Finding Local Diffusion Schrödinger Bridge using Kolmogorov-Arnold Network

Supplementary Material

A. Experimental setup

This section presents the experimental setup for DDIM and FM applied to x_θ and $\epsilon_\theta$in the reverse path π^{2n+2} as defined in Sec. 3.1. The LDSB models both the forward and reverse paths as follows:

$$\begin{aligned}
&\pi^{2n+1} : x_t^{2n+1} = \prod_{i=0}^{t} A^{2n+1}(i) x_{0_\theta} + \prod_{i=t}^{T} B^{2n+1}(i) \epsilon_\theta \\
&\pi^{2n+2} : x_t^{2n+2} = \prod_{i=0}^{t} A^{2n+2}(i) x_0 + \prod_{i=t}^{T} B^{2n+2}(i) \epsilon \\
&\forall n : \pi^{2n+1}, \pi^{2n+2} \in \mathcal{P}_{A,B}, t \in \{0, 1, ..., T\}
\end{aligned} \tag{1}$$

where the reverse path is start from P_{prior} (Gaussian noise), the x_{0_θ} and ϵ_θ are predicted by the denoising network D_θ.

For DDIM, the output of the denoising network is predicted Gaussian noise, and x_t is known. Hence, the x_{0_θ} and ϵ_θ is :

$$\begin{aligned}
\epsilon_\theta &= D_\theta(x_t, t) \\
x_{0_\theta} &= (x_t - \prod_{i=t}^{T} B(i) D_\theta(x_t, t)) / \prod_{i=0}^{t} A(i)
\end{aligned} \tag{2}$$

For FM, the denoising network outputs the direction $x_0 - x_T$, where the $x_0 \sim P_{data}$ and $x_T \sim P_{prior}$. The x_t is known. Hence, the x_{0_θ} and ϵ_θ is obtained after reparameterization :

$$\begin{aligned}
\epsilon_\theta &= \frac{x_t - \prod_{i=0}^{t} A(i) D_\theta(x_t, t)}{\prod_{i=0}^{t} A(i) + \prod_{i=t}^{T} B(i)} \\
x_{0_\theta} &= \frac{x_t + \prod_{i=t}^{T} B(i) D_\theta(x_t, t)}{\prod_{i=0}^{t} A(i) + \prod_{i=t}^{T} B(i)}
\end{aligned} \tag{3}$$

Specifically, FM originally sets $f_A = t/T$ and $f_B = 1 - t/T$, as well as the sum of f_A and f_B is 1. However, LDSB does not constrain the sum of f_A and f_B during prediction, which can directly affect the value range of x_{0_θ} and ϵ_θ as denominators. Experiments found that sampling stability can be improved by scalingf_A and f_B isometrically to ensure their sum is 1, as demonstrated in CelebA-HQ in FM.

B. Results of Optimized Diffusion Paths

The specific weight (f_A and f_B) of the diffusion paths optimized by LDSB are provided in Tab. 1 and Tab. 2. The experiments utilize the pre-trained denoising network parameters officially provided by DDIM (CIFAR and CelebA) and Flow Matching (CIFAR and CelebA-HQ).

Table 1. Original diffusion path in DDIM and optimised diffusion path using LDSB

(a) 5-Step path

Time Steps	DDIM (Origin)		CIFAR10 (LDSB)		CelebA (LDSB)	
	f_A	f_B	f_A	f_B	f_A	f_B
0	0.99995	0.01000	1.00000	0.00923	1.00000	0.01142
200	0.81015	0.58622	0.80274	0.58399	0.74435	0.59543
400	0.43997	0.89801	0.43314	0.89487	0.34136	0.85574
600	0.15990	0.98713	0.10833	0.98532	0.10544	0.97506
800	0.03883	0.99925	0.02153	0.99570	0.02064	0.99842

(b) 10-Step path

Time Steps	DDIM (Origin)		CIFAR10 (LDSB)		CelebA (LDSB)	
	f_A	f_B	f_A	f_B	f_A	f_B
0	0.99995	0.01000	1.00000	0.01043	1.00000	0.00983
100	0.94612	0.32382	0.94370	0.32281	0.94515	0.31762
200	0.81015	0.58622	0.80653	0.58421	0.80662	0.57952
300	0.62770	0.77845	0.62379	0.77525	0.62069	0.77264
400	0.43997	0.89801	0.43538	0.89460	0.43651	0.89361
500	0.27892	0.96031	0.27906	0.95708	0.27553	0.95814
600	0.15990	0.98713	0.16213	0.97921	0.15572	0.98646
700	0.08288	0.99656	0.07561	0.98634	0.07938	0.99920
800	0.03883	0.99925	0.02897	0.98634	0.03670	0.99920
900	0.01644	0.99986	0.00924	1.00000	0.01548	1.00000

(c) 20-Step path

Time Steps	DDIM (Origin)		CIFAR10 (LDSB)		CelebA (LDSB)	
	f_A	f_B	f_A	f_B	f_A	f_B
0	0.99995	0.01000	1.00000	0.01018	1.00000	0.00928
50	0.98486	0.17335	0.98477	0.17398	0.98500	0.16057
100	0.94612	0.32382	0.94510	0.32440	0.94755	0.31583
150	0.88650	0.46272	0.88500	0.46257	0.88846	0.45574
200	0.81015	0.58622	0.80802	0.58566	0.81150	0.58104
250	0.72210	0.69179	0.72006	0.69098	0.72364	0.68790
300	0.62770	0.77845	0.62547	0.77755	0.62969	0.77391
350	0.53215	0.84665	0.53080	0.89658	0.53403	0.89456
400	0.43997	0.89801	0.43963	0.89658	0.44200	0.89456
450	0.35474	0.93497	0.35278	0.95804	0.35683	0.95803
500	0.27892	0.96031	0.27576	0.95804	0.28106	0.95803
550	0.21386	0.97686	0.21368	0.97286	0.21615	0.97421
600	0.15990	0.98713	0.16443	0.97447	0.16307	0.98435
650	0.11658	0.99318	0.12495	0.97539	0.12089	0.99406
700	0.08288	0.99656	0.09488	0.97583	0.08692	0.99729
750	0.05745	0.99835	0.07023	0.97625	0.06112	0.99787
800	0.03883	0.99925	0.04959	0.97625	0.04030	0.99787
850	0.02559	0.99967	0.03349	0.97668	0.02365	0.99799
900	0.01644	0.99986	0.02109	0.97761	0.01282	0.99881
950	0.01030	0.99995	0.01228	1.00000	0.00682	1.00000

Table 2. Original diffusion path in Flow Matching and optimised diffusion path using LDSB

(a) 5-Step path

Time Steps	Flow Matching (Origin)		CIFAR10 (LDSB)		CelebA (LDSB)	
	f_A	f_B	f_A	f_B	f_A	f_B
0	1.00000	0.00000	1.00000	0.00000	1.00000	0.00000
200	0.80000	0.20000	0.77570	0.20977	0.80200	0.19800
400	0.60000	0.40000	0.58961	0.40658	0.61247	0.38753
600	0.40000	0.60000	0.40095	0.60201	0.42058	0.57942
800	0.20000	0.80000	0.05246	0.80053	0.08983	0.91017

(b) 10-Step path

Time Steps	Flow Matching (Origin)		CIFAR10 (LDSB)		CelebA (LDSB)	
	f_A	f_B	f_A	f_B	f_A	f_B
0	1.00000	0.00000	1.00000	0.00000	1.00000	0.00000
100	0.90000	0.10000	0.88914	0.10762	0.90011	0.09989
200	0.80000	0.20000	0.78573	0.20539	0.79990	0.20010
300	0.70000	0.30000	0.67939	0.30450	0.69992	0.30008
400	0.60000	0.40000	0.57500	0.40346	0.60003	0.39997
500	0.50000	0.50000	0.47184	0.50230	0.49999	0.50001
600	0.40000	0.60000	0.36996	0.60141	0.40011	0.59989
700	0.30000	0.70000	0.26685	0.70045	0.29987	0.70013
800	0.20000	0.80000	0.15307	0.80057	0.19913	0.80087
900	0.10000	0.90000	0.05306	0.89695	0.04486	0.95514

(c) 20-Step path

Time Steps	Flow Matching (Origin)		CIFAR10 (LDSB)		CelebA (LDSB)	
	f_A	f_B	f_A	f_B	f_A	f_B
0	1.00000	0.00000	1.00000	0.00000	1.00000	0.00000
50	0.95000	0.05000	0.94195	0.06487	0.94874	0.05126
100	0.90000	0.10000	0.87925	0.11170	0.89906	0.10094
150	0.85000	0.15000	0.82802	0.15753	0.84911	0.15089
200	0.80000	0.20000	0.77480	0.21011	0.79919	0.20081
250	0.75000	0.25000	0.71373	0.25986	0.74926	0.25074
300	0.70000	0.30000	0.66295	0.30991	0.69949	0.30051
350	0.65000	0.35000	0.60922	0.36024	0.64945	0.35055
400	0.60000	0.40000	0.55540	0.41058	0.59907	0.40093
450	0.55000	0.45000	0.49805	0.45857	0.54852	0.45148
500	0.50000	0.50000	0.44299	0.50801	0.49825	0.50175
550	0.45000	0.55000	0.38999	0.55719	0.44753	0.55247
600	0.40000	0.60000	0.33685	0.60583	0.39704	0.60296
650	0.35000	0.65000	0.28416	0.65499	0.34660	0.65340
700	0.30000	0.70000	0.23635	0.70432	0.29606	0.70394
750	0.25000	0.75000	0.17714	0.75340	0.24553	0.75447
800	0.20000	0.80000	0.12667	0.80305	0.19562	0.80438
850	0.15000	0.85000	0.08734	0.85257	0.14861	0.85139
900	0.10000	0.90000	0.03767	0.90136	0.09043	0.90957
950	0.05000	0.95000	0.00316	0.95189	0.03957	0.96043

Figure 1. Comparison results of LDSB optimisation over DDIM on CIFAR10 32 × 32. The NFE is 5, 10 and 20.

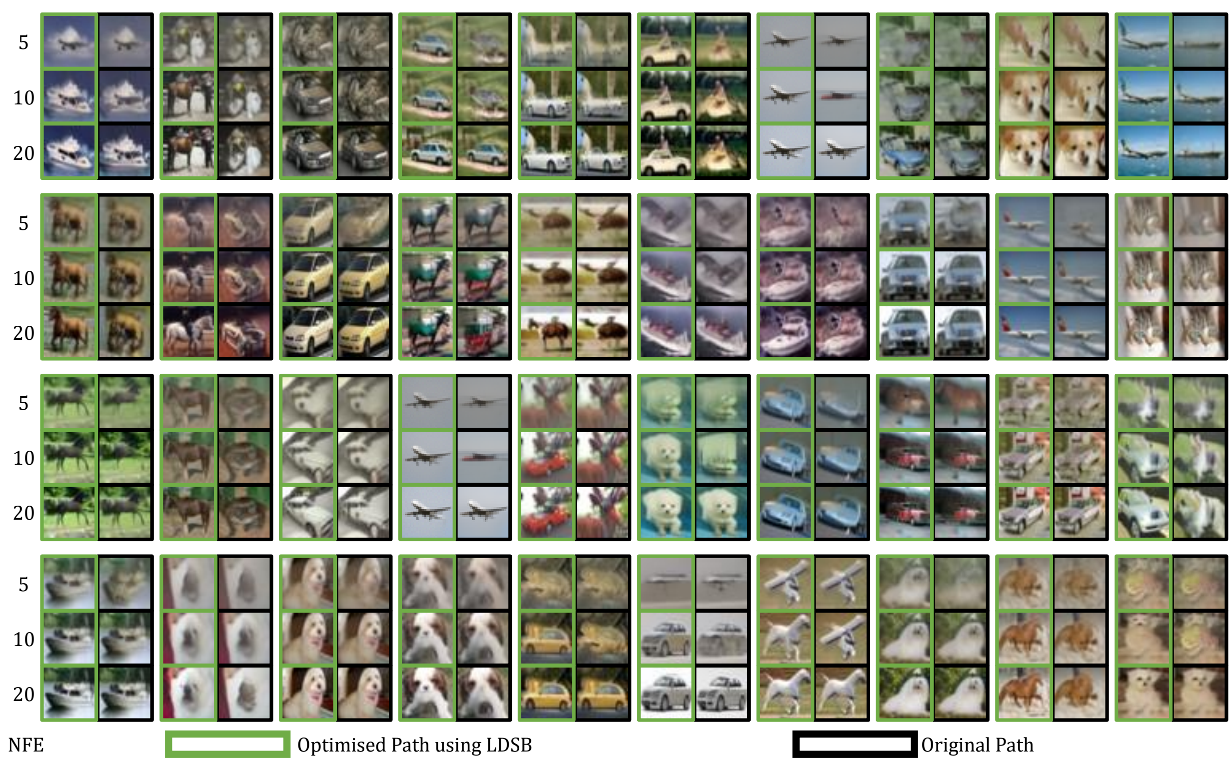

Figure 2. Comparison results of LDSB optimisation over Flow Matching on CIFAR10 32 × 32. The NFE is 5, 10 and 20.

Figure 3. Comparison results of LDSB optimisation over DDIM on CelebA 64×64. The NFE is 5, 10 and 20.

Figure 4. Comparison results of LDSB optimisation over Flow Matching on CelebA-HQ 256 × 256. The NFE is 5, 10 and 20.