BELLMAN OPTIMAL STEP-SIZE STRAIGHTENING OF FLOW-MATCHING MODELS

Bao Nguyen, Binh Nguyen, Viet Anh Nguyen

Velocity function

The estimator v_θ allows us to flow from the distribution π_0 (noises) to the distribution π_1 (real images) through the following equation:

$$X_1 = X_0 + \int_0^1 v_\theta(X_t, t)dt, \tag{1}$$

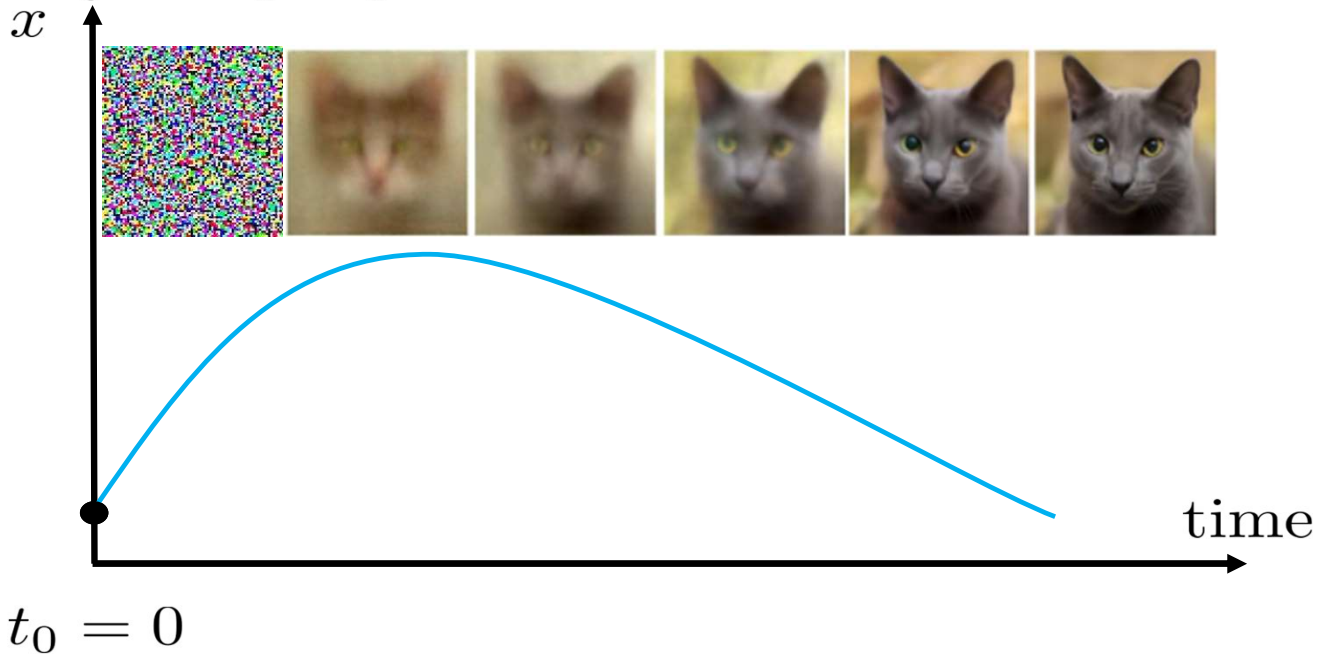
where $X_0 \sim \pi_0$ and $X_1 \sim \pi_1$.

Velocity function

The estimator v_θ allows us to flow from the distribution π_0 (noises) to the distribution π_1 (real images) through the following equation:

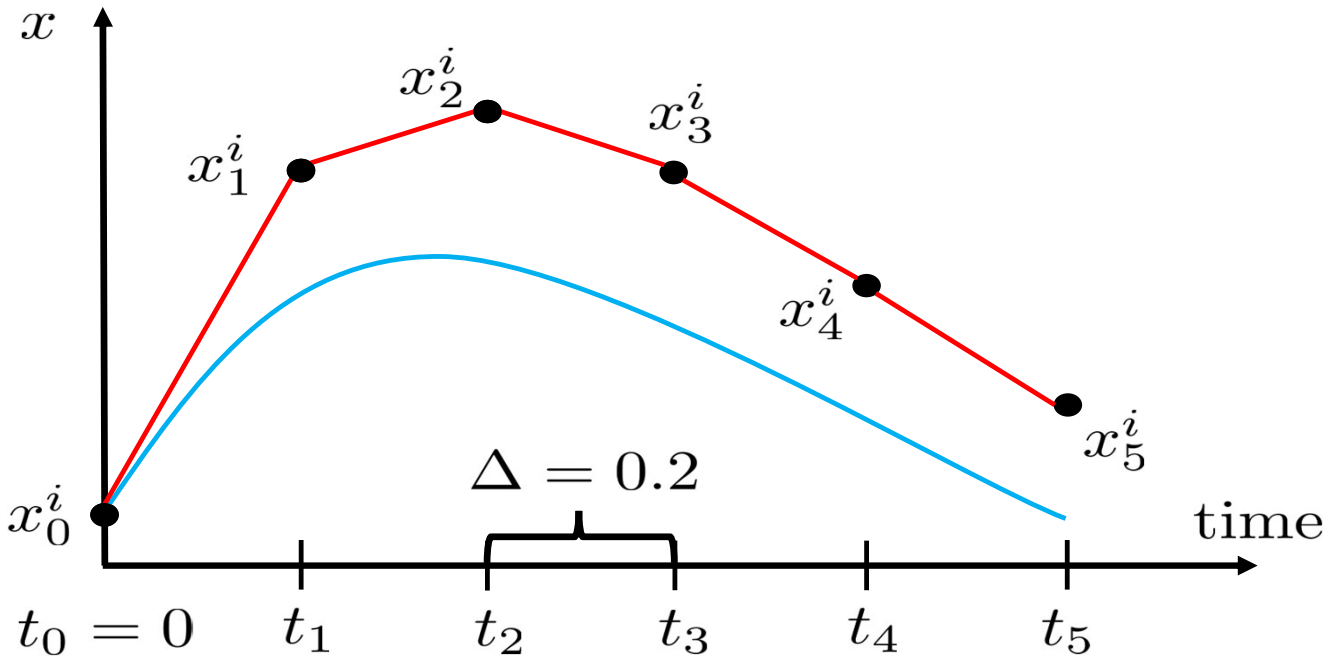
$$X_1 = X_0 + \int_0^1 v_\theta(X_t, t)dt, \tag{1}$$

where $X_0 \sim \pi_0$ and $X_1 \sim \pi_1$.



Euler method

$$x^i_{k/K} = x^i_{(k-1)/K} + v_\theta(x^i_{(k-1)/K}, (k-1)\Delta) \times \Delta \quad \forall k = 1, \ldots, K,$$



Contributions

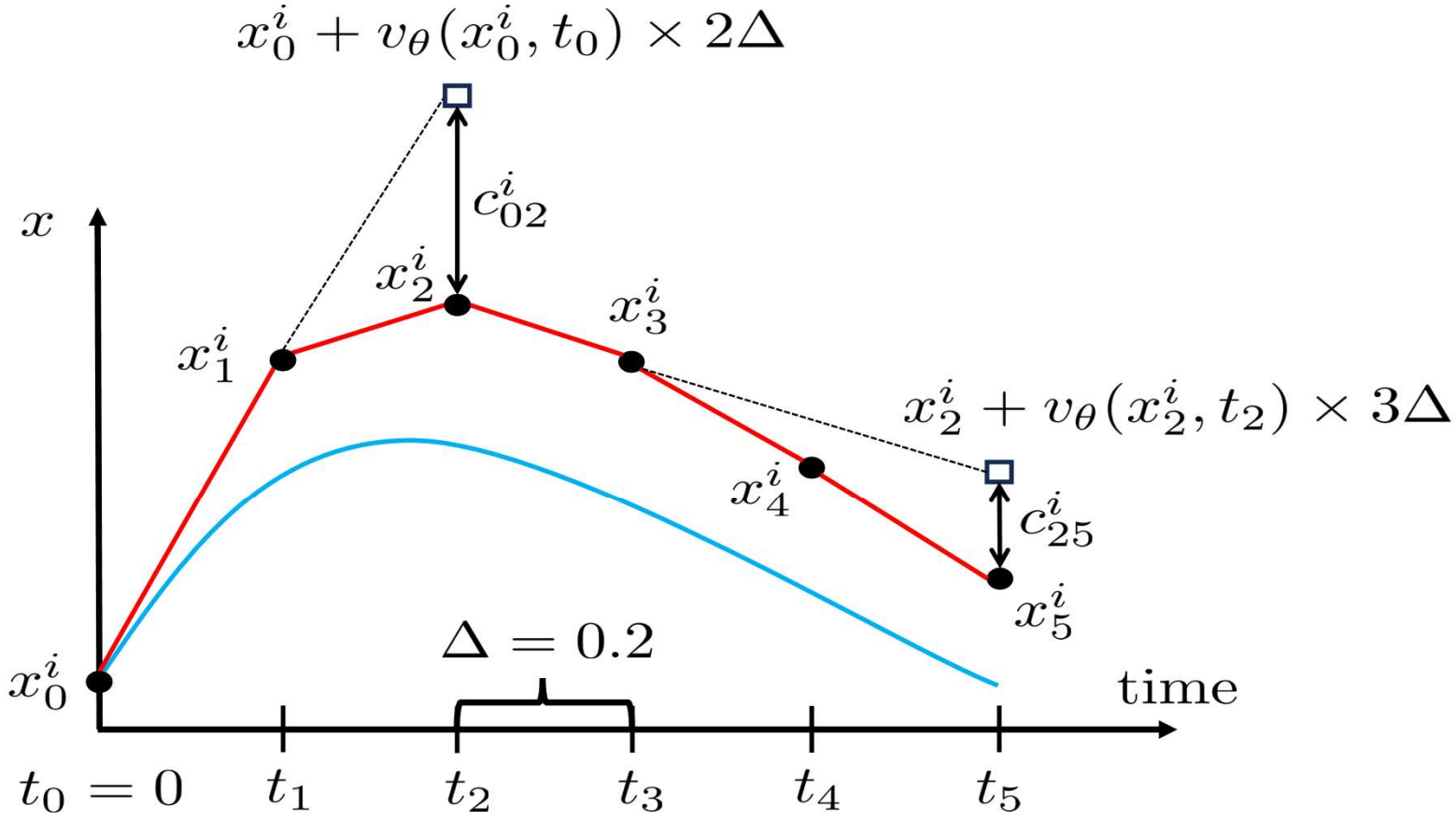
Our work focuses on fine-tuning flow-matching models to generate **high-fidelity** images for any target of K number of function evaluations (NFEs) in **low-resource** scenarios.

The main contribution is a two-phase algorithm where:

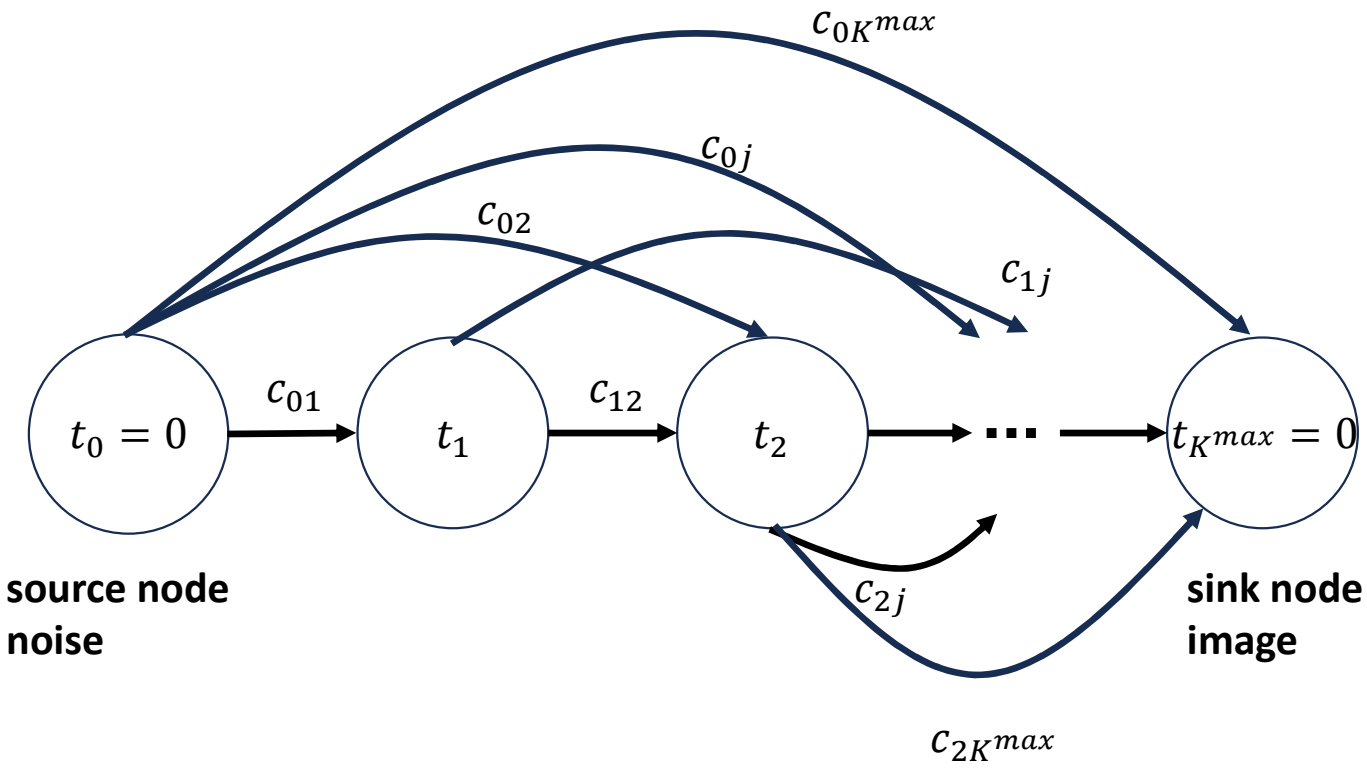
- Phase 1: A **training-free** algorithm based on dynamic programming to find optimal time schedules for sampling which takes **2 minutes** for resolution 256×256.
- Phase 2: A fine-tuning method based on found time schedules, which take **12 hours on an A5000 GPU.**

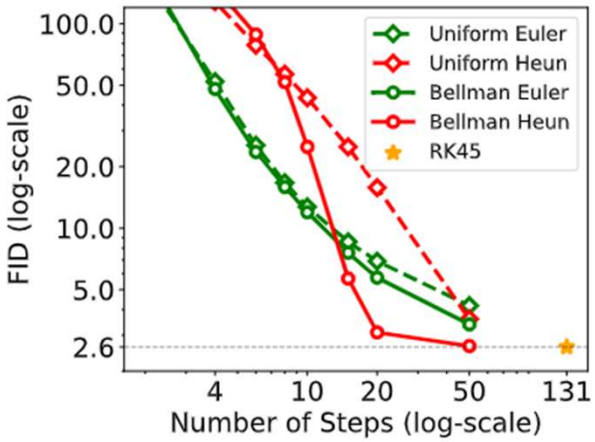
Sampling error estimation

$$c_{jk} = \frac{1}{N}\sum_{i=1}^{N} c_{jk}^{i}, \quad \text{where} \quad c_{jk}^{i} = \|x_{t_k}^{i} - x_{t_j}^{i} - v_\theta(x_{t_j}^{i}, t_j) \times (t_k - t_j)\|_2^2,$$

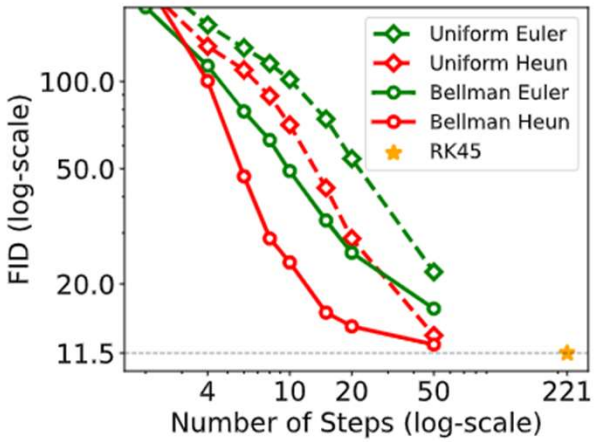


Optimal sampling stepsizes

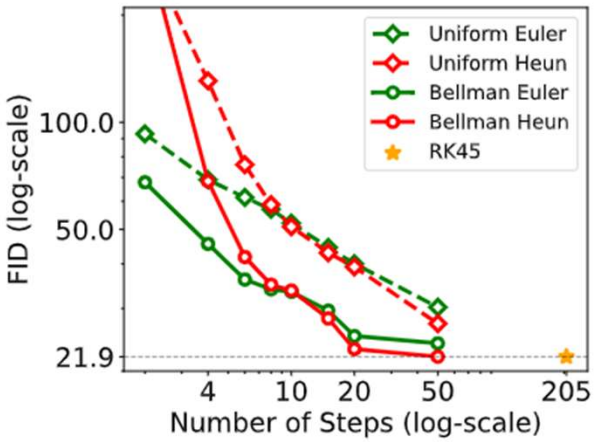




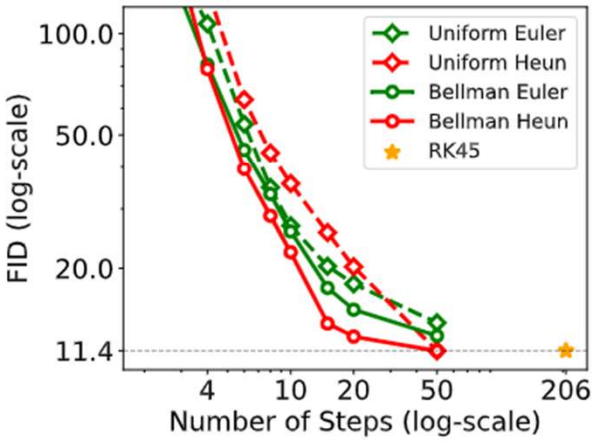
(a) CIFAR-10



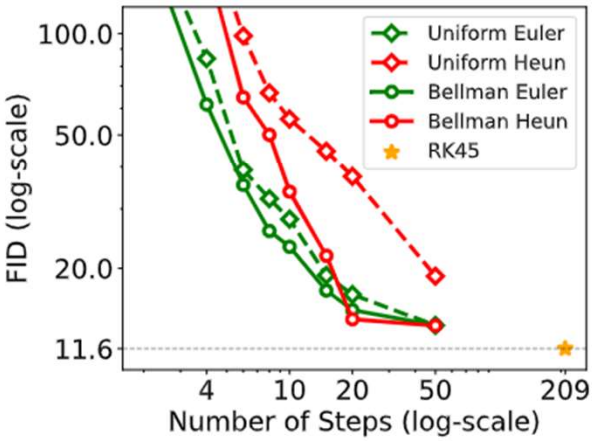
(b) CELEBA-HQ



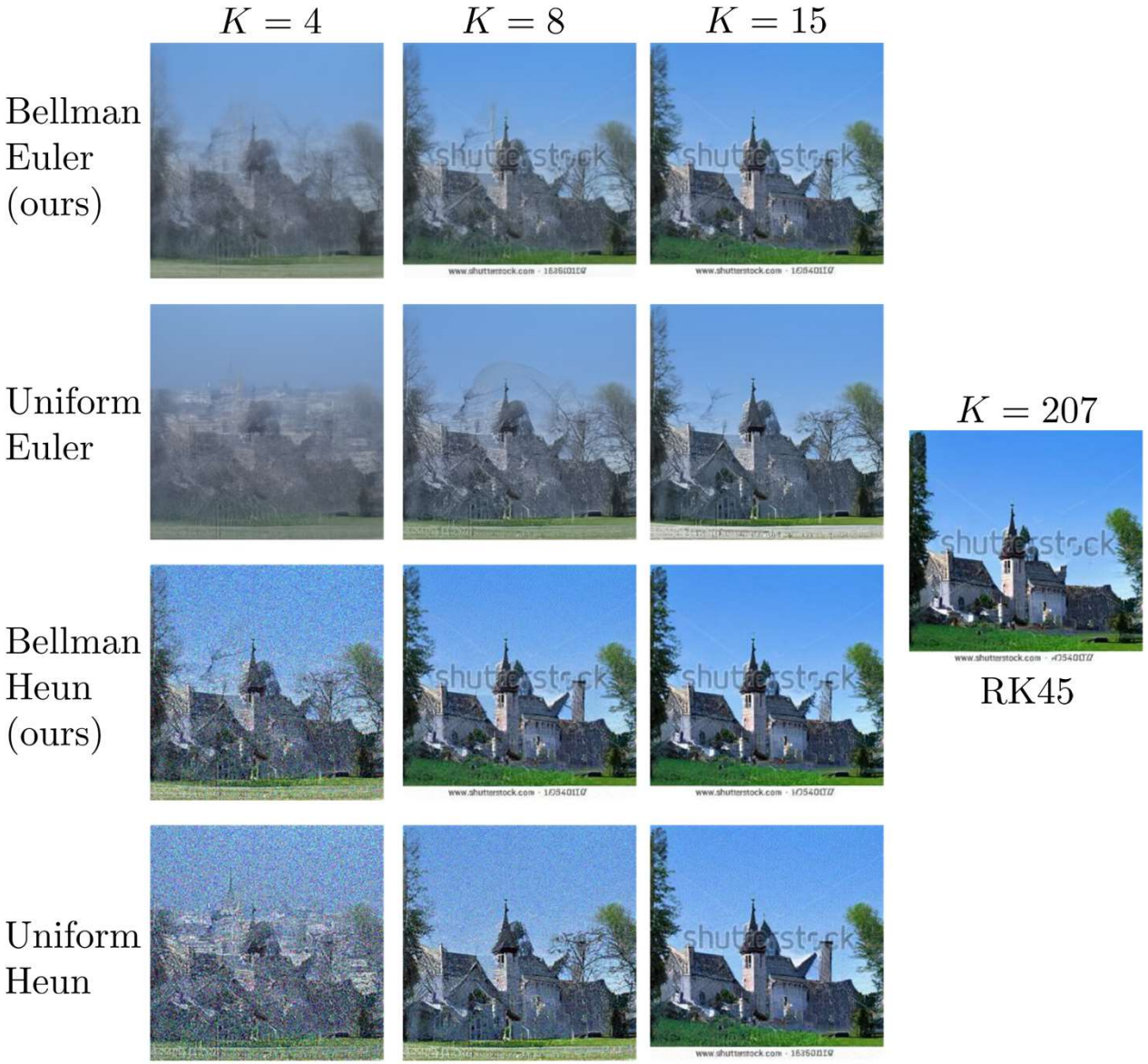
(c) AFHQ-Cat



(d) LSUN-Church

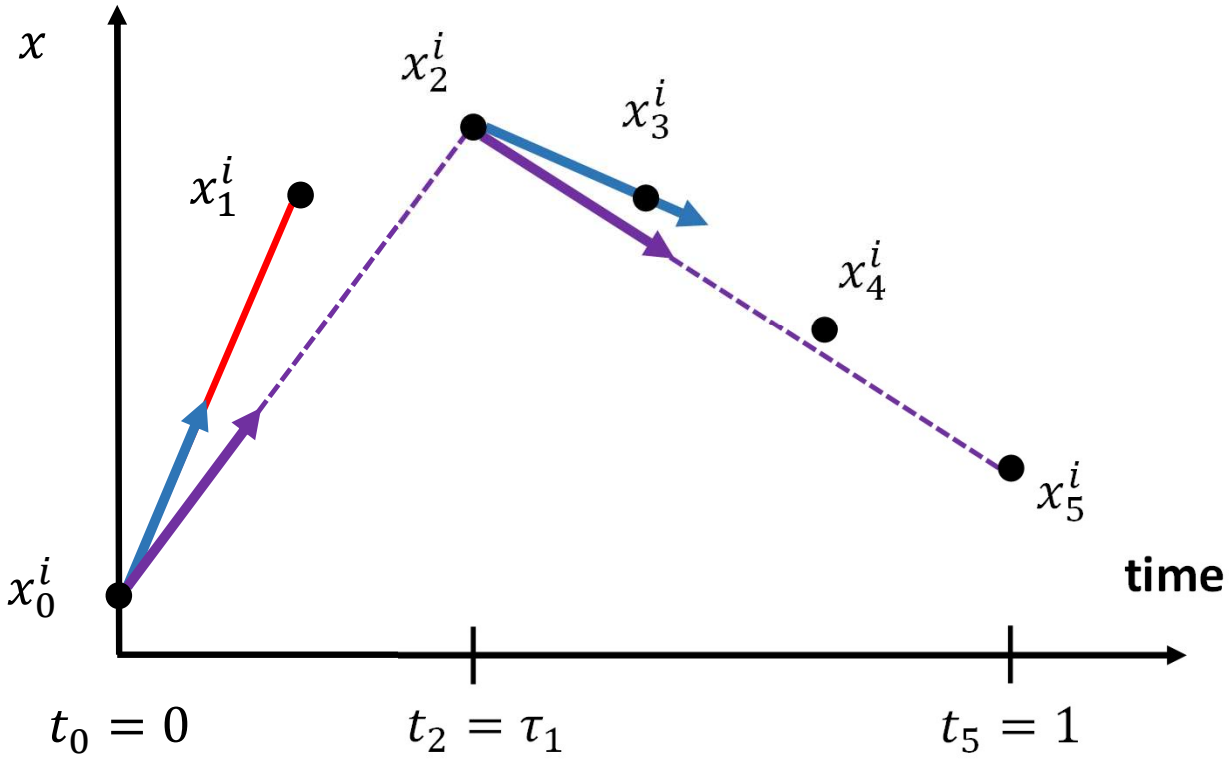


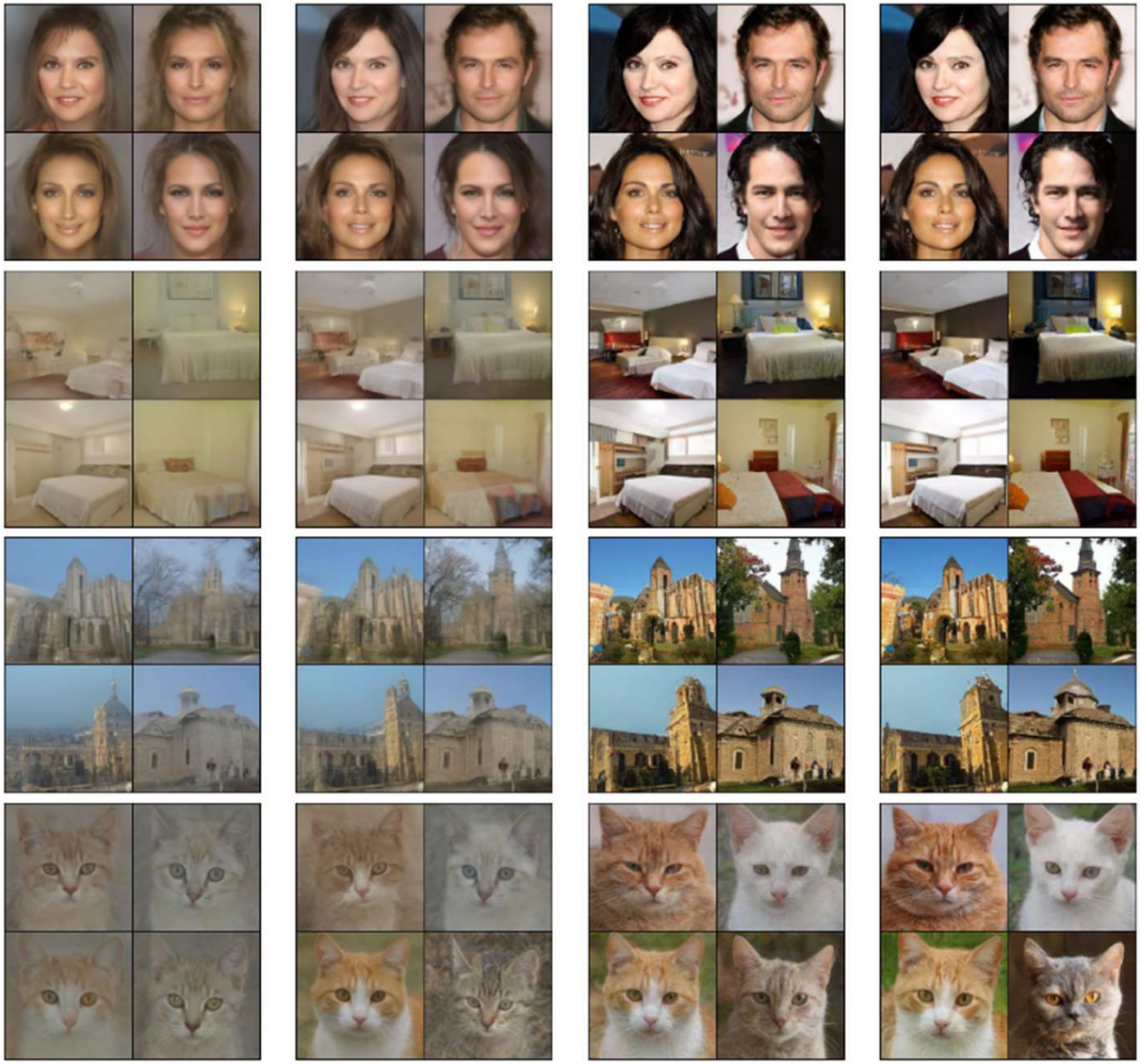
(e) LSUN-Bedroom


$K = 4$
$K = 8$
$K = 15$
Bellman Euler (ours)
Uniform Euler
Bellman Heun (ours)
Uniform Heun
$K = 207$
RK45

Straightening flows with Bellman step-sizes

$$\min_{\theta} \frac{1}{n} \sum_{i=1}^{n} \sum_{k=0}^{K-1} \left\| v_\theta(x^i_{\tau_k}, \tau_k) - \frac{x^i_{\tau_{k+1}} - x^i_{\tau_k}}{\tau_{k+1} - \tau_k} \right\|_2^2 .$$





(a) Euler (6 NFEs) (b) Bellman (6 NFEs) (c) BOSS (6 NFEs) (d) RK45 (208 NFEs)