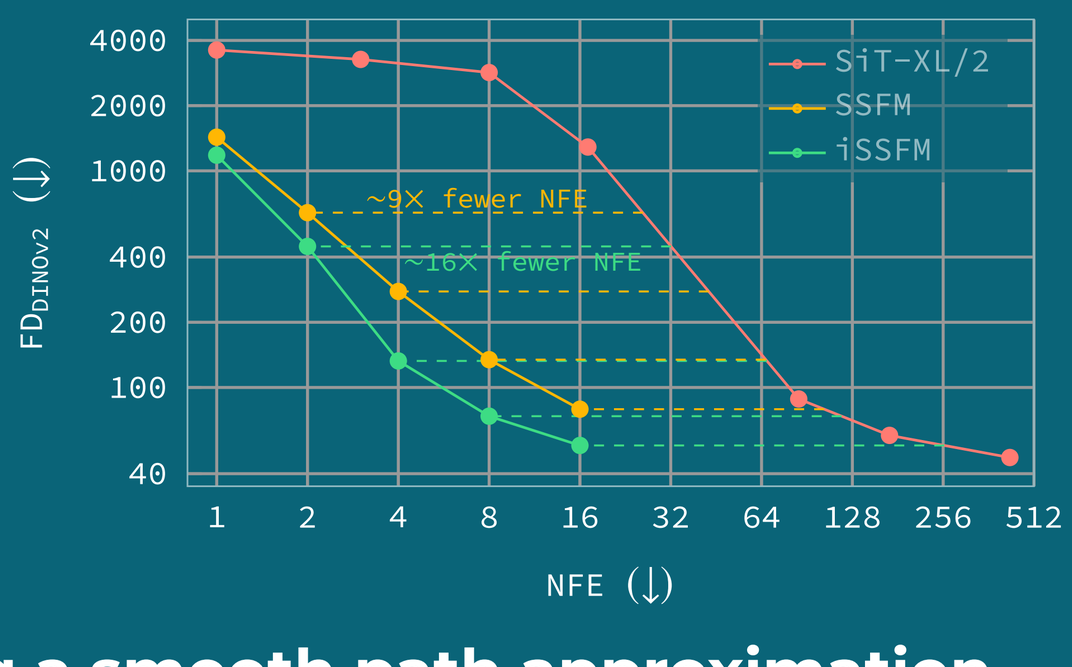
Improved Training of Strong Stochastic Flow Maps

Zander W. Blasingame[1] Antonio Franca[1,2] Niklas Rindtorff[1,2] Alexander Tong[1]

[1]AITHYRA [2]Technische Universität Wien

AITHYRA





Train with JVP-based losses using a smooth path approximation

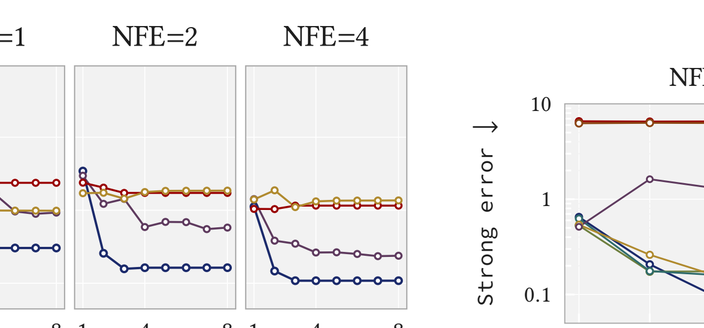
Motivation

A *strong stochastic flow map* (SSFM) (McCallum et al., 2026) models the pathwise solution:

$$\Psi^{\theta}_{s,t}(\boldsymbol{X}_s; \boldsymbol{W}_{[s,t]}(\omega)) \approx \boldsymbol{X}_s(\omega) + \int_s^t \mathrm{d}r \left[2\boldsymbol{u}_r(\boldsymbol{X}_r(\omega)) - \frac{\dot{\alpha}_r}{\alpha_r}\boldsymbol{X}_r(\omega)\right] + \int_s^t \mathrm{d}\boldsymbol{W}_r(\omega)\, \nu_r. \quad (1)$$

The Brownian path $\boldsymbol{W}_{[s,t]}$ is *rough*, making JVP-based losses ill-posed.

Consider a Wong-Zakai type argument where $\boldsymbol{W}^{(N)}_r \to \boldsymbol{W}_r$ as $N \to \infty$



Karhunen-Loève (crimson) vs shifted Legendre (blue) reconstructions.

I.e., we train on the smooth surrogate flow:

$$\Psi^{\theta}_{s,t}(\boldsymbol{X}_s; \boldsymbol{W}^{(N)}_{[s,t]}(\omega)) \approx \boldsymbol{X}_s(\omega) + \int_s^t \mathrm{d}r \left[2\boldsymbol{u}_r(\boldsymbol{X}_r(\omega)) - \frac{\dot{\alpha}_r}{\alpha_r}\boldsymbol{X}_r(\omega) + \nu_r \dot{\boldsymbol{W}}^{(N)}_r(\omega)\right]. \quad (2)$$

We use the following design principles to guide the design of *improved strong stochastic flow maps* (iSSFM):

1. Choose a smooth high-order approximation of the Brownian path which has closed-form composition laws.
2. Choose an objective which ensures that the neural network learns to use the path.
3. Choose a parameterization such that the drift matches the pre-trained flow model.

Toy experiments

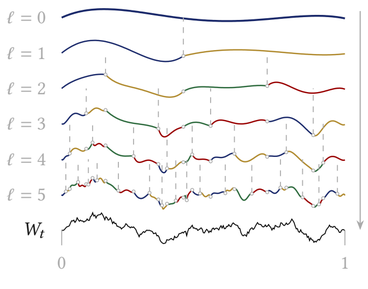
Brownian representations (2D checkerboard).

Training losses (1D SDE).

Algebra of the shifted Legendre polynomials

The shifted Legendre expansion of the Brownian path (Foster, Lyons, and Oberhauser, 2020; McCallum et al., 2026) is given by:

$$\boldsymbol{W}^{(N,s,t)}_{u,v} := \sum_{n=0}^{N-1} \frac{2n+1}{t-s} \boldsymbol{I}^{(n)}_{s,t} \int_u^v \mathrm{d}r\, \widetilde{P}_n\left(\frac{r-s}{t-s}\right), \quad \boldsymbol{I}^{(n)}_{s,t} \sim \mathcal{N}\left(0, \frac{t-s}{2n+1}\boldsymbol{I}\right). \quad (3)$$



Brownian-tree refinement of one path by degree-3 shifted Legendre polynomials composed at arbitrary split points.

Theorem (Composition law for arbitrary splits). For $0 \le n < N$ and any $u \in (s,t)$ with $\lambda = \frac{u-s}{t-s}$:

$$\boldsymbol{I}^{(n)}_{s,t} = \sum_{m=0}^{n} \left[A_{n,m}(\lambda)\, \boldsymbol{I}^{(m)}_{s,u} + B_{n,m}(\lambda)\, \boldsymbol{I}^{(m)}_{u,t}\right],$$

with closed-form coefficients

$$A_{n,m}(\lambda) = (-1)^n(2m+1)\sum_{k=m}^{n}(-\lambda)^k \tfrac{(n+k)!}{(n-k)!(k-m)!(k+m+1)!},$$

$$B_{n,m}(\lambda) = (-1)^m(2m+1)\sum_{k=m}^{n}(-(1-\lambda))^k \tfrac{(n+k)!}{(n-k)!(k-m)!(k+m+1)!},$$

and $\lambda = \frac{1}{2}$ recovers the midpoint law of McCallum et al. (2026).

Strong stochastic flow maps on ImageNet-256

We fine-tune a pre-trained DMF-XL/2+ (Lee, Yu, and Shin, 2026) (distilled from SiT-XL/2+REPA (Yu et al., 2025)) into a strong stochastic flow map.



Rollouts at 4–16 NFE, one Brownian path per row.

iSSFM preserves the *strong* solution property across NFEs:



Strong error w.r.t. 32-step rollout.

Loss	Brownian path	FID (↓)
Semigroup (SSFM)	Shifted Legendre	<u>14.68</u>
Semigroup (KL)	KL + increment	24.92
dESD	Shifted Legendre	15.95
dLSD	Shifted Legendre	23.81
WZ-LMD	Shifted Legendre	214.91
WZ-iMF (iSSFM)	Shifted Legendre	**6.75**

FID at 4 NFE (N=3) for different losses.

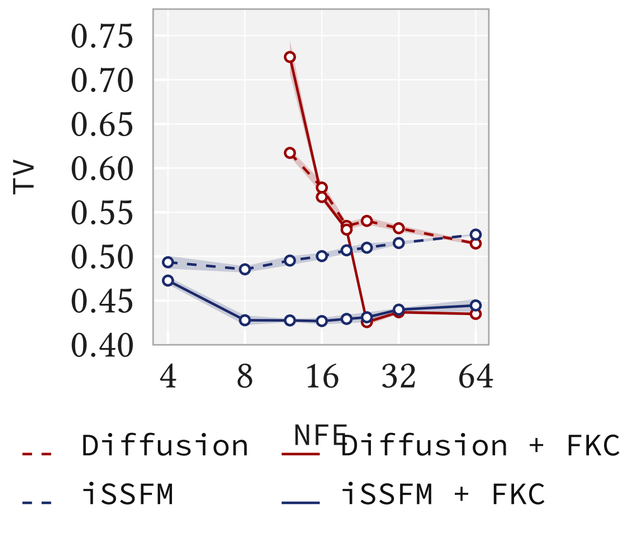
Downstream generative and human-preference evaluations:

Table 1. Head-to-head with SSFM on ImageNet-256 (200k training iterations).

	FID (↓)		$\mathrm{FD}_{\mathrm{DINOv2}}$ (↓)		ImageReward (↑)		PickScore (↑)		HPSv2.1 (↑)	
NFE	SSFM	**iSSFM**	SSFM	**iSSFM**	SSFM	**iSSFM**	SSFM	**iSSFM**	SSFM	**iSSFM**
1	128.44	**115.96**	1430.3	**1181.0**	−1.932	−1.990	**18.65**	18.41	**0.1612**	0.1407
2	41.71	**36.11**	641.3	**447.9**	−1.110	**−1.014**	**19.26**	19.22	**0.1883**	0.1867
4	12.97	**5.36**	277.5	**132.6**	−0.627	**−0.407**	19.77	**20.04**	0.2083	**0.2155**
8	4.22	**2.50**	134.4	**73.7**	−0.351	**−0.220**	20.13	**20.35**	0.2223	**0.2277**
16	2.23	**2.15**	79.4	**54.0**	−0.193	**−0.129**	20.36	**20.50**	0.2313	**0.2343**

Feynman–Kac correctors

FKC (Skreta et al., 2025) sample targets along adjusted marginals (*e.g.*, annealed $\pi_t \propto p_t^{\beta}$). Correction weights $\{w_t\}_{t\in[0,1]}$ evolve **along the particle's own trajectory** (*i.e.*, requiring a **strong** map); we learn the coupled state $(\boldsymbol{X}_t, w_t)$ with an augmented map on $\mathbb{R}^d \oplus \mathbb{R}$.



Total variation vs. NFE.

Key takeaways

1. **Shifted Legendre** + **Wong-Zakai** unlock JVP-based (iMF) losses.
2. iSSFM **outperforms SSFM at every NFE**: ∼2× lower FID at 4 NFE, ∼16× fewer NFE than the teacher.
3. Strong maps enable **few-step Feynman-Kac correctors**, stable at 4 NFE where diffusion samplers diverge.