.g., optimal transport) couplings during GFT while preserving correctness. Empirically, GFT improves convergence stability and shortens probability paths, resulting in faster inference, while maintaining generation quality comparable to standard fine-tuning. Our results position GFT as a theoretically grounded and practically effective alternative for scalable adaptation of flow matching models under distribution shift.

## 1 INTRODUCTION

Recent advances in flow-based generative modeling (Lipman et al., 2023; Liu et al., 2023) have achieved remarkably high generation quality, and have seen widespread use in natural image generation (Esser et al., 2024), as well as scientific applications such as molecule generation (Jing et al., 2024; Klein et al., 2023; Song et al., 2023), biological sequence modeling (Stark et al., 2024; Davis et al., 2024), and atmospheric forecasting (Fotiadis et al., 2025; Bao et al., 2024). In settings characterized by limited computational resources, data-scarcity, or evolving data distributions, fine-tuning a pretrained flow matching model is often the most effective approach for achieving accurate generation. Unlike training from random initialization, fine-tuning must reconcile the competing objectives of adapting to a new target distribution while preserving the structural properties (e.g. short probability paths) learned during pretraining. Such considerations are important for flow matching models, where path lengths directly influence generation efficiency, and small distributional changes can substantially alter the path lengths. As a result, effective fine-tuning requires principled mechanisms that enable controlled adaptation without sacrificing stability, accuracy, or inference efficiency.

Prior work on flow matching fine-tuning is dominated by reward-based methods. These rely on an external function that is constructed or trained to reward desirable behavior, such as human preference (Black et al., 2024; Clark et al., 2024) or physical constraints (Tauberschmidt et al., 2025). Notably, Adjoint Matching (Domingo-Enrich et al., 2025) casts fine-tuning as a Stochastic Optimal Control problem, and optimizes a control vector field which linearly adapts the pretrained model. Other methods (Fan et al., 2025b;a) extend fine-tuning to non-differentiable rewards, and demonstrate that constraining the deviation from the pretrained model can preserve generation diversity.

Despite this progress, several key limitations remain. First, fine-tuning with access to target distribution samples, eliminating the reliance on an external reward function, is comparatively underexplored. Second, reward-based fine-tuning methods require a memoryless noise schedule to provably

converge to the correct target distribution (Domingo-Enrich et al., 2025), meaning that training cannot leverage optimal transport (OT) couplings between source and target samples. This restriction has significant practical implications, as OT-based training can substantially improve generation efficiency (Tong et al., 2024; Pooladian et al., 2023). Further, reward-based optimal control methods for fine-tuning optimize an auxiliary vector field that is added linearly to the pretrained drift (Domingo-Enrich et al., 2025). These approaches are therefore incompatible with many popular nonlinear fine-tuning techniques, including full fine-tuning and low-rank adaptation (LoRA) (Hu et al., 2022). Finally, standard fine-tuning, defined as continued training after changing the target data distribution, uses the pretrained model as a warm-start, but otherwise ignores its learned dynamics.

We propose a principled framework for fine-tuning flow-based generative models, called Gradual Fine-Tuning (GFT). GFT defines a temperature-controlled sequence of intermediate objectives that smoothly interpolate between the pretrained and target dynamics, approaching the true target as the temperature approaches zero. We generalize GFT to Conditional Flow Matching (CFM), establishing equivalent results for GFT when conditioning on arbitrary source-target couplings. Crucially, this allows for the use of training techniques that enhance inference efficiency. GFT is also explicitly developed to accommodate nonlinear fine-tuning, making it compatible with widely used parameter fine-tuning techniques such as full fine-tuning and LoRA. We empirically demonstrate the strong performance of GFT on several natural distribution shift datasets. Across all settings, GFT achieves more stable convergence and consistently shorter probability paths without sacrificing generation quality relative to standard fine-tuning. This work presents the first systematic investigation of annealed gradual fine-tuning for flow matching models using a temperature parameter, and establishes a foundation for scalable and efficient adaptation of flow-based generative models.

## 2 Background and Preliminaries

**Flow Matching** Flow Matching (Lipman et al., 2023; Liu et al., 2023) is a training algorithm for dynamical generative models, a class of models which iteratively simulate a continuous process between a source distribution $p_0$ and a data-defined target distribution $p_1$. Flow Matching defines a generative model by learning a time-varying vector field $v_\theta(X_t, t) : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$ which transports a family of probability densities $\{p_t\}_{t\in[0,1]}$ along a deterministic path: $dX_t = v_\theta(X_t, t)dt$. The trained model $v_\theta$ simulates an ordinary differential equation (ODE) which transports mass from $p_0$ to $p_1$, allowing for simulated sampling from the target distribution. To achieve a well-defined continuous transport process, each intermediate density $p_t$ must satisfy the continuity equation with the vector field $v_\theta(\cdot, t)$ (Appendix B). Despite achieving impressive generation quality in many applications, the iterative nature of flow-based generative models significantly reduces their inference efficiency compared to earlier single-step generative models (Kingma & Welling, 2014; Goodfellow et al., 2014). However, several variations of the training algorithm, such as minimizing vector field curvature (Liu et al., 2023) or sampling according to OT plans (Tong et al., 2024), have proven to lead to highly efficient inference-time generation.

**Stochastic View of Flow Matching** Despite their deterministic formulation, flow matching models can be embedded into a class of stochastic differential equations (SDEs) with time marginals that match $p_t$. Substituting the continuity equation into the Fokker-Planck equation (which describes SDE density evolution) and solving for the SDE drift term recovers the exact ODE drift, added to a product of the diffusion coefficient and the score function of the density $p_t$ (details shown in Appendix C). This has motivated the exploration of the relationship between flow matching and diffusion models (Song et al., 2021), as well as the related stochastic interpolant model (Albergo et al., 2025). We recover the flow matching ODE by setting the diffusion term to 0.

## 3 Gradual Fine-Tuning

### 3.1 Problem Setup

We assume that we have access to a pretrained flow matching model with parameters $\theta_0$ which generates samples from a distribution $p_1$. The model has been trained to simulate the base continuous process

$$\mathbb{P}_{\theta_0} : dX_t = v_{\theta_0}(X_t, t)dt + \sigma_t dB_t, \qquad X_0 \sim p_0 \tag{1}$$

where $v_{\theta_0}$ is the vector field transforming samples from the source distribution $p_0$ to $p_1$, and $B_t$ is a standard $d$-dimensional Brownian motion. We now want to fine-tune this model to generate samples from a new target distribution $q$, from which we have samples $\mathcal{D}$. The target fine-tuned continuous process is defined as

$$\mathbb{P}_q : dX_t = v_q(X_t, t)dt + \sigma_t dB_t, \qquad X_0 \sim p_0. \tag{2}$$

where $v_q$, the vector field generating samples from $q$, will be approximated by the neural network $v_\theta$ to induce the fine-tuned process $\mathbb{P}_\theta$. Here, $\mathbb{P}_{\theta_0}, \mathbb{P}_q$, and $\mathbb{P}_\theta$ are probability path measures on the path space $C([0,1], \mathbb{R}^d)$ induced by their corresponding SDEs. For each process, the distribution of the terminal state $X_1$ generated by the SDE is defined by its corresponding terminal marginal. In particular, under $\mathbb{P}_{\theta_0}$ we have $X_1 \sim p_1$, and under $\mathbb{P}_q$ we have $X_1 \sim q$. The goal of fine-tuning is to adapt the base process such that the terminal marginal of $\mathbb{P}_\theta$ matches the target distribution $q$. We now present GFT, a fine-tuning method which preserved optimality under arbitrary source-target couplings, and is inherent compatibility with nonlinear fine-tuning methods.

### 3.2 Gradual Fine-Tuning Objective

Consider the following optimization over SDE-induced path measures:

$$\begin{aligned} \min_\theta \quad & \mathrm{KL}(\mathbb{P}_\theta \| \mathbb{P}_q) + \beta \mathrm{KL}(\mathbb{P}_\theta \| \mathbb{P}_{\theta_0}) \\ \text{s.t.} \quad & dX_t = v_\theta(X_t, t)dt + \sigma_t dB_t \ . \end{aligned} \tag{3}$$

The first term of the cost expression promotes alignment with the target process $\mathbb{P}_q$, while the second term acts as a divergence regularizer that penalizes large deviations from the pretrained dynamics. The coefficient $\beta \in \mathbb{R}_+$ is a temperature parameter that controls the strength of this regularization. The GFT objective can therefore be viewed as an interpolation between two transport processes, one which guides samples to the terminal distribution by matching the target process $\mathbb{P}_q$, and another which maintains the pretrained dynamics by penalizing deviation from $\mathbb{P}_{\theta_0}$. Balancing these costs suggests that the optimal fine-tuned model should be a weighted average of the base and target dynamics, controlled by the temperature $\beta$. We formalize this intuition in the following theorem.

**Theorem 1.** *Let $\mathbb{P}_{\theta_0}, \mathbb{P}_q$, and $\mathbb{P}_\theta$ be path measures induced by SDEs with drift terms $v_{\theta_0}, v_q$, and $v_\theta$, respectively. Assume that these processes share the diffusion coefficient $\sigma_t$. Then, for a given temperature $\beta$, the vector field minimizing the GFT objective (3) is given by the convex combination*

$$v^*_\theta(X_t, t) = \left(\frac{1}{1+\beta}\right) v_q(X_t, t) + \left(\frac{\beta}{1+\beta}\right) v_{\theta_0}(X_t, t) \ .$$

*Proof given in Appendix F.*

Theorem 1 states that minimizing the GFT objective (3) results in a closed-form and continuous optimal drift which is a weighted arithmetic mean of the pretrained and target vector fields. The weighting is directly controlled by the temperature $\beta$, resulting in limiting behavior which recovers the pretrained and target dynamics

$$\lim_{\beta \to \infty} v^*_\theta = v_{\theta_0} \ , \qquad \lim_{\beta \to 0^+} v^*_\theta = v_q \ . \tag{4}$$

**Annealed Regularization.** The limiting behavior of the optimal vector field $v^*_\theta$ provides a clear motivation for the gradual nature of our method. To control the tradeoff between preserving pretrained knowledge and adapting to the new target distribution, GFT uses a time-dependent temperature $\beta_s$, where $s$ denotes the current fine-tuning step. At the start of fine-tuning, a high temperature strongly penalizes deviation from the base model, keeping updates close to the pretrained drift and preventing premature collapse towards the target. As training progresses, the temperature $\beta_s$ is gradually annealed toward $0^+$, smoothly relaxing this constraint and allowing the fine-tuned drift $v_\theta$ to asymptotically recover the target vector field $v_q$ (4). This progression creates a gradual transition between the pretrained and target dynamics, improving training stability and yielding improved inference efficiency. Similar gradual adaptation methods have previously been shown to stabilize and improve classical domain adaptation (He et al., 2024), and related trust region methods have recently been considered for diffusion model fine-tuning (Blessing et al., 2025).

**Tractable Reformulation.** Although the GFT objective equation 3 provides a simple and elegant formulation for gradual fine-tuning of flow-based models, it cannot be directly optimized for two reasons. First, it implicitly depends on the vector field $v_q$ through the path measure $\mathbb{P}_q$. The use of $v_q$ requires knowledge of the score or density of the target distribution $q$, which is impossible under the assumption that $q$ is only known through a set of samples $\mathcal{D}$. Second, the objective is defined over the space of path measures, which are not differentiable with respect to the model parameters $\theta$. Fundamentally, the fine-tuned neural network represents the drift term of the SDE which induces the path measure $\mathbb{P}_\theta$, not the measure itself. For direct optimization, the objective must be rewritten in terms of the corresponding drift terms of each path measure. In the following proposition, we demonstrate that the application of Girsanov's Theorem achieves this result.

**Proposition 1.** *Let $\mathbb{P}_f$ and $\mathbb{P}_g$ be path measures induced by SDEs with drift terms $f_t$ and $g_t$, respectively. Assume that these processes share the diffusion coefficient $\sigma_t$, are mutually absolutely continuous, and draw initial samples from the same source distribution. Following Girsanov's theorem, the KL divergence between these path measures is given by:*

$$KL(\mathbb{P}_f \| \mathbb{P}_g) = \mathbb{E}_{X_{[0,1]} \sim \mathbb{P}_f} \left[ \frac{1}{2} \int_0^1 \|\sigma_t^{-1}(f_t(X_t) - g_t(X_t))\|^2 dt \right] \tag{5}$$

*Proof given in Appendix D.*

We apply Proposition 1 to (3) to write the tractable form of the GFT objective (details given in Appendix H). Note that we set $\sigma_t = I$, as we find that this leads to empirically strong results while maintaining simplicity.

$$\mathcal{L}(\theta) = \mathbb{E}_{\mathbb{P}_\theta} \left[ \frac{1}{2} \int_0^1 (\|v_\theta - v_q\|^2 + \beta \|v_\theta - v_{\theta_0}\|^2) dt \right] \tag{6}$$

While this reformulation successfully shifts optimization from the path space to the space of vector fields, enabling direct optimization of the neural network $v_\theta$, the dependence on $v_q$ still exists. In the next section, we eliminate this dependency by following the CFM framework to reformulate the marginal GFT objective as a conditional objective.

### 3.3 Optimality Under Source-Target Coupling

A key property of GFT is that conditioning on arbitrary couplings between the source and target distributions does not alter the unconditional optimization results. This enables the use of OT couplings during fine-tuning, creating an opportunity to substantially improve generation efficiency.

Let $\pi \in \Pi(p_0, q)$ be any coupling between the source and target distributions with marginals $\int \pi(X_0, X_1) dX_1 = p_0(X_0)$ and $\int \pi(X_0, X_1) dX_0 = q(X_1)$. Suppose that training samples are drawn jointly, and define $Z := (X_0, X_1) \sim \pi$. We can then define the conditional GFT objective,

$$\mathcal{L}_\pi(\theta) = \mathbb{E}_{Z \sim \pi} \left[ KL(\mathbb{P}_\theta(\cdot|Z) \| \mathbb{P}_q(\cdot|Z)) + \beta KL(\mathbb{P}_\theta(\cdot|Z) \| \mathbb{P}_{\theta_0}(\cdot|Z)) \right] \tag{7}$$

where for $\phi = \theta, \theta_0, q$, the path measure $\mathbb{P}_\phi(\cdot|Z)$ is induced by the conditional SDE $dX_t = v_\phi(X_t, t|Z)dt + \sigma_t dB_t$. For a fixed $Z$, minimizing this objective yields a conditional optimal vector field (details given in Appendix G).

$$v_\theta^*(X_t, t|Z) = \left( \frac{1}{1+\beta} \right) v_q(X_t, t|Z) + \left( \frac{\beta}{1+\beta} \right) v_{\theta_0}(X_t, t|Z) \, . \tag{8}$$

Importantly, this expression is not directly usable in training, as it depends on the unknown conditional target drift. CFM resolves this by choosing a simple, tractable conditional vector field $u_q$ which marginalizes to the the true target, such that $\mathbb{E}_{Z \sim \pi}[u_q(X_t, t|Z)] = v_q(X_t, t)$ (Lipman et al., 2023). A common choice for $u_q$ is the linear interpolation path between the given source and target samples, which satisfies this marginalization property over (Liu et al., 2023). Crucially, the fine-tuned model $v_\theta$ is unconditional, and does not take $Z$ as input during inference. Instead, $v_\theta$ is trained to approximate the marginalized optimal conditional vector field:

$$v_\theta(X_t, t) \approx \mathbb{E}_{Z \sim \pi}[v_\theta^*(X_t, t|Z)] \, . \tag{9}$$

Applying Proposition 1 to equation 7 and substituting $u_q$ yields the tractable GFT loss:

$$\mathcal{L}_\pi(\theta) = \mathbb{E}_{Z \sim \pi} \mathbb{E}_{X_t \sim \mathbb{P}(\cdot|Z)} \left[ \frac{1}{2} \int_0^1 (\|v_\theta(X_t, t) - u_q(X_t, t|Z)\|^2 + \beta \|v_\theta(X_t, t) - v_{\theta_0}(X_t, t|Z)\|^2) dt \right] . \tag{10}$$

This loss is now tractable, and can be optimized directly over the parameters of the neural network $v_\theta$. When the unconditional base and target vector fields are defined as the marginals of their conditional counterparts, we can show that the gradients of the conditional objective $\mathcal{L}_\pi$ with respect to the model parameters $\theta$ will be exactly equivalent to the gradients of the unconditional GFT objective $\mathcal{L}$ (proof given in Appendix G). Optimizing this conditional loss therefore preserves the theoretical guarantees of the unconditional objective, while enabling the use of arbitrary couplings, including OT, during fine-tuning without introducing bias.

Note that unlike pretraining, which considers a single pair of source and target distributions, fine-tuning involves the shared source distribution $p_0$, the pretraining target $p_1$, and the fine-tuning target $q$. Our objective is ultimately to generate samples from $q$ by transporting mass from $p_0$, making $(p_0, q)$ the natural pair to couple during fine-tuning. When $q$ is close to $p_1$, OT paths from the coupling $(p_0, p_1)$, as might be used during pretraining, give a good approximation for short paths from $p_0$ to $q$. In this setting, the pretrained model provides a strong initialization for the target dynamics (see results over small distribution shifts, Section 4.1).

## 4 EXPERIMENTS

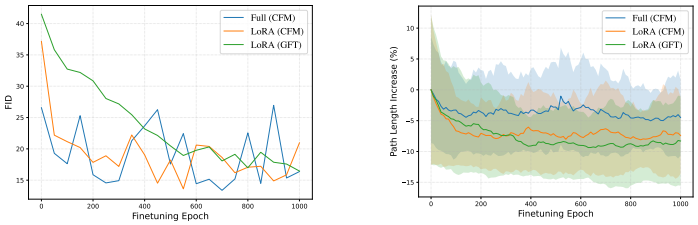
We experimentally evaluate GFT by fine-tuning a pretrained flow matching model, comparing the performance of GFT to standard fine-tuning with the CFM objective (Lipman et al., 2023). Note that we leave comparison against reward-based SOC fine-tuning methods for future work, as the reward function corresponding to GFT is intractable under a memoryless noise schedule (Appendix L). To evaluate GFT under natural distribution shifts, we test it against standard fine-tuning on three WILDS benchmark dataset (Koh et al., 2021): the Camelyon17 dataset of histopathology images (Bándi et al., 2019), the RxRx1 dataset of fluorescent microscopy images (Sypetkowski et al., 2023), and the FMoW 3-channel satellite image dataset (Christie et al., 2018). Each image in these datasets belongs to a single group of samples with common natural factors such as experimental batches, geographical region, or time of acquisition. The training and validation split of each of these datasets are comprised of disjoint sets of groups, creating a distribution shift between them. To ensure rigorous evaluation, we hold out a random 20% of both the training and validation splits of each dataset for exclusive use in calculating Fréchet Inception Distance (FID).

We define two distinct fine-tuning scenarios to evaluate GFT across varying degrees of distribution shift. **Cross-domain adaptation.** We fine-tune a model pretrained on Cifar-10 to the training split of a WILDS dataset, and calculate FID on the held-out training samples. This represents a large shift from a general-purpose prior to a specific scientific domain. **In-domain adaptation.** Starting from models after cross-domain adaptation, we perform a second round of fine-tuning on the corresponding validation split, calculating FID on the held-out validation samples. This simulates a subtler distribution shift where adaptation is required within the same scientific context.

In addition to FID and average path length, we report three metrics to quantify the speed and stability of convergence. Instantaneous variance quantifies local volatility by taking the average of sliding window variance calculations relative to a radial basis function mean. Convergence rate measures adaptation speed by taking the average of the absolute slopes of linear regressions performed over

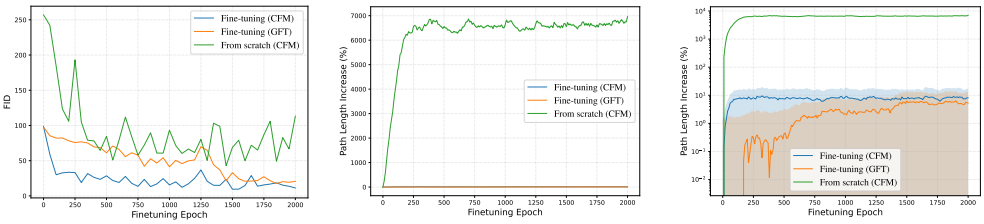
Table 1: Convergence and stability analysis on the Camelyon17 dataset. The first row shows results from cross-domain adaptation (one method is trained from random initialization), and the second shows results from in-domain adaptation. Extended results in Appendix J.4.

| | **Objective** | **Inst. Variance** $\downarrow$ | **Conv. Rate** $\uparrow$ | **Spearman** $\rho \downarrow$ |
|---|---|---|---|---|
| Cross-Domain Adaptation | Full (CFM) | 99.242 | $3.133 \times 10^{-2}$ | -0.663 |
| | Full from R.I. (CFM) | 859.632 | $\mathbf{7.801 \times 10^{-2}}$ | -0.394 |
| | LoRA (CFM) | 70.604 | $2.504 \times 10^{-2}$ | -0.597 |
| | Full (GFT) | **63.132** | $3.647 \times 10^{-2}$ | **-0.932** |
| In-Domain Adaptation | Full (CFM) | 19.387 | $4.693 \times 10^{-3}$ | -0.253 |
| | LoRA (CFM) | 13.134 | $1.182 \times 10^{-2}$ | -0.499 |
| | LoRA (GFT) | **8.429** | $\mathbf{2.213 \times 10^{-2}}$ | **-0.966** |




(a) FID. (b) Path length.

Figure 1: Comparison of fine-tuning methods for in-domain adaptation on Camelyon17. The shaded region of the path length graph (b) represents one standard deviation from the mean. Further results shown in Appendix J.2.




(a) FID. (b) Path length. (c) Path length (log scale).

Figure 2: Cross-domain adaptation results, as opposed to training from random initialization. Fine-tuning and pretraining are done on the Camelyon17 train dataset. The shaded regions of the path length graphs (b, c) represent one standard deviation from the mean.

sliding windows. Finally, we report Spearman correlation between training epochs and FID to assess global consistency. Further details on practical implementation, including datasets, performance metrics, and path length calculations, are given in Appendix I.

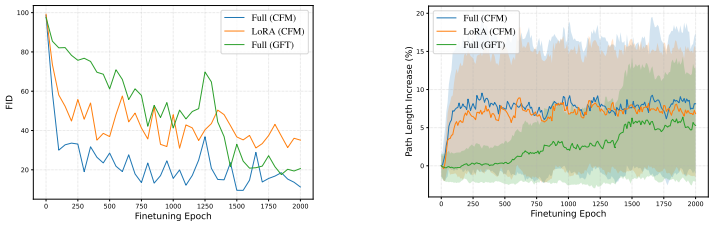
### 4.1 Fine-tuning for In-Domain Adaptation

We begin our empirical evaluation with an investigation of fine-tuning for in-domain adaptation. These experiments perform 1000 epochs of fine-tuning on the validation split of each dataset, starting from a model pretrained on the corresponding train set. This allows for comparison of GFT with standard fine-tuning across small distribution shifts within the same scientific context. For this setting, where the pretraining and fine-tuning target distribution are similar, low rank updates demonstrate superior performance. LoRA with the CFM objective matches the accuracy of full fine-tuning throughout training, and LoRA with the GFT objective reaches a similar accuracy after about 600 epochs (Figure 1a)). However, GFT converges with significantly higher stability than CFM (table 1). GFT achieves the lowest instantaneous variance, indicating low oscillation throughout fine-tuning. GFT's Spearman correlation coefficient is close to -1, pointing to strong monotonic convergence towards optimal model parameters. In addition to high accuracy and stable convergence, GFT achieves a lower average path length than both CFM methods, resulting in efficient inference-time generation.

### 4.2 Cross-Domain Adaptation vs Random Initialization

We now begin our evaluation of fine-tuning performance for cross-domain adaptation by comparing the fine-tuning of a base Cifar-10 model on the Camelyon17 train dataset to training on the same dataset from random initialization (Figure 2). The Camelyon17 dataset, which consists of biological images with highly specialized textures and structures, is a significant distribution shift from Cifar-10 in image size, resolution, and material.

FID analysis highlights the clear advantage of fine-tuning over training from random initialization (Figure 2a). Before fine-tuning has commenced, the base Cifar-10 model achieves a substantially




(a) FID. (b) Path length.

Figure 3: Comparison of fine-tuning methods for cross-domain adaptation on Camelyon17. The shaded region of the path length graph (b) represents one standard deviation from the mean. Further results shown in Appendix J.1.

lower FID than random initialization, despite the large difference in the generative distribution. Fine-tuning with either of the tested objectives maintains this advantage over 2000 epochs (a more thorough comparison of these methods is described in subsequent sections). Despite its higher convergence rate, training from random initialization is highly unstable compared to fine-tuning, as measured by both instantaneous velocity and Spearman correlation coefficient (Table 1).

The variation in path length also strongly points to the benefits of fine-tuning (Figures 2b, 2c). Training from random initialization results in a sharp rise in average path length, which does not decrease over the scope of the experiment. Although fine-tuning from the pretrained model also results in a slight increase in average path length, the relative scale is significantly smaller than training from scratch. This suggests that the pretrained Cifar-10 vector field provides a structural prior that remains partially valid even under large shifts.

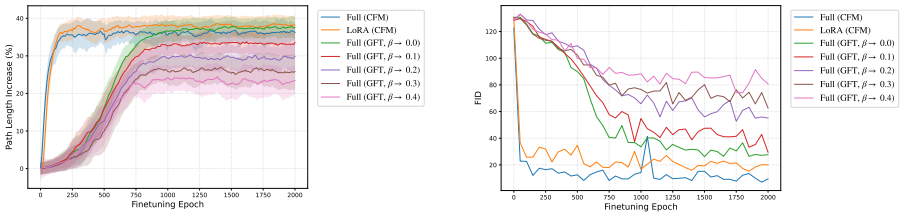
### 4.3 Fine-tuning for Cross-Domain Adaptation

Having established the advantages of fine-tuning a pretrained model over training from random initialization, we now perform a more focused comparison of GFT to standard fine-tuning for cross-domain adaptation. The evolution of FID in these experiments demonstrates the key differences between LoRA and full fine-tuning in the setting of large distribution shifts, as well as the benefits of GFT (Figure 3a). LoRA fine-tuning fails to match the accuracy of full fine-tuning, indicating that low-rank updates are too restrictive to capture the substantial differences between the pretrained and fine-tuned target distributions. Although full fine-tuning with GFT requires more epochs to reach high accuracy generation than CFM, it attains comparable FID after approximately 2000 epochs, and does so with notably improved stability (Table 1). This behavior is consistent with the use of the gradual cooling schedule, which moderates updates early in training leading to more stable optimization. In addition, GFT yields shorter average probability path lengths with lower variance (Figure 3b), indicating better preservation of the pretrained vector field and enabling more efficient generation at inference time.

### 4.4 Effect of Regularization

To demonstrate the effect of regularization to the pretrained model, we conduct a set of experiments with the GFT objective using cooling schedules which vary only in the minimum temperature $\beta$ reached at the end of fine-tuning. Since higher temperatures correspond to a stronger constraint in divergence from the pretrained model (4), the minimum value of the cooling schedule determines the strength of regularization at the final stages of fine-tuning, and therefore dictates the final distribution from which the model generates images.

The outcome of this experiment indicates that variations in regularization result in uniform and highly predictable changes in average path length (Figure 4a). Both full fine-tuning and LoRA with the CFM objective result in an immediate, sharp increase in path length, which then stabilizes after about 250 epochs. Full fine-tuning with the GFT objective however results in a smoother and more gradual increase for all settings. Setting the minimum temperature to zero eventually results in a similar increase in path length as the CFM objective. However, every uniform increase in the minimum $\beta$ value results in a very consistent and predictable decrease in final average path length.




(a) Path length.

(b) FID.

Figure 4: Cross-domain adaptation on RxRx1, showing the regularization effect of varying cooling schedules. Every experiment using the GFT objective has identical hyperparameters, apart from the minimum $\beta$ value reached by the cooling schedule. The shaded region of the path length graph (a) represents one standard deviation from the mean.

Each GFT experiment converges to its final path length at approximately the same epoch, and ends with very similar path length variance.

The corresponding FID evolution in these experiments reveals an inverse relationship between path length and accuracy (Figure 4b). Increasing the minimum temperature reached by the cooling schedule results in a uniformly higher final FID. According to these results, the tradeoff between FID and path length is likely directly caused by the strength of regularization, and can therefore be explicitly controlled by varying the cooling schedule.

## 5 Discussion

We have introduced **GFT**, a principled framework for fine-tuning flow matching generative models. GFT is derived from an optimization over the path space induced by stochastic flows, and yields a closed-form and continuous optimal drift that interpolates between pretrained and target dynamics through a time-dependent annealing cooling schedule. By deriving a fine-tuning objective directly from the continuous formulation of flow-based generative models, we address several critical gaps in the existing literature, most notably the incompatibility with nonlinear parameter updates and optimal transport training methods. By resolving these restrictions, GFT represents a shift towards controllable and reliable adaptation that is effective, yet simple to implement.

The theoretical reults given in Section 3.2 establish that the optimal fine-tuning drift according to GFT is a convex combination of the pretrained and target vector fields, providing justification for the gradual nature of our method. Using the standard CFM objective during fine-tuning treats the pretrained model as a weight initialization, ignoring its learned dynamics in favor of rapid convergence to the target. GFT instead treats the base model as an informative prior that can be gradually updated throughout fine-tuning. In Section 3.3, we demonstrate that the same theoretical results hold under OT-based training, enabling the prioritization of efficient inference while maintaining correctness.

Our empirical results show that, beyond theoretical consistency, this shift towards gradual converges has significant practical benefits. GFT exhibits highly monotonic and robust convergence, as quantified by both Spearman correlation and instantaneous variance. In practical settings, this greatly reduces the need for early stopping and intermediate checkpoint saving, as the model is highly likely to consistently improve throughout the fine-tuning process. We attribute this stability to the cooling schedule, which constrains model updates early in training, then gradually incorporates information from the new target distribution. Another major contribution of this work is the demonstrated reduction in probability path length, and thereby sampling efficiency. The results shown in Figure 4 reveal a uniform and predictable trade-off between generation accuracy and inference speed, which can be easily controlled by varying the cooling schedule.

We have presented Gradual Fine-Tuning, the first framework to both theoretically and empirically explore fine-tuning flow matching models using target samples without the need for memoryless noise schedules or linear drift assumptions. By bridging the gap between high-quality pretraining and fine-tuning, GFT provides a robust solution for adapting generative models to complex, data-driven domains which require nuanced control of the outcome of fine-tuning.

## REFERENCES

Michael S. Albergo, Nicholas M. Boffi, and Eric Vanden-Eijnden. Stochastic interpolants: A unifying framework for flows and diffusions, 2025. URL https://arxiv.org/abs/2303.08797.

Feng Bao, Zezhong Zhang, and Guannan Zhang. An ensemble score filter for tracking high-dimensional nonlinear dynamical systems. *Computer Methods in Applied Mechanics and Engineering*, 432:117447, 2024.

Kevin Black, Michael Janner, Yilun Du, Ilya Kostrikov, and Sergey Levine. Training diffusion models with reinforcement learning. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=YCWjhGrJFD.

Denis Blessing, Julius Berner, Lorenz Richter, Carles Domingo-Enrich, Yuanqi Du, Arash Vahdat, and Gerhard Neumann. Trust region constrained measure transport in path space for stochastic optimal control and inference. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025. URL https://openreview.net/forum?id=6RlbOEcOS4.

Péter Bándi, Oscar Geessink, Quirine Manson, Marcory Van Dijk, Maschenka Balkenhol, Meyke Hermsen, Babak Ehteshami Bejnordi, Byungjae Lee, Kyunghyun Paeng, Aoxiao Zhong, Quanzheng Li, Farhad Ghazvinian Zanjani, Svitlana Zinger, Keisuke Fukuta, Daisuke Komura, Vlado Ovtcharov, Shenghua Cheng, Shaoqun Zeng, Jeppe Thagaard, Anders B. Dahl, Huangjing Lin, Hao Chen, Ludwig Jacobsson, Martin Hedlund, Melih Çetin, Eren Halıcı, Hunter Jackson, Richard Chen, Fabian Both, Jörg Franke, Heidi Küsters-Vandevelde, Willem Vreuls, Peter Bult, Bram van Ginneken, Jeroen van der Laak, and Geert Litjens. From detection of individual metastases to classification of lymph node status at the patient level: The camelyon17 challenge. *IEEE Transactions on Medical Imaging*, 38(2):550–560, 2019. doi: 10.1109/TMI.2018.2867350.

Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David Duvenaud. Neural ordinary differential equations. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, pp. 6572–6583, Red Hook, NY, USA, 2018. Curran Associates Inc.

Gordon Christie, Neil Fendley, James Wilson, and Ryan Mukherjee. Functional map of the world. In *CVPR*, 2018.

Kevin Clark, Paul Vicol, Kevin Swersky, and David J. Fleet. Directly fine-tuning diffusion models on differentiable rewards. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=1vmSEVL19f.

Oscar Davis, Samuel Kessler, Mircea Petrache, İsmail İlkan Ceylan, Michael Bronstein, and Avishek Joey Bose. Fisher flow matching for generative modeling over discrete data. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang (eds.), *Advances in Neural Information Processing Systems*, volume 37, pp. 139054–139084. Curran Associates, Inc., 2024. doi: 10.52202/079017-4413. URL https://proceedings.neurips.cc/paper_files/paper/2024/file/fadec8f2e65f181d777507d1df69b92f\-Paper-Conference.pdf.

Carles Domingo-Enrich, Michal Drozdzal, Brian Karrer, and Ricky T. Q. Chen. Adjoint matching: Fine-tuning flow and diffusion generative models with memoryless stochastic optimal control. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=xQBRrtQM8u.

Patrick Esser, Sumith Kulal, Andreas Blattmann, Rahim Entezari, Jonas Müller, Harry Saini, Yam Levi, Dominik Lorenz, Axel Sauer, Frederic Boesel, Dustin Podell, Tim Dockhorn, Zion English, and Robin Rombach. Scaling rectified flow transformers for high-resolution image synthesis. In Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp (eds.), *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 12606–12633. PMLR, 21–27 Jul 2024. URL https://proceedings.mlr.press/v235/esser24a.html.

Jiajun Fan, Chaoran Cheng, Shuaike Shen, Xiangxin Zhou, and Ge Liu. Fine-tuning flow matching generative models with intermediate feedback, 2025a. URL https://arxiv.org/abs/2510.18072.

Jiajun Fan, Shuaike Shen, Chaoran Cheng, Yuxin Chen, Chumeng Liang, and Ge Liu. Online reward-weighted fine-tuning of flow matching with wasserstein regularization. In *The Thirteenth International Conference on Learning Representations*, 2025b. URL https://openreview.net/forum?id=2IoFFexvuw.

Stathi Fotiadis, Noah D Brenowitz, Tomas Geffner, Yair Cohen, Michael Pritchard, Arash Vahdat, and Morteza Mardani. Adaptive flow matching for resolving small-scale physics. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=YJ1My9ttEN.

Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K.Q. Weinberger (eds.), *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc., 2014. URL https://proceedings.neurips.cc/paper_files/paper/2014/file/f033ed80deb0234979a61f95710dbe25-Paper.pdf.

Aaron J Havens, Benjamin Kurt Miller, Bing Yan, Carles Domingo-Enrich, Anuroop Sriram, Daniel S. Levine, Brandon M Wood, Bin Hu, Brandon Amos, Brian Karrer, Xiang Fu, Guan-Horng Liu, and Ricky T. Q. Chen. Adjoint sampling: Highly scalable diffusion samplers via adjoint matching. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=6Eg1OrHmg2.

Yifei He, Haoxiang Wang, Bo Li, and Han Zhao. Gradual domain adaptation: Theory and algorithms. *Journal of Machine Learning Research*, 25(361):1–40, 2024. URL http://jmlr.org/papers/v25/23-1180.html.

Edward J Hu, yelong shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. LoRA: Low-rank adaptation of large language models. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=nZeVKeeFYf9.

Bowen Jing, Bonnie Berger, and Tommi Jaakkola. Alphafold meets flow matching for generating protein ensembles. In *Forty-first International Conference on Machine Learning*, 2024.

Diederik P. Kingma and Max Welling. Auto-Encoding Variational Bayes. In *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*, 2014.

S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220 (4598):671–680, 1983. doi: 10.1126/science.220.4598.671. URL https://www.science.org/doi/abs/10.1126/science.220.4598.671.

Leon Klein, Andreas Krämer, and Frank Noe. Equivariant flow matching. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine (eds.), *Advances in Neural Information Processing Systems*, volume 36, pp. 59886–59910. Curran Associates, Inc., 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/file/bc827452450356f9f558f4e4568d553b\-Paper-Conference.pdf.

Pang Wei Koh, Shiori Sagawa, Henrik Marklund, Sang Michael Xie, Marvin Zhang, Akshay Balsubramani, Weihua Hu, Michihiro Yasunaga, Richard Lanas Phillips, Irena Gao, Tony Lee, Etienne David, Ian Stavness, Wei Guo, Berton A. Earnshaw, Imran S. Haque, Sara Beery, Jure Leskovec, Anshul Kundaje, Emma Pierson, Sergey Levine, Chelsea Finn, and Percy Liang. WILDS: A benchmark of in-the-wild distribution shifts. In *International Conference on Machine Learning (ICML)*, 2021.

Alex Krizhevsky. Learning multiple layers of features from tiny images. Technical report, 2009.

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.

Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. In *International Conference on Learning Representations (ICLR)*, 2023.

Byoungwoo Park, Juho Lee, and Guan-Horng Liu. Functional adjoint sampler: Scalable sampling on infinite dimensional spaces, 2025. URL https://arxiv.org/abs/2511.06239.

Aram-Alexandre Pooladian, Heli Ben-Hamu, Carles Domingo-Enrich, Brandon Amos, Yaron Lipman, and Ricky T. Q. Chen. Multisample flow matching: straightening flows with minibatch couplings. In *Proceedings of the 40th International Conference on Machine Learning*, ICML'23. JMLR.org, 2023.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=PxTIG12RRHS.

Yuxuan Song, Jingjing Gong, Minkai Xu, Ziyao Cao, Yanyan Lan, Stefano Ermon, Hao Zhou, and Wei-Ying Ma. Equivariant flow matching with hybrid probability transport for 3d molecule generation. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine (eds.), *Advances in Neural Information Processing Systems*, volume 36, pp. 549–568. Curran Associates, Inc., 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/file/01d64478381c33e29ed611f1719f5a37-Paper\-Conference.pdf.

Hannes Stark, Bowen Jing, Chenyu Wang, Gabriele Corso, Bonnie Berger, Regina Barzilay, and Tommi Jaakkola. Dirichlet flow matching with applications to DNA sequence design. In *ICLR 2024 Workshop on Machine Learning for Genomics Explorations*, 2024. URL https://openreview.net/forum?id=ehYe5bz8H3.

Maciej Sypetkowski, Morteza Rezanejad, Saber Saberian, Oren Kraus, John Urbanik, James Taylor, Ben Mabey, Mason Victors, Jason Yosinski, Alborz Rezazadeh Sereshkeh, Imran Haque, and Berton Earnshaw. RxRx1: A Dataset for Evaluating Experimental Batch Correction Methods . In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 4285–4294, Los Alamitos, CA, USA, June 2023. IEEE Computer Society. doi: 10.1109/CVPRW59228.2023.00451. URL https://doi.ieeecomputersociety.org/10.1109/CVPRW59228.2023.00451.

Jan Tauberschmidt, Sophie Fellenz, Sebastian J. Vollmer, and Andrew B. Duncan. Physics-constrained fine-tuning of flow-matching models for generation and inverse problems, 2025. URL https://arxiv.org/abs/2508.09156.

Alexander Tong, Nikolay Malkin, Kilian FATRAS, Lazar Atanackovic, Yanlei Zhang, Guillaume Huguet, Guy Wolf, and Yoshua Bengio. Simulation-free schrödinger bridges via score and flow matching. In *ICML Workshop on New Frontiers in Learning, Control, and Dynamical Systems*, 2023. URL https://openreview.net/forum?id=adkj23mvB0.

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.

Tonghe Zhang, Chao Yu, Sichang Su, and Yu Wang. Reinflow: Fine-tuning flow matching policy with online reinforcement learning. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025. URL https://openreview.net/forum?id=ACagRwCCqu.

# A Related Works

**Fine-Tuning for Flow Matching** Flow Matching provides an efficient framework for training Continuous Normalizing Flows (Chen et al., 2018) by learning vector fields that transport probability mass between distributions (Lipman et al., 2023; Liu et al., 2023). Although substantial effort has focused on improving base training objectives and inference efficiency, comparatively little work has addressed the problem of fine-tuning pretrained flow matching models under distribution shifts. Existing work is dominated by reward-based approaches of fine-tuning, which adapt the pretrained model to a tilted target distribution implicitly defined by a reward function. Adjoint Matching (Domingo-Enrich et al., 2025) formulates fine-tuning as a stochastic optimal control problem and introduces a control vector field that linearly modifies the pretrained dynamics. Related methods extend the reward fine-tuning paradigm to non-differentiable rewards and online settings (Fan et al., 2025b;a), or perform fine-tuning through reinforcement learning methods by injecting noise into the deterministic flow (Zhang et al., 2025). While these methods are well suited to preference objectives, they fundamentally differ from our focus on fine-tuning flow matching models with access to target distribution samples.

**Regularization and Inference Efficiency** Previous reward-based fine-tuning methods have demonstrated the use of divergence regularization to limit deviation from the pretrained model, including KL regularization (Domingo-Enrich et al., 2025), or Wasserstein distance (Fan et al., 2025b). However, existing convergence guarantees require the use of a memoryless noise schedule, preventing derivative methods from conditioning on any jointly drawn source-target samples during training, including OT couplings (Havens et al., 2025; Park et al., 2025). This restriction has critical practical consequences, as OT-based training is known to significantly enhance inference efficiency by encouraging short probability path lengths (Tong et al., 2024; Pooladian et al., 2023). Prior fine-tuning formulations also assume a linear relationship between the pretrained and fine-tuned vector fields (Domingo-Enrich et al., 2025), rendering them incompatible with widely used nonlinear fine-tuning techniques, including full fine-tuning and low-rank adaptation (LoRA) (Hu et al., 2022). In contrast, our approach directly admits nonlinear model updates, and establishes theoretical guarantees that remain valid under arbitrary source–target couplings.

# B Preliminaries

## B.1 Flow Matching

Flow Matching (Lipman et al., 2023; Liu et al., 2023) is a training algorithm for dynamical generative models, a class of models which iteratively simulate a continuous process between a source distribution $p_0$ and a data-defined target distribution $p_1$. Flow Matching defines a generative model by learning a time-varying vector field $v_\theta(X_t, t) : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$ which transports a family of probability densities $\{p_t\}_{t\in[0,1]}$ along a deterministic path:

$$dX_t = v_\theta(X_t, t)dt, \qquad X_0 \sim p_0 \tag{11}$$

The trained model $v_\theta$ simulates an ordinary differential equation (ODE) which transports mass from $p_0$ to $p_1$, allowing for simulated sampling from the target distribution. To achieve a well-defined continuous transport process, each intermediate density $p_t$ must satisfy the continuity equation with the vector field $v_t(\cdot) := v_\theta(\cdot, t)$.

$$\partial_t p_t + \nabla_x \cdot (p_t v_t) = 0 \tag{12}$$

Despite achieving impressive generation quality in many applications, the iterative nature of flow-based generative models significantly reduces their inference efficiency compared to earlier single-step generative models (Kingma & Welling, 2014; Goodfellow et al., 2014). However, several variations of the training algorithm, such as minimizing vector field curvature (Liu et al., 2023) or sampling according to OT plans (Tong et al., 2024), have proven to lead to highly efficient inference-time generation.

## B.2 Stochastic View of Flow Matching

Despite their deterministic formulation, flow matching models can be embedded into a class of stochastic differential equations (SDEs) with time marginals that match $p_t$. Consider the following

SDE

$$dX_t = f_t(X_t)dt + g_t(X_t)dB_t, \qquad X_0 \sim p_0 \tag{13}$$

with drift term $f : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$ and diffusion term $g : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^{d\times d}$. Assuming isotropic diffusion, such that $g_t g_t^\intercal = \sigma_t^2 I$, the Fokker-Planck equation describing the change in the density $p_t$ according to this SDE is

$$\partial_t p_t = -\nabla_x \cdot (f_t p_t) + \frac{1}{2}\nabla_x \cdot \nabla_x \cdot (\sigma_t^2 I p_t). \tag{14}$$

In order to match the ODE density evolution, we substitute the continuity equation (12) and solve for the drift term $f_t$ (details shown in Appendix C).

$$f_t = v_t + \frac{1}{2}\sigma_t^2 \nabla_x \log p_t + Z_t \tag{15}$$

$Z_t$ is a divergence free velocity which can simply be set to 0. This formulation explicitly introduces the score function of the density $p_t$, and has therefore motivated the exploration of the relationship between flow matching and diffusion models (Song et al., 2021), as well as the related stochastic interpolant model (Albergo et al., 2025). We recover the exact deterministic flow matching ODE by setting the diffusion coefficient $\sigma_t$ to 0.

## C Recovering the Score in the Stochastic View of Flow Matching

Substituting the continuity equation (12) into the Fokker-Planck equation (14) yields

$$\nabla_x \cdot (p_t v_t) = \nabla_x \cdot (f_t p_t) - \frac{1}{2}\nabla_x \cdot \nabla_x \cdot (\sigma_t^2 I p_t). \tag{16}$$

Since the diffusion coefficient $\sigma_t^2$ depends only on the timestep $t$, we can move it outside the divergence operators.

$$\nabla_x \cdot (p_t v_t) = \nabla_x \cdot (f_t p_t) - \frac{1}{2}\sigma_t^2(\nabla_x \cdot \nabla_x \cdot p_t) \tag{17}$$

We rearrange terms to isolate the drift term divergence.

$$\nabla_x \cdot (f_t p_t) = \nabla_x \cdot (p_t v_t) + \frac{1}{2}\sigma_t^2(\nabla_x \cdot \nabla_x \cdot p_t) \tag{18}$$

As the divergence operator commutes with scalar multiplication, the terms on the right hand side can be combined.

$$\nabla_x \cdot (f_t p_t) = \nabla_x \cdot \left[p_t v_t + \frac{1}{2}\sigma_t^2 \nabla_x p_t\right] \tag{19}$$

Two vector fields have equivalent divergence when they differ only by a divergence-free component $Z_t$, such that $\nabla_x \cdot (p_t Z_t) = 0$. Therefore, this equivalence can be rewritten without the leading divergence operator.

$$f_t p_t = p_t v_t + \frac{1}{2}\sigma_t^2 \nabla_x p_t + p_t Z_t \tag{20}$$

Finally, we divide by $p_t$, using the identity that $\frac{\nabla_x p_t}{p_t} = \nabla \log p_t$.

$$f_t = v_t + \frac{1}{2}\sigma_t^2 \nabla_x \log p_t + Z_t \tag{21}$$

## D Derivation of Path Length by KL Divergence

We derive a distance measure between the probability distribution of the paths generated by two SDEs using Girsanov's Theorem. Let $\mathbb{P}_f$ and $\mathbb{P}_g$ be the path measures of two SDEs on the time interval $[0, 1]$, with initial samples from the same source distribution.

$$\mathbb{P}_f : \quad dX_t = f_t(X_t, t)dt + \sigma_t dB_t, \qquad X_0 \sim p_0 \tag{22}$$

$$\mathbb{P}_g : \quad dX_t = g_t(X_t, t)dt + \sigma_t dB_t, \qquad X_0 \sim p_0 \tag{23}$$

We assume that $\mathbb{P}_f$ and $\mathbb{P}_g$ are mutually absolutely continuous, and that the diffusion coefficient $\sigma_t$ is non-zero. The KL-divergence between the probability measures of these processes is

$$KL(\mathbb{P}_f \| \mathbb{P}_g) = \mathbb{E}_{\mathbb{P}_f}\left[\log \frac{d\mathbb{P}_f}{d\mathbb{P}_g}\right] = -\mathbb{E}_{\mathbb{P}_f}\left[\log \frac{d\mathbb{P}_g}{d\mathbb{P}_f}\right]. \tag{24}$$

Girsanov's Theorem provides the formulation for the Radon-Nikodym derivative $\frac{d\mathbb{P}_g}{d\mathbb{P}_f}$, which specifies how the probability measure $\mathbb{P}_g$ changes with respect to $\mathbb{P}_f$.

$$\frac{d\mathbb{P}_g}{d\mathbb{P}_f} = \exp\left(\int_0^1 \sigma_t^{-1}(g_t - f_t)^\intercal dB_t - \frac{1}{2}\int_0^1 \|\sigma_t^{-1}(g_t - f_t)\|^2 dt\right) \tag{25}$$

$$\log\left(\frac{d\mathbb{P}_g}{d\mathbb{P}_f}\right) = \int_0^1 \sigma_t^{-1}(g_t - f_t)^\intercal dB_t - \frac{1}{2}\int_0^1 \|\sigma_t^{-1}(g_t - f_t)\|^2 dt \tag{26}$$

We now take the expectation with respect to the target process.

$$KL(\mathbb{P}_f \| \mathbb{P}_g) = -\mathbb{E}_{\mathbb{P}_f}\left[\int_0^1 \sigma_t^{-1}(g_t - f_t)^\intercal dB_t\right] + \mathbb{E}_{\mathbb{P}_f}\left[\frac{1}{2}\int_0^1 \|\sigma_t^{-1}(g_t - f_t)\|^2 dt\right] \tag{27}$$

Note that since the integral in the first term is with respect to a Brownian motion, this term is a martingale under the measure $\mathbb{P}_f$, and has an expected value of 0. Therefore, our final expression for Proposition 1 is

$$KL(\mathbb{P}_f \| \mathbb{P}_g) = \mathbb{E}_{\mathbb{P}_f}\left[\frac{1}{2}\int_0^1 \|\sigma_t^{-1}(f_t - g_t)\|^2 dt\right]. \tag{28}$$

## E Relationship Between Path Length and Fisher Information

Consider the practical pretrained and fine-tuned processes

$$\mathbb{P}_{\theta_0} : \quad dX_t = v_{\theta_0}(X_t, t)dt + \sigma_t dB_t, \qquad X_0 \sim p_0 \tag{29}$$

$$\mathbb{P}_{\theta} : \quad dX_t = v_{\theta}(X_t, t)dt + \sigma_t dB_t, \qquad X_0 \sim p_0 \tag{30}$$

Applying Proposition 1 to these processes gives us a measure of the divergence between the probability distribution of their generated paths. In this section we set $\sigma_t = 1$ for simplicity.

$$KL(\mathbb{P}_\theta \| \mathbb{P}_{\theta_0}) = \mathbb{E}_{\mathbb{P}_\theta}\left[\frac{1}{2}\int_0^1 \|(v_\theta - v_{\theta_0})\|^2 dt\right] \tag{31}$$

With fine-tuning methods such as LoRA and head adaptation, we can assume that $\mathbb{P}_\theta \approx \mathbb{P}_{\theta_0}$ throughout the fine-tuning process, and we can further simplify this expression using first-order Taylor approximation.

$$v_\theta - v_{\theta_0} \approx \mathbf{J}_t(X_t, t) \cdot (\theta - \theta_0) \tag{32}$$

Here, $\mathbf{J}_t(X_t, t) = \nabla_\theta v_\theta(X_t, t)\big|_{\theta_0} \in \mathbb{R}^{D \times P}$ is the Jacobian matrix of the vector field $v$ with respect to parameters $\theta$, evaluated at $\theta_0$. Note that when fine-tuning with LoRA or head adaptation, $\Delta\theta = \theta - \theta_0$, which is the guidance of the fine-tuning process. Substituting this back into the $KL$ expression yields

$$KL(\mathbb{P}_\theta \| \mathbb{P}_{\theta_0}) \approx \mathbb{E}_{\mathbb{P}_\theta}\left[\frac{1}{2}\int_0^1 \|\mathbf{J}_t \Delta\theta\|^2 dt\right] \tag{33}$$

$$= \mathbb{E}_{\mathbb{P}_\theta}\left[\frac{1}{2}\int_0^1 (\Delta\theta^\intercal \mathbf{J}_t^\intercal \mathbf{J}_t \Delta\theta) dt\right]. \tag{34}$$

Note that the guidance $\Delta\theta$ is not dependent on $x_t$ or $t$.

$$= \frac{1}{2}\Delta\theta^\intercal \left(\int_0^1 \mathbb{E}_{\mathbb{P}_\theta}\left[\mathbf{J}_t^\intercal \mathbf{J}_t\right] dt\right) \Delta\theta \tag{35}$$

Notice that the integral term is related to the Fisher Information matrix centered at $\theta_0$, $\mathbf{F}_{\theta_0} \in \mathbb{R}^{P \times P}$.

$$\mathbf{F}_{\theta_0} = \int_0^1 \mathbb{E}_{\mathbb{P}_\theta}[\mathbf{J}_t^\intercal \mathbf{J}_t] dt \tag{36}$$

This is the time-integrated expected Fisher Information matrix over the time range $[0, 1]$ for the path measure $\mathbb{P}_\theta$. For practical implementation, we can rewrite the integral as an expectation over a uniformly sampled time $\tau \sim \mathcal{U}[0, 1]$.

$$\mathbf{F}_{\theta_0} = \mathbb{E}_{\tau \sim \mathcal{U}[0,1], X_\tau \sim \mathbb{P}_\theta}[\mathbf{J}_\tau^\intercal \mathbf{J}_\tau] \tag{37}$$

Therefore, the final simplified approximated divergence between the probability distribution of the paths generated by the base and fine-tuned models is

$$KL(\mathbb{P}_\theta \| \mathbb{P}_{\theta_0}) \approx \frac{1}{2}\Delta\theta^\intercal \mathbf{F}_{\theta_0} \Delta\theta. \tag{38}$$

## F Minimizing the Gradual Fine-Tuning Objective

The GFT objective function is defined as

$$\mathcal{L}(\theta) = KL(\mathbb{P}_\theta \| \mathbb{P}_q) + \beta KL(\mathbb{P}_\theta \| \mathbb{P}_{\theta_0}) \tag{39}$$

Applying Proposition 1 to both terms with $\sigma_t = 1$ yields

$$\mathcal{L}(\theta) = \mathbb{E}_{\mathbb{P}_\theta}\left[\frac{1}{2}\int_0^1 \|v_\theta - v_q\|^2 dt\right] + \beta \cdot \mathbb{E}_{\mathbb{P}_\theta}\left[\frac{1}{2}\int_0^1 \|v_\theta - v_{\theta_0}\|^2 dt\right] \tag{40}$$

$$= \mathbb{E}_{\mathbb{P}_\theta}\left[\frac{1}{2}\int_0^1 \|v_\theta - v_q\|^2 dt + \frac{\beta}{2}\int_0^1 \|v_\theta - v_{\theta_0}\|^2 dt\right] \tag{41}$$

$$= \mathbb{E}_{\mathbb{P}_\theta}\left[\int_0^1 \frac{1}{2}\|v_\theta - v_q\|^2 + \frac{\beta}{2}\|v_\theta - v_{\theta_0}\|^2 dt\right] \tag{42}$$

Minimizing this function over $\mathbb{P}_\theta$ is equivalent to finding the optimal drift $v_\theta^*$. We do this by minimizing the expression inside the integral.

$$C = \frac{1}{2}\|v_\theta - v_q\|^2 + \frac{\beta}{2}\|v_\theta - v_{\theta_0}\|^2 \tag{43}$$

$$\frac{dC}{dv_\theta} = (v_\theta - v_q) + \beta(v_\theta - v_{\theta_0}) = 0 \tag{44}$$

$$(1+\beta)v_\theta - v_q - \beta v_{\theta_0} = 0 \tag{45}$$

$$v_\theta^*(X_t, t) = \left(\frac{1}{1+\beta}\right) v_q(X_t, t) + \left(\frac{\beta}{1+\beta}\right) v_{\theta_0}(X_t, t) \tag{46}$$

While the expectation in the objective function depends on $v_\theta$ through the path measure $\mathbb{P}_\theta$, the pointwise minimizer of the integrand remains consistent. Since the optimizer $v_\theta^*$ is a convex combination of continuous vector fields, the resulting optimal path measure $\mathbb{P}_\theta^*$ is well-defined and achieves the lower bound of the objective.

## G Gradient Equivalence between Total and OT-Conditional KL Divergence

Let $\pi \in \Pi(p_0, q)$ be any coupling between the source and target distributions that admits the marginals $\int \pi(X_0, X_1)dX_1 = p_0(X_0)$ and $\int \pi(X_0, X_1)dX_0 = q(X_1)$. For each drawn pair $(X_0, X_1) \sim \pi$, we define the conditional probability path $p_t(X_t|X_0, X_1)$ with the corresponding conditional vector field $v_\theta(X_t|X_0, X_1)$, which satisfy the continuity equation.

$$\partial_t p_t(X_t|X_0, X_1) + \nabla_{X_t} \cdot (p_t(X_t|X_0, X_1) v_\theta(X_t|X_0, X_1)) = 0 \tag{47}$$

The unconditional path measure $\mathbb{P}_\theta$ is then obtained by marginalizing over $\pi$.

$$p_t(X_t) = \int p_t(X_t|X_0, X_1) d\pi(X_0, X_1) \tag{48}$$

$$v_t(X_t) = \frac{\int v_t(X_t|X_0, X_1) p_t(X_t|X_0, X_1) d\pi(X_0, X_1)}{p_t(X_t)} \tag{49}$$

We use this to define the conditional GFT objective for an arbitrary coupling $\pi$.

$$\mathcal{L}_\pi(\theta) = \mathbb{E}_{(X_0,X_1)\sim\pi}\left[KL(\mathbb{P}_\theta(\cdot|X_0, X_1)\|\mathbb{P}_q(\cdot|X_0, X_1)) + \beta KL(\mathbb{P}_\theta(\cdot|X_0, X_1)\|\mathbb{P}_{\theta_0}(\cdot|X_0, X_1))\right] \tag{50}$$

Here, $\mathbb{P}_\theta$ denotes the path measure induced by the conditional SDE:

$$dX_t = v_\theta(X_t, t|X_0, X_1)dt + \sigma_t dB_t. \tag{51}$$

The conditional path measure $\mathbb{P}_q$ is defined similarly using the unknown conditional target vector field $v_q(X_t, t|X_0, X_1)$. Note that despite this new conditioning, the base process $\mathbb{P}_{\theta_0}$ remains fixed. We apply Proposition 1 to the conditional GFT objective to write it in terms of these conditional vector fields.

$$\mathcal{L}_\pi(\theta) = \mathbb{E}_{(X_0,X_1)\sim\pi}\mathbb{E}_{X_{[0,1]}\sim\mathbb{P}_\theta(\cdot|X_0,X_1)}\Bigg[\frac{1}{2}\int_0^1 \|v_\theta(X_t,t|X_0,X_1) - v_q(X_t,t|X_0,X_1)\|^2 dt + \frac{\beta}{2}\int_0^1 \|v_\theta(X_t,t|X_0,X_1) - v_{\theta_0}(X_t,t|X_0,X_1)\|^2 dt\Bigg] \quad (52)$$

$$= \mathbb{E}_{(X_0,X_1)\sim\pi}\mathbb{E}_{X_{[0,1]}\sim\mathbb{P}_\theta(\cdot|X_0,X_1)}\Bigg[\int_0^1 \frac{1}{2}\Bigg[\|v_\theta(X_t,t|X_0,X_1) - v_q(X_t,t|X_0,X_1)\|^2 dt + \beta\|v_\theta(X_t,t|X_0,X_1) - v_{\theta_0}(X_t,t|X_0,X_1)\|^2\Bigg] dt\Bigg] \quad (53)$$

Following the same logic used to minimize the unconditional objective (Appendix F) the integrand is minimized by the following conditional drift pointwise for each $(X_0, X_1)$, yielding the conditional minimizer.

$$v_\theta^*(X_t,t|X_0,X_1) = \left(\frac{1}{1+\beta}\right) v_q(X_t,t|X_0,X_1) + \left(\frac{\beta}{1+\beta}\right) v_{\theta_0}(X_t,t|X_0,X_1) \quad (54)$$

Note that when the unconditional vector fields $v_q$ and $v_{\theta_0}$ are defined as the marginalization of their conditional counterparts over the same coupling $\pi$, the unconditional optimum is precisely the marginalization of the conditional optimum. However, this optimum is still not usable for training, as it is written in terms of the unknown target vector field $v_q$. We can solve this intractable dependency by following CFM (Lipman et al., 2023), and replacing $v_q$ with a chosen conditional vector field $u_q$ which marginalizes to the true target:

$$\mathbb{E}_{(X_0,X_1)\sim\pi}[u_q(X_t,t|Z)] = v_q(X_t,t). \quad (55)$$

Crucially, the fine-tuned model $v_\theta$ is unconditional, and is trained to approximate the expected marginal optimal vector field

$$v_\theta(X_t,t) \approx \mathbb{E}_{(X_0,X_1)\sim\pi}[v_\theta^*(X_t,t|X_0,X_1)]. \quad (56)$$

Applying Proposition 1 and substituting the chosen tractable conditional target vector field, we get the final GFT loss:

$$\mathcal{L}_\pi(\theta) = \mathbb{E}_{(X_0,X_1)\sim\pi}\mathbb{E}_{X_t\sim\mathbb{P}(\cdot|X_0,X_1)}\left[\frac{1}{2}\int_0^1 (\|v_\theta(X_t,t) - u_q(X_t,t|X_0,X_1)\|^2 + \beta\|v_\theta(X_t,t) - v_{\theta_0}(X_t,t|X_0,X_1)\|^2) dt\right]. \quad (57)$$

As a result, the gradient of the conditional objective $\mathcal{L}_\pi$ with respect to the model parameters $\theta$ will coincide with the gradient of the original unconditional objective $\mathcal{L}$ when the model is parameterized to satisfy the marginalization property. This follows from the same arguments used to justify the original CFM objective (Lipman et al., 2023). Differentiating with respect to the model parameters $\theta$ gives

$$\nabla_\theta\mathcal{L}_\pi(\theta) = \mathbb{E}_{(X_0,X_1)\sim\pi}\mathbb{E}_{X_{[0,1]}\sim\mathbb{P}(\cdot|X_0,X_1)}\Bigg[\int_0^1 \langle (v_\theta(X_t,t) - u_q(X_t,t|X_0,X_1)) + \beta(v_\theta(X_t,t) - v_{\theta_0}(X_t,t|X_0,X_1)), \nabla_\theta v_\theta(X_t,t)\rangle dt\Bigg]. \quad (58)$$

By the construction of the conditional vector fields, for $\phi = \theta, \theta_0, q$, we have

$$v_\phi(X_t, t) = \mathbb{E}_{(X_0, X_1)\sim\pi}\left[v_\phi(X_t, t|X_0, X_1)|X_t\right]. \tag{59}$$

Using these identities and applying the tower property of conditional expectations,

$$\nabla_\theta \mathcal{L}_\pi(\theta) = \mathbb{E}_{X_{[0,1]}\sim\mathbb{P}}\left[\int_0^1 \langle (v_\theta(X_t,t) - \mathbb{E}_\pi[u_q(X_t,t|X_t)]) + \beta(v_\theta(X_t,t) - \mathbb{E}_\pi[v_{\theta_0}(X_t,t|X_t)]), \nabla_\theta v_\theta(X_t,t)\rangle dt\right]. \tag{60}$$

This is exactly the gradient of $\mathcal{L}(\theta)$ when the unconditional objective uses the marginal vector fields, so we can conclude that

$$\nabla_\theta \mathcal{L}_\pi(\theta) = \nabla_\theta \mathcal{L}(\theta). \tag{61}$$

## H Practical Implementation of Gradual Fine-Tuning

Our proposed objective function consists of two KL divergence terms.

$$\mathcal{L}(\theta) = KL(\mathbb{P}_\theta \| \mathbb{P}_q) + \beta KL(\mathbb{P}_\theta \| \mathbb{P}_{\theta_0}) \tag{62}$$

We can therefore apply Proposition 1 to both terms to rewrite the expression.

$$KL(\mathbb{P}_\theta \| \mathbb{P}_q) = \mathbb{E}_{\mathbb{P}_\theta}\left[\frac{1}{2}\int_0^1 \|\sigma_t^{-1}(v_\theta - v_q)\|^2 dt\right] \tag{63}$$

$$KL(\mathbb{P}_\theta \| \mathbb{P}_{\theta_0}) = \mathbb{E}_{\mathbb{P}_\theta}\left[\frac{1}{2}\int_0^1 \|\sigma_t^{-1}(v_\theta - v_{\theta_0})\|^2 dt\right] \tag{64}$$

Here, $v_{\theta_0}$ is the pretrained base vector field, and $v_q$ is the optimal vector field that transports samples from the source distribution $p_0$ to the target fine-tuning distribution $q$. Since this vector field is unknown, it can be defined in the same way that the CFM (Lipman et al., 2023) defines a target vector field to be matched. We show empirically that the diffusion coefficient $\sigma_t$ can simply be set to 1 while maintaining high performance. Note that both terms of the objective are now written using the drift terms of the base and target processes, and that the parameters $\theta$ directly determine the path measure $\mathbb{P}_\theta$. The otherwise intractable objective therefore now optimizes directly over the parameters $\theta$. We can now write a tractable version of the proposed objective.

$$\mathcal{L}(\theta) = \mathbb{E}_{\mathbb{P}_\theta}\left[\frac{1}{2}\int_0^1 \|v_\theta - v_q\|^2 dt\right] + \beta \cdot \mathbb{E}_{\mathbb{P}_\theta}\left[\frac{1}{2}\int_0^1 \|v_\theta - v_{\theta_0}\|^2 dt\right] \tag{65}$$

## I Implementation Details

We use the exact U-Net architecture given by Tong et al. and Tong et al., which has been pretrained on the Cifar-10 dataset Krizhevsky (2009) for 400,000 epochs. Minibatch OT couplings are used during pretraining and fine-tuning of the model, in accordance with the convergence results given in Section 3.3. All GFT experiments use an inverse sigmoid $\beta_s$ cooling schedule.

### I.1 Datasets

Camelyon17 is a dataset of H&E-stained histopathology images of lymph node sections of breast cancer patients. The dataset includes both tumor and non-tumor images. The distribution shift between each split of this dataset is caused experimental batch effects and population variations. The Camelyon17 train set has a total of 302,436 samples, and 60,490 of these samples are held out from training for use in the FID calculations. The Camelyon17 validation set is substantially smaller

than the train dataset, with a total of 34,904 samples, 6,981 of which are held out for use in the FID calculations.

RxRx1 is a dataset of fluorescent microscopy images of human tissue sections. The distribution shift between each split of this dataset is caused experimental batch effects, such as variations in temperature, humidity, and reagent concentration. The RxRx1 experiments are conducted on a compressed version of the dataset to enable running on a single GPU. The RxRx1 train set has a total of 40,612 samples, 8,123 of which are held out for FID calculations. The RxRx1 validation set is smaller than the train dataset, with a total of 9,854 samples, and 1,971 of these samples are held out from fine-tuning for use in the FID calculations.

FMoW is a dataset of global satellite images of regions with varying land use. The experiments shown below are conducted on the RGB image version of the dataset, rather than the multispectral data. The distribution shift between each split of this dataset is caused temporal evolution. The FMoW experiments are conducted on a compressed version of the dataset to enable running on a single GPU. The FMoW train set has a total of 76,863 samples, and 15,373 are held out for use in the FID calculations. The FMoW validation set has a total of 19,915 samples, 3,983 of which are held out from fine-tuning for FID calculations.

**Stability and Convergence Analysis** We report three metrics to quantitatively analyze the convergence speed and stability of GFT. *Instantaneous Variance* measures the local variance in accuracy convergence. At 50 epoch intervals, the instantaneous mean is found using a Radial Basis Function (RBF) kernel over a window of 10 FID measurements. This mean is then used to calculate variance with the same weighting and window size. The reported value is the average of all instantaneous variances over the scope of fine-tuning. *Average convergence rate* is calculated by performing linear regression on windows of 10 consecutive FID measurements. The reported value is the average of the absolute value of the resulting slopes. Finally, *Spearman correlation* is calculated between fine-tuning epochs and FID. For a stable fine-tuning process, we expect a strong negative correlation. A Spearman correlation coefficient close to -1 would indicate that as fine-tuning progresses, the model performance consistently improves with minimal regressive spikes or oscillation.

**Path Length Calculation** As a proxy to estimate the generation efficiency of each method we test, we report average probability path length. We can define the path length of a single sample through a vector field as its energy cost, or its accumulated kinetic energy over time. By the Benamou-Brenier formulation of optimal transport, this can be calculated by integrating the square of the velocity magnitude over the path of the given sample.

$$L(x_0) = \int_0^1 \|\frac{dx}{dt}\|^2 dt = \int_0^1 \|\mathbf{v}(x_t, t)\|^2 dt \tag{66}$$

Given a prior distribution over initial samples $p(x_0)$, we can get the expected length of the vector field.

$$\mathbb{E}_{p(x_0)}[L(x_0)] = \mathbb{E}_{x_0 \sim p(x_0)} \left[ \int_0^1 \|\mathbf{v}(x_t, t)\|^2 dt \right] \tag{67}$$

The expectation over the prior distribution $p_0$ can be approximated by averaging over a minibatch of $M$ samples. We can apply Euler discretization with some chosen step size $\Delta t$ to approximate the integral.

$$L(\theta) = \frac{1}{M} \sum_{j=1}^{M} \sum_{i=1}^{N} \|\mathbf{v}_\theta(x_i^{(j)}, t_i)\|^2 \cdot \Delta t \tag{68}$$

We estimate average path length for a model with parameters $\theta$ with 100 discretization steps and 64 initial samples.

# J Additional Results

## J.1 Extended Results on Large Distribution Shifts

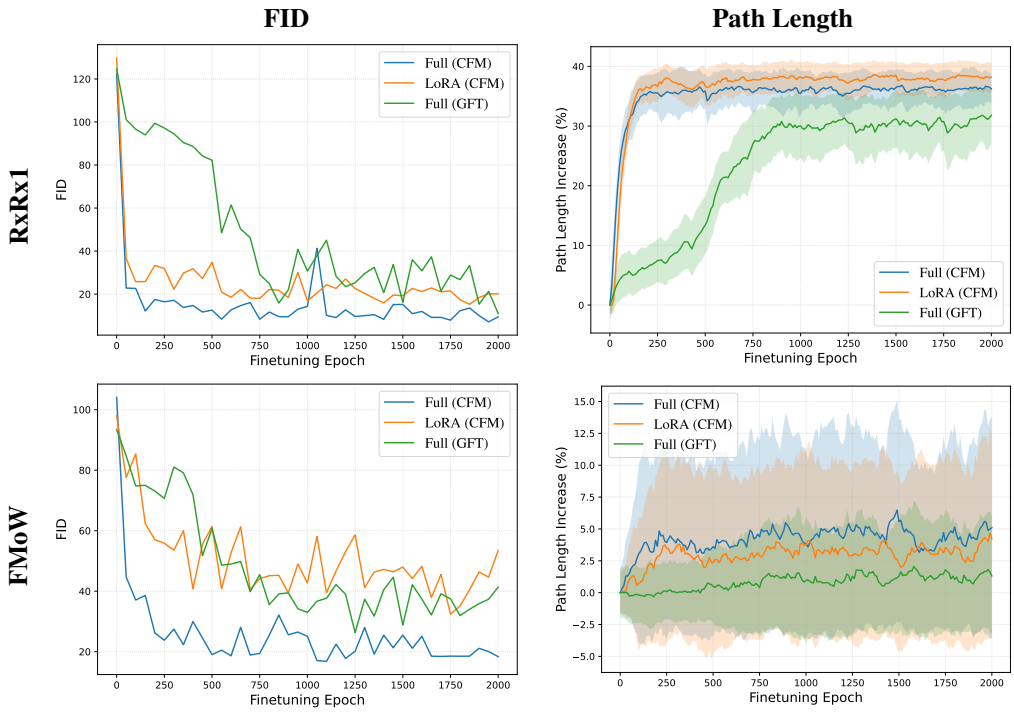



Figure 5: Fine-tuning performance from a pretrained Cifar-10 model to various WILDS train datasets. Rows corresponding to individual datasets, and the columns show FID and probability path lengths through 2000 epochs of fine-tuning. The shaded regions of the path length graphs correspond to one standard deviation from the mean.

### J.2 Extended Results on Small Distribution Shifts

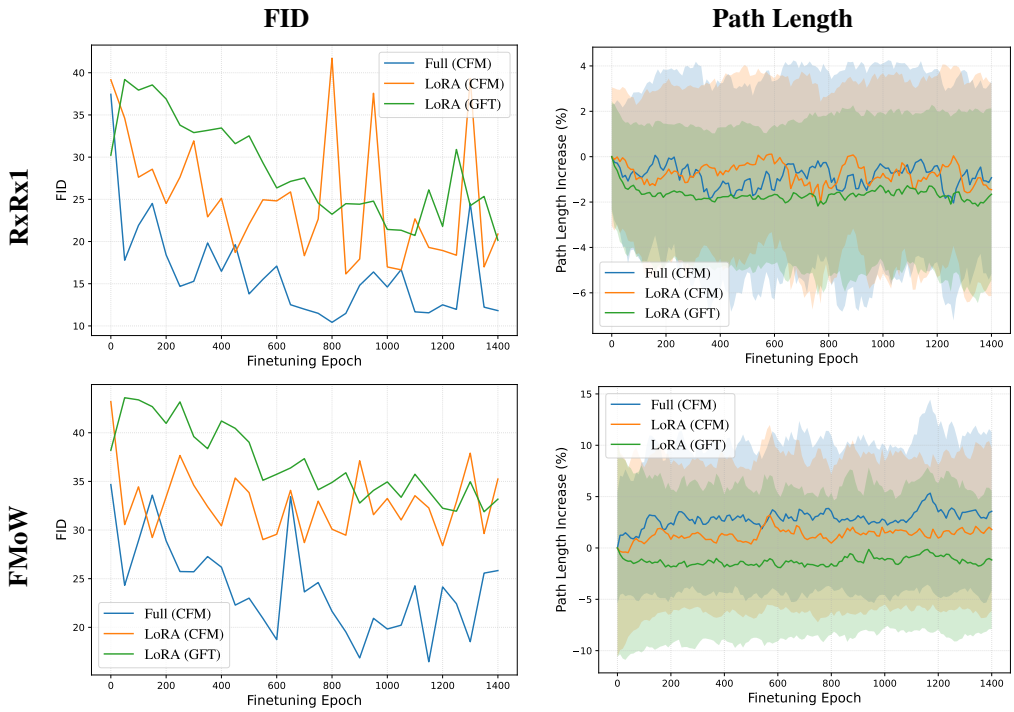



Figure 6: Fine-tuning performance from a fully fine-tuned model to two WILDS validation datasets. Rows corresponding to individual datasets, and the columns show FID and probability path lengths through fine-tuning. The shaded region of the path length graphs show one standard deviation from the mean.

### J.3 GENERATED IMAGES

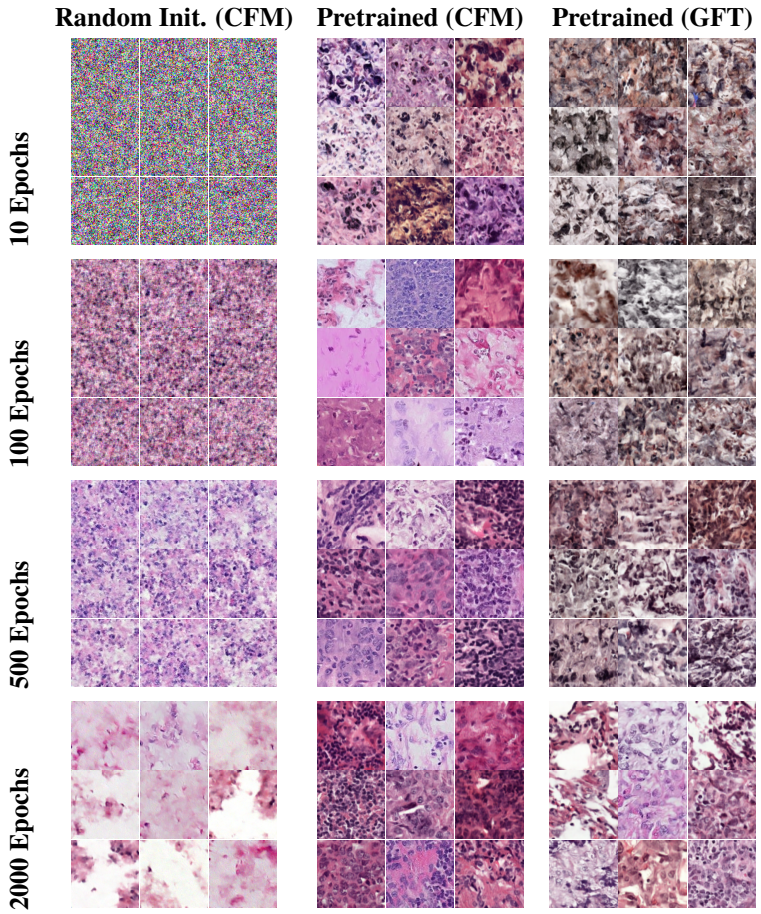



Figure 7: Generated images at various training stages from models trained or fine-tuned on the Camelyon17 train dataset. Rows correspond to the number of training/fine-tuning steps, and columns correspond to training initialization and objective.

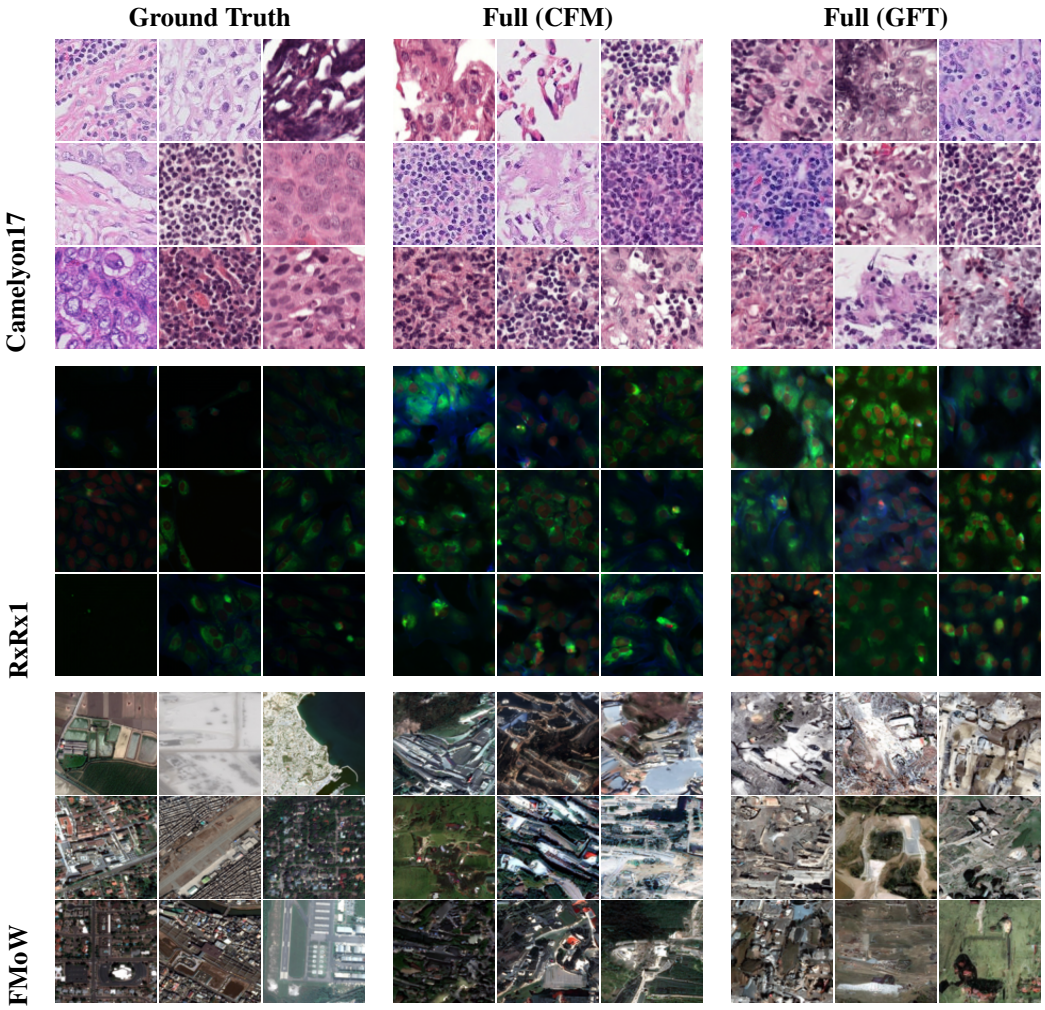



Figure 8: Generated images after 2000 epochs of fine-tuning on three WILDS datasets, starting from a pretrained Cifar-10 model.

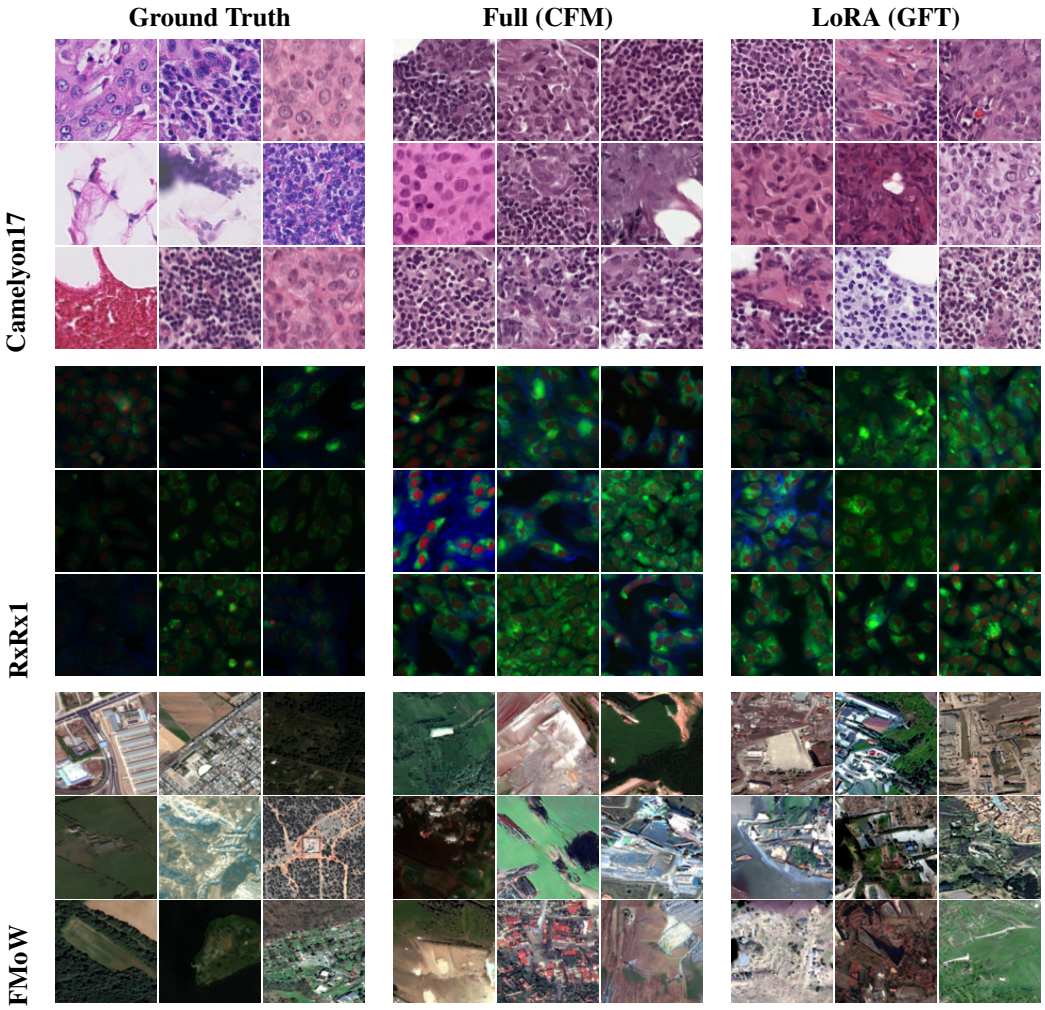



Figure 9: Generated images after fine-tuning on the validation splits of three WILDS datasets. Fine-tuning starts from a pretrained Cifar-10 model which has been fully fine-tuned on the train set of the corresponding WILDS dataset.

### J.4 Extended Stability and Convergence Results

Table 2: Convergence stability comparison over large distribution shifts. Each method starts from a pretrained Cifar-10 model, and is fine-tuned on a WILDS train set.

| | **Objective** | **Inst. Variance** $\downarrow$ | **Avg. Conv. Rate** $\uparrow$ | **Spearman** $\rho \downarrow$ |
|---|---|---|---|---|
| RxRx1 | Full (CFM) | 191.685 | $3.110 \times 10^{-2}$ | -0.560 |
| | LoRA (CFM) | 166.996 | $3.058 \times 10^{-2}$ | -0.602 |
| | Full (GFT) | **117.470** | $\mathbf{4.820 \times 10^{-2}}$ | **-0.800** |
| FMoW | Full (CFM) | 90.369 | $\mathbf{2.827 \times 10^{-2}}$ | -0.584 |
| | LoRA (CFM) | 66.178 | $2.154 \times 10^{-2}$ | -0.552 |
| | Full (GFT) | **43.226** | $2.419 \times 10^{-2}$ | **-0.759** |

Table 3: Convergence stability comparison over small distribution shifts. Each method starts from the base Cifar-10 model after it has been fully fine-tuned on a WILDS train set, and is fine-tuned for the corresponding validation set.

| | **Objective** | **Inst. Variance** $\downarrow$ | **Avg. Conv. Rate** $\uparrow$ | **Spearman** $\rho \downarrow$ |
|---|---|---|---|---|
| RxRx1 | Full (CFM) | 10.654 | $9.779 \times 10^{-3}$ | -0.165 |
| | LoRA (CFM) | 45.366 | $\mathbf{1.014 \times 10^{-2}}$ | -0.329 |
| | LoRA (GFT) | **2.347** | $3.743 \times 10^{-3}$ | **-0.518** |
| FMoW | Full (CFM) | 12.334 | $4.456 \times 10^{-3}$ | -0.550 |
| | LoRA (CFM) | 9.236 | $1.406 \times 10^{-3}$ | -0.125 |
| | LoRA (GFT) | **2.881** | $\mathbf{6.157 \times 10^{-3}}$ | **-0.884** |

## K Reward vs Sample-based Fine-Tuning Distribution

Reward-based fine-tuning methods assume access to a known reward model $r(x)$, and the goal of fine-tuning is to train a model to generate samples from the tilted distribution

$$p^*(x) \propto p^{\text{base}}(x) \exp(r(x))$$

In the case where we are directly given a fine-tuning target distribution $q$ (in the form of a dataset of samples from $q$), our goal is now to train a model which generates samples from

$$q(x) = p(x)w(x), \qquad w(x) = \frac{q(x)}{p(x)} \tag{69}$$

In practice, we do not know the true target distribution $q$. However, we can prove that the density ratio $w$ is exactly the Radon-Nikodym derivative.

*proof.*

Let $\mathbb{P}_p$ and $\mathbb{P}_q$ be the path measures associated with the continuous processes which generate samples from the base and fine-tuned target distributions $p$ and $q$, respectively. In order to apply the Radon-Nikodym theorem, we assume that $\mathbb{P}_q$ is absolutely continuous with respect to $\mathbb{P}_p$. Then, there exists a unique function $w(x)$ such that

$$\mathbb{P}_q(A) = \int_A w(x) d\mathbb{P}_p$$

where $A$ is any measurable set in the sample space. The function $w$ is the Radon-Nikodym derivative, and is denoted as

$$w = \frac{d\mathbb{P}_q}{d\mathbb{P}_p}$$

Given that $\mathbb{P}_p, \mathbb{P}_q$ are defined over a space with an underlying Lebesgue base measure $\lambda$, we can write them in terms of their probability density functions.

$$\mathbb{P}_p = \int p(x) d\lambda \qquad \mathbb{P}_q = \int q(x) d\lambda$$
$$p(x) = \frac{d\mathbb{P}_p}{d\lambda} \qquad q(x) = \frac{d\mathbb{P}_q}{d\lambda}$$

We can then apply the chain rule for Radon-Nikodym derivatives to recover the density ratio.

$$w(x) = \frac{d\mathbb{P}_q}{d\mathbb{P}_p}(x) = \frac{d\mathbb{P}_q/d\lambda}{d\mathbb{P}_p/d\lambda}(x) = \frac{q(x)}{p(x)}$$

## L Corresponding Reward Function

Reward-based methods fine-tune the base model to achieve generation from a tilted distribution defined by an external reward function $r(x)$.

$$p^*_\theta(x) \propto p_{\theta_0}(x) \exp(r(x)) \tag{70}$$

The reward function which results in an equivalent optimum to our proposed objective function is

$$r(x) = \frac{1}{1+\beta} \log\left(\frac{q(x)}{p_{\theta_0}(x)}\right) \tag{71}$$

We can easily verify this by substituting this expression into the tilted distribution.

$$p^*_\theta(x) \propto p_{\theta_0}(x) \exp\left(\frac{1}{1+\beta} \log\left(\frac{q(x)}{p_{\theta_0}(x)}\right)\right) \tag{72}$$

$$= p_{\theta_0}(x) \left(\frac{q(x)}{p_{\theta_0}(x)}\right)^{\frac{1}{1+\beta}} \tag{73}$$

$$= p_{\theta_0}(x)^{\frac{\beta}{1+\beta}} q(x)^{\frac{1}{1+\beta}} \tag{74}$$

Since the distribution $q(x)$ is unknown, this reward function is not tractable in it's current form. However, in Appendix K we prove the relationship between the density ratio and the Radon-Nikodym derivative, and in Appendix D we give the expression for the logarithm of the Radon-Nikodym derivative. Putting these together, we can rewrite the reward function as:

$$r(x) = \frac{1}{1+\beta} \left(\int_0^1 \sigma_t^{-1}(v_q - v_{\theta_0})^\intercal dB_t - \frac{1}{2}\int_0^1 \|\sigma_t^{-1}(v_q - v_{\theta_0})\|^2 dt\right) \tag{75}$$

The reward function is now written in terms of the drift terms of the base, target, and fine-tuned processes. As a result, this reward is only tractable when it is conditioned on given source and target samples. This is because the vector field $v_q$ is unknown, and can only be defined conditionally given known samples. In the setting where a memoryless noise schedule is required to achieve unbiased convergence, this reward function is therefore intractable.

## M Equivalence Between Total and Source-Conditional KL Divergence

Let $\mathbb{P}_f$ and $\mathbb{P}_g$ be the path measures of two SDEs on the time interval $[0, 1]$, with initial samples from the same source distribution and the same diffusion coefficient $\sigma_t$.

$$\mathbb{P}_f: \quad dX_t = f_t(X_t, t)dt + \sigma_t dB_t, \qquad X_0 \sim p_0 \tag{76}$$

$$\mathbb{P}_g: \quad dX_t = g_t(X_t, t)dt + \sigma_t dB_t, \qquad X_0 \sim p_0 \tag{77}$$

We assume that the standard conditions of Girsanov's Theorem hold, such that $\mathbb{P}_f$ and $\mathbb{P}_g$ are mutually absolutely continuous. The path measure $\mathbb{P}$ for any stochastic process starting at the state $X_0$ can be decomposed into the product of the initial distribution $p_0(X_0)$ and the conditional law of the paths given this starting point, $\mathbb{P}(\cdot|X_0)$.

$$\mathbb{P}(\cdot) = \int \mathbb{P}(\cdot|X_0)p_0(X_0)dX_0. \tag{78}$$

The path measure $\mathbb{P}$ then admits a disintegration with respect to $X_0$.

$$d\mathbb{P} = p_0(X_0)dX_0 \cdot d\mathbb{P}(A|X_0) \tag{79}$$

Since the path measures $\mathbb{P}_f$ and $\mathbb{P}_g$ share the same source distribution, applying this chain rule to both yields

$$d\mathbb{P}_f = p_0(X_0)dX_0 \cdot d\mathbb{P}_f(\cdot|X_0) \tag{80}$$

$$d\mathbb{P}_g = p_0(X_0)dX_0 \cdot d\mathbb{P}_g(\cdot|X_0) \tag{81}$$

We can then write the expression for the conditional Radon-Nikodym derivative of $\mathbb{P}_f$ and $\mathbb{P}_g$, pointwise over $X_0$.

$$\frac{d\mathbb{P}_f}{d\mathbb{P}_g} = \frac{p_0(X_0)}{p_0(X_0)} \times \frac{d\mathbb{P}_f(\cdot|X_0)}{d\mathbb{P}_g(\cdot|X_0)} \tag{82}$$

$$\frac{d\mathbb{P}_f}{d\mathbb{P}_g} = \frac{d\mathbb{P}_f(\cdot|X_0)}{d\mathbb{P}_g(\cdot|X_0)} \tag{83}$$

Therefore, the Radon-Nikodym derivative of the path measures conditioned on a given source sample $X_0$ is equivalent to the total derivative when the path measures share the same initial distribution. It follows from the proof given in appendix D that

$$KL(\mathbb{P}_f(\cdot|X_0)\|\mathbb{P}_g(\cdot|X_0)) = -\mathbb{E}_{\mathbb{P}_f(\cdot|X_0)}\left[\int_0^1 \sigma_t^{-1}(g_t - f_t)^\intercal dB_t\right] + \mathbb{E}_{\mathbb{P}_f(\cdot|X_0)}\left[\frac{1}{2}\int_0^1 \|\sigma_t^{-1}(g_t - f_t)\|^2 dt\right] \tag{84}$$

The stochastic integral is a Martingale under the path measure $\mathbb{P}_f$, and its expectation vanishes.

$$KL(\mathbb{P}_f(\cdot|X_0)\|\mathbb{P}_g(\cdot|X_0)) = \mathbb{E}_{\mathbb{P}_f(\cdot|X_0)}\left[\frac{1}{2}\int_0^1 \|\sigma_t^{-1}(g_t - f_t)\|^2 dt\right] \tag{85}$$

We can take the expectation over initial states $X_0$ to get

$$\mathbb{E}_{X_0\sim p_0}\left[KL(\mathbb{P}_f(\cdot|X_0)\|\mathbb{P}_g(\cdot|X_0))\right] = \mathbb{E}_{X_0\sim p_0}\left[\mathbb{E}_{\mathbb{P}_f(\cdot|X_0)}\left[\frac{1}{2}\int_0^1 \|\sigma_t^{-1}(g_t - f_t)\|^2 dt\right]\right] \tag{86}$$

By the Law of Total Expectation, nesting these expectations is equivalent to taking the expectation over the full path measure $\mathbb{P}_f$. We therefore find the final identity

$$KL(\mathbb{P}_f\|\mathbb{P}_g) = \mathbb{E}_{X_0\sim p_0}\left[KL(\mathbb{P}_f(\cdot|X_0)\|\mathbb{P}_g(\cdot|X_0))\right] = \mathbb{E}_{\mathbb{P}_f}\left[\frac{1}{2}\int_0^1 \|\sigma_t^{-1}(g_t - f_t)\|^2 dt\right] \tag{87}$$

## N Equivalence Between Total and OT-Conditional KL Divergence

Let $\mathbb{P}_f$ and $\mathbb{P}_g$ be the path measures of two SDEs on the time interval $[0, 1]$, with initial samples from the same source distribution and the same diffusion coefficient $\sigma_t$.

$$\mathbb{P}_f: \quad dX_t = f_t(X_t, t)dt + \sigma_t dB_t, \qquad X_0 \sim p_0 \tag{88}$$

$$\mathbb{P}_g : \quad dX_t = g_t(X_t, t)dt + \sigma_t dB_t, \qquad X_0 \sim p_0 \tag{89}$$

We assume that the standard conditions of Girsanov's Theorem hold, such that $\mathbb{P}_f$ and $\mathbb{P}_g$ are mutually absolutely continuous. Let $\pi \in \Pi(p_0, q)$ be any coupling between the source and target distributions that admits the marginals $\int \pi(X_0, X_1)dX_1 = p_0(X_0)$ and $\int \pi(X_0, X_1)dX_0 = q(X_1)$. Consider the following expected KL divergence over jointly drawn samples:

$$\mathbb{E}_{(X_0,X_1)\sim\pi}\left[KL(\mathbb{P}_f(\cdot|X_0)\|\mathbb{P}_g(\cdot|X_0, X_1))\right] \tag{90}$$

Using the definition of $\pi$, we can write the integral form of this expectation.

$$\mathcal{L}_\pi = \int\int KL(\mathbb{P}_f(\cdot|X_0)\|\mathbb{P}_g(\cdot|X_0, X_1))\pi(X_0, X_1)dX_0dX_1 \tag{91}$$

Expanding the definition of the KL divergence for path measures:

$$\mathcal{L}_\pi = \int \pi(X_0, X_1)\left(\int \log \frac{d\mathbb{P}_f(\cdot|X_0)}{d\mathbb{P}_g(\cdot|X_0, X_1)} d\mathbb{P}_f(\cdot|X_0)\right) dX_0dX_1 \tag{92}$$

For the target process $g$, the relationship between the measure $\mathbb{P}_g(\cdot|X_0, X_1)$ and the source-conditioned measure $\mathbb{P}_g(\cdot|X_0)$ is given by Bayes Theorem,

$$d\mathbb{P}_g(\cdot|X_0, X_1) = \frac{d\mathbb{P}_g(\cdot|X_0)p_g(X_1|\cdot)}{p_g(X_1|X_0)} \tag{93}$$

where $p_g(X_1|\cdot)$ is the density of the terminal state of a path. Since we condition directly on $X_1$, this terminal-state likelihood evaluates to a Dirac measure, which is independent of $f$. We now substitute this into the Radon-Nikodym derivative.

$$\log \frac{d\mathbb{P}_f(\cdot|X_0)}{d\mathbb{P}_g(\cdot|X_0, X_1)} = \log \frac{d\mathbb{P}_f(\cdot|X_0)}{d\mathbb{P}_g(\cdot|X_0)} + \log \frac{p_g(X_1|X_0)}{p_g(X_1|\cdot)} \tag{94}$$

Finally, we plug this back into the coupled objective and distribute the expectation.

$$\mathcal{L}_\pi = \mathbb{E}_\pi\left[KL(\mathbb{P}_f(\cdot|X_0)\|\mathbb{P}_g(\cdot|X_0))\right] + \mathbb{E}_\pi\mathbb{E}_{\mathbb{P}_f(\cdot|X_0)}\left[\log \frac{p_g(X_1|X_0)}{p_g(X_1|\cdot)}\right] \tag{95}$$

$$= \int p_0(X_0)KL(\mathbb{P}_f(\cdot|X_0)\|\mathbb{P}_g(\cdot|X_0))dX_0 + \mathcal{C}(\pi, g) \tag{96}$$

## O TERMINAL DISTRIBUTION OBJECTIVE

We examine a related objective function, which works on the terminal distributions of the generative process rather than the path measures. Consider the following constrained SDE,

$$\min_{p_\theta} \left(KL(p_\theta\|q) + \beta KL(p_\theta\|p_{\theta_0})\right) \tag{97}$$

$$\text{s.t. } dX_t = v_\theta(X_t, t)dt + \sigma_t dB_t \tag{98}$$

where $p_\theta$ and $p_{\theta_0}$ are the terminal distributions from which the fine-tuned and pretrained models generates samples, respectively, and $\beta$ is a scalar hyperparameter. The objective function corresponding to this constraint is

$$\mathcal{L}(p_\theta) = KL(p_\theta\|q) + \beta KL(p_\theta\|p_{\theta_0}). \tag{99}$$

We want to show that minimizing this objective function results in sampling from the desired target distribution $q$. We start by expressing the constraint in integral form with respect to the Lebesgue base measure $\lambda$ for the continuous Euclidean space $\mathbb{R}^d$.

$$\mathcal{L}(p_\theta) = \int p_\theta(X) \log \frac{p_\theta(X)}{q(X)} d\lambda(X) + \int \beta \cdot p_\theta(X) \log \frac{p_\theta(X)}{p_{\theta_0}(X)} d\lambda(X) \tag{100}$$

We minimize $\mathcal{L}(p_\theta)$ subject to the constraint that $p_\theta$ is a valid probability distribution: $\int_{\mathcal{X}} p_\theta(X) d\lambda(X) = 1$. To do so, we introduce a Lagrange multiplier $\mu$ for this constraint, and form the Lagrangian $L(p_\theta)$.

$$L(p_\theta) = \mathcal{L}(p_\theta) - \mu \left( \int_{\mathcal{X}} p_\theta(X) d\lambda(X) - 1 \right) \tag{101}$$

$$L(p_\theta) = \int_{\mathcal{X}} \left[ p_\theta(X) \log \frac{p_\theta(X)}{q(X)} + \beta \cdot p_\theta(X) \log \frac{p_\theta(X)}{p_{\theta_0}(X)} - \mu p_\theta(X) \right] d\lambda(X) + \mu \tag{102}$$

We now minimize by setting the derivative with respect to $p_\theta$ to zero, and solving to find the optimal $p^*_\theta$.

$$\frac{dL}{dp_\theta} = (\log p_\theta + 1 - \log q) + \beta(\log p_\theta + 1 - \log p_{\theta_0}) - \mu = 0 \tag{103}$$

Finally, we solve for the optimal $p^*_\theta(X)$.

$$(1+\beta) \log p_\theta + 1 + \beta - \log q - \beta \log p_{\theta_0} - \mu = 0 \tag{104}$$

$$(1+\beta) \log p_\theta = \log q + \beta \log p_{\theta_0} + \mu - 1 - \beta \tag{105}$$

$$\log p_\theta = \frac{1}{1+\beta}(\log q) + \frac{\beta}{1+\beta}(\log p_{\theta_0}) + \frac{\mu - 1 - \beta}{1+\beta} \tag{106}$$

$$p_\theta = \exp\left( \frac{1}{1+\beta}(\log q) + \frac{\beta}{1+\beta}(\log p_{\theta_0}) + \frac{\mu - 1 - \beta}{1+\beta} \right) \tag{107}$$

$$p^*_\theta = q^{\frac{1}{1+\beta}} \cdot p_{\theta_0}^{\frac{\beta}{1+\beta}} \exp\left( \frac{\mu - 1 - \beta}{1+\beta} \right) \tag{108}$$

We set $Z$ to be a constant that absorbs all terms not related to $X$. Then, the final optimal distribution is

$$p^*_\theta(X) = Z \cdot q(X)^{\frac{1}{1+\beta}} \cdot p_{\theta_0}(X)^{\frac{\beta}{1+\beta}}. \tag{109}$$

**Cooling Schedule** Note that the exponents $\frac{1}{1+\beta}$ and $\frac{\beta}{1+\beta}$ sum to 1, meaning that we have defined a weighted average of the pretrained and fine-tuned target distributions. This means that hyperparameter $\beta$ can be adjusted to achieve a fine-tuned model that generates from a user-defined target distribution.

$$p^*_\theta(X) \propto \begin{cases} q(X) & \text{if } \beta = 0 \\ \sqrt{p_{\theta_0}(X) q(X)} & \text{if } \beta = 1 \\ p_{\theta_0}(X) & \text{if } \beta = \infty \end{cases} \tag{110}$$

This means that the second KL term is acting as a regularization for the change from the pretrained model, and the hyperparameter $\beta$ controls the strength of this regularization. $\beta$ can therefore be used as a cooling schedule for simulated annealing (Kirkpatrick et al., 1983) from the pretrained distribution $p_{\theta_0} \approx p$ to the fine-tuned target distribution $p_\theta \approx q$.