Generative Fractional Diffusion Models

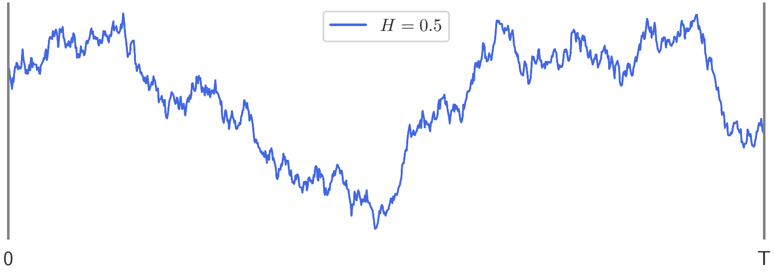
Gabriel Nobis, Maximilian Springenberg, Marco Aversa, Michael Detzel, Rembert Daems, Roderick Murray-Smith, Shinichi Nakajima,

Sebastian Lapuschkin, Stefano Ermon, Tolga Birdal, Manfred Opper, Christoph Knochenhauer, Luis Oala, Wojciech Samek

Brownian Motion (BM)

Brownian motion $B = (B_t)_{t \in [0,T]}$ is a centered Gaussian process with *independent increments*.

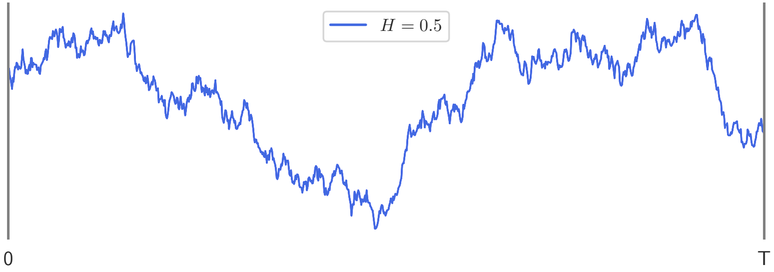
Fractional Brownian Motion (fBM)



Fractional Brownian motion $B^H = (B_t^H)_{t\in[0,T]}$ with Hurst index $H \in (0,1)$ is a centered Gaussian process that posses *correlated increments* where we have for

$H = 0.5$: Brownian motion

Fractional Brownian Motion (fBM)

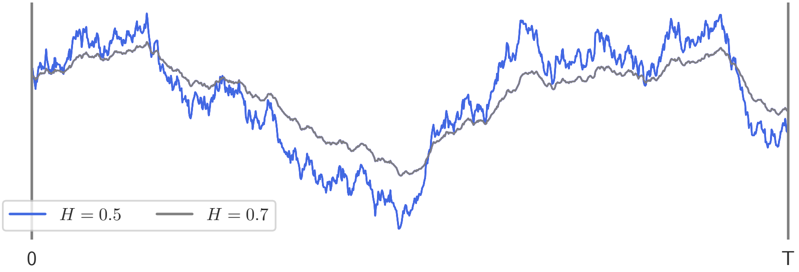


Fractional Brownian motion $B^H = (B_t^H)_{t\in[0,T]}$ with Hurst index $H \in (0,1)$ is a centered Gaussian process that posses *correlated increments* where we have for

$H = 0.5$: Brownian motion

$H > 0.5$: positively correlated increments

Fractional Brownian Motion (fBM)

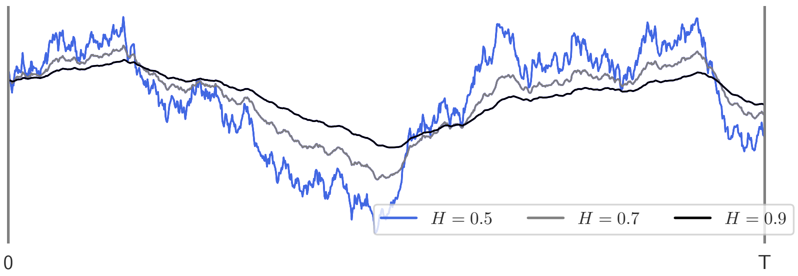


Fractional Brownian motion $B^H = (B_t^H)_{t\in[0,T]}$ with Hurst index $H \in (0,1)$ is a centered Gaussian process that posses *correlated increments* where we have for

$H = 0.5$: Brownian motion

$H > 0.5$: positively correlated increments

Fractional Brownian Motion (fBM)

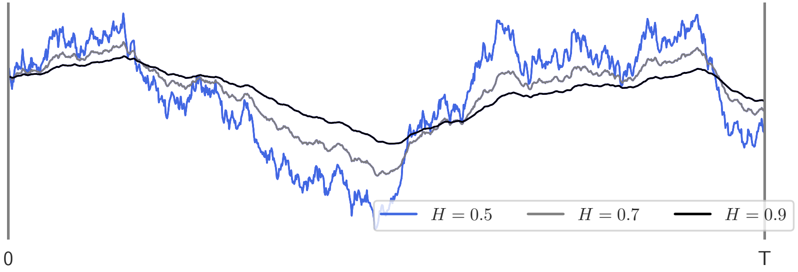


Fractional Brownian motion $B^H = (B_t^H)_{t \in [0,T]}$ with Hurst index $H \in (0, 1)$ is a centered Gaussian process that posses *correlated increments* where we have for

$H = 0.5$: Brownian motion

$H > 0.5$: positively correlated increments

Fractional Brownian Motion (fBM)



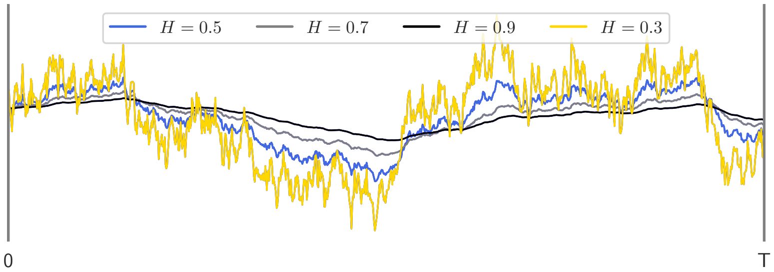
Fractional Brownian motion $B^H = (B_t^H)_{t \in [0,T]}$ with Hurst index $H \in (0, 1)$ is a centered Gaussian process that posses *correlated increments* where we have for

$H = 0.5$: Brownian motion

$H > 0.5$: positively correlated increments

$H < 0.5$: negatively correlated increments

Fractional Brownian Motion (fBM)



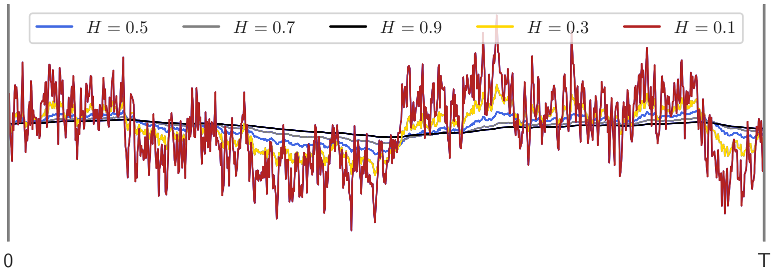
Fractional Brownian motion $B^H = (B_t^H)_{t \in [0,T]}$ with Hurst index $H \in (0, 1)$ is a centered Gaussian process that posses *correlated increments* where we have for

$H = 0.5$: Brownian motion

$H > 0.5$: positively correlated increments

$H < 0.5$: negatively correlated increments

Fractional Brownian Motion (fBM)



Fractional Brownian motion $B^H = (B_t^H)_{t \in [0,T]}$ with Hurst index $H \in (0,1)$ is a centered Gaussian process that posses *correlated increments* where we have for

$H = 0.5$: Brownian motion
$H > 0.5$: positively correlated increments
$H < 0.5$: negatively correlated increments

A Diffusion Model Driven by fBM?

$$X_t = \text{"}X_0 + \int_0^t \mu(u) X_u \mathrm{d}u + \int_0^t g(u) \mathrm{d}B_u^H\text{"}$$

B^H is neither a Markov-process nor a Semimartingale

A Diffusion Model Driven by fBM?

$$X_t = \text{"}X_0 + \int_0^t \mu(u) X_u \mathrm{d}u + \int_0^t g(u) \mathrm{d}B_u^H\text{"}$$

B^H is neither a Markov-process nor a Semimartingale

$\Rightarrow$ No Markov property or Kolmogorov equations (Fokker-Planck) to derive the reverse-time model

Markovian Approximation of fBM (MA-fBM)

Define for every $\gamma \in (0, \infty)$ the Ornstein-Uhlenbeck process Y^γ following

$$\mathrm{d}Y_t^\gamma = -\gamma Y_t^\gamma \mathrm{d}t + \mathrm{d}B_t, \quad Y_0^\gamma = 0.$$

Markovian Approximation of fBM (MA-fBM)

Define for every $\gamma \in (0, \infty)$ the Ornstein-Uhlenbeck process Y^γ following

$$\mathrm{d}Y_t^\gamma = -\gamma Y_t^\gamma \mathrm{d}t + \mathrm{d}B_t, \quad Y_0^\gamma = 0.$$

Given a Hurst index H and a geometrically spaced grid $\gamma_k = r^{k-n}$, define

$$B_t^H = \begin{cases} \int_0^\infty \left(Y_t^\gamma - Y_0^\gamma\right) \nu_1(\gamma)\mathrm{d}\gamma \\ -\int_0^\infty \partial_\gamma \left(Y_t^\gamma - Y_0^\gamma\right) \nu_2(\gamma)\mathrm{d}\gamma \end{cases}$$

Markovian Approximation of fBM (MA-fBM)

Define for every $\gamma \in (0, \infty)$ the Ornstein-Uhlenbeck process Y^γ following

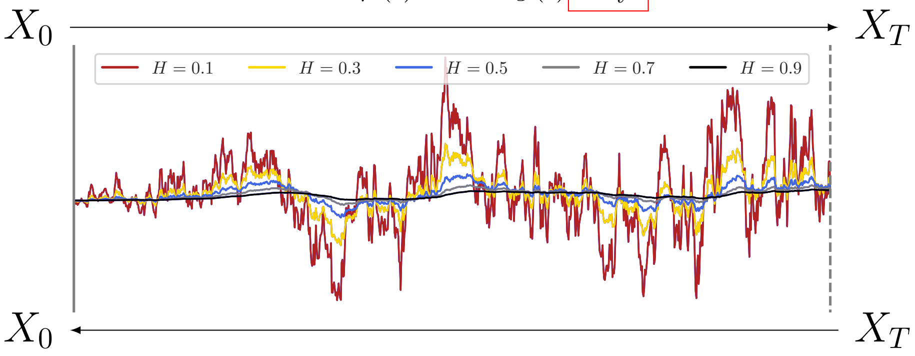
$$\mathrm{d}Y_t^\gamma = -\gamma Y_t^\gamma \mathrm{d}t + \mathrm{d}B_t, \quad Y_0^\gamma = 0.$$

Given a Hurst index H and a geometrically spaced grid $\gamma_k = r^{k-n}$, define

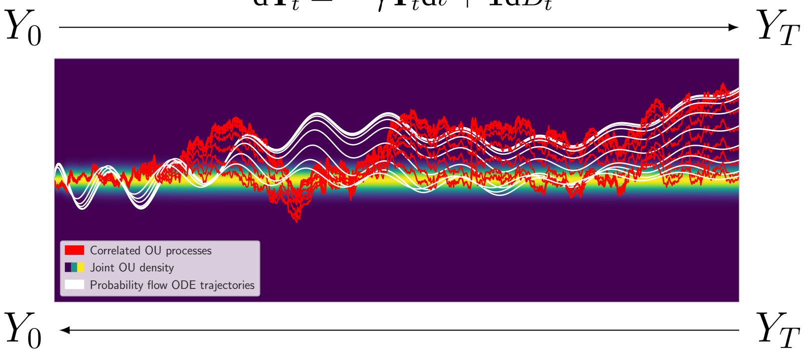
$$B_t^H = \begin{cases} \int_0^\infty \left(Y_t^\gamma - Y_0^\gamma\right) \nu_1(\gamma) \mathrm{d}\gamma \\ -\int_0^\infty \partial_\gamma \left(Y_t^\gamma - Y_0^\gamma\right) \nu_2(\gamma) \mathrm{d}\gamma \end{cases} \approx \boxed{\sum_{k=1}^{K} \omega_k (Y_t^k - Y_0^k) =: \hat{B}_t^H}.$$

Forward Dynamics

$$\mathrm{d}X_t = \mu(t)X_t\mathrm{d}t + g(t)\,\mathrm{d}\hat{B}_t^H$$



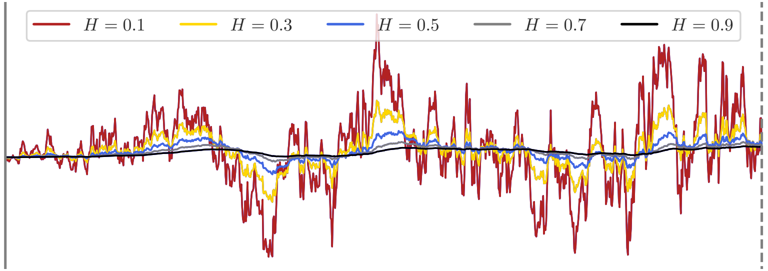
$$\mathrm{d}\mathbf{Y}_t = -\boldsymbol{\gamma}\mathbf{Y}_t\mathrm{d}t + \mathbf{1}\mathrm{d}B_t$$



Forward Dynamics

$$\mathrm{d}X_t = \mu(t)X_t\mathrm{d}t + g(t)\,\mathrm{d}\hat{B}_t^H$$

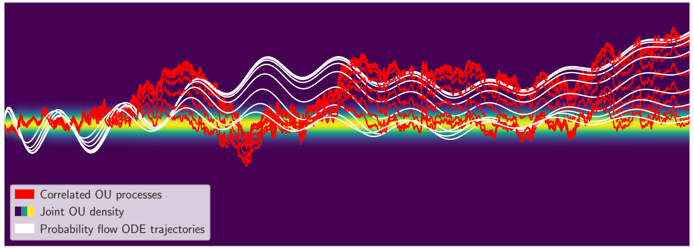
$X_0 \longrightarrow X_T$



$X_0 \longleftarrow X_T$

$$\mathrm{d}\mathbf{Y}_t = -\boldsymbol{\gamma}\mathbf{Y}_t\mathrm{d}t + \mathbf{1}\mathrm{d}B_t$$

$Y_0 \longrightarrow Y_T$

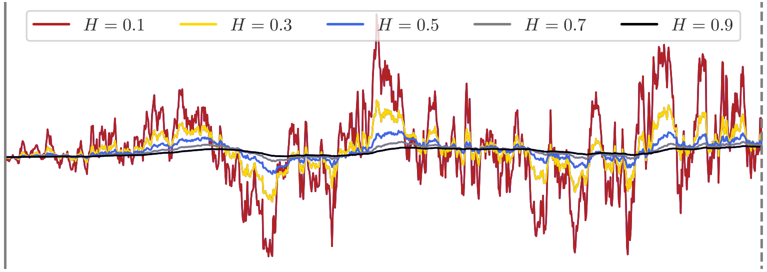


$Y_0 \longleftarrow Y_T$

$$X_t|x_0 = c(t)\left(x_0 + \int_0^t \alpha(t,s)\mathrm{d}B_s\right) \sim \mathcal{N}(c(t)x_0, c^2(t)\sigma^2(t))$$

A Diffusion Model Driven by MA-fBM - The Reverse Process

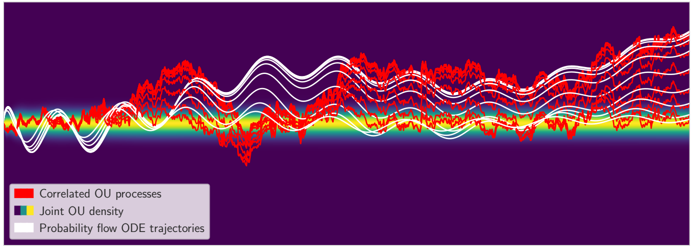
$$X_0 \xrightarrow{\mathrm{d}X_t = \mu(t)X_t\mathrm{d}t + g(t)\mathrm{d}\hat{B}_t^H} X_T$$



$$X_0 \xleftarrow{} X_T$$

$$\mathrm{d}\overline{\mathbf{Z}}_t = \left[\overline{\mathbf{F}}(t)\overline{\mathbf{Z}}_t - \overline{\mathbf{G}}(t)\overline{\mathbf{G}}(t)^T \nabla_{\mathbf{z}} \log \overline{p}_t(\overline{\mathbf{Z}}_t)\right] \mathrm{d}t + \overline{\mathbf{G}}(t)\mathrm{d}\overline{B}_t$$

$$Y_0 \xrightarrow{\mathrm{d}\mathbf{Y}_t = -\boldsymbol{\gamma}\mathbf{Y}_t\mathrm{d}t + \mathbf{1}\mathrm{d}B_t} Y_T$$



$$Y_0 \xleftarrow{} Y_T$$

$$\mathrm{d}\overline{\mathbf{Y}}_t = [-\boldsymbol{\gamma}\overline{\mathbf{Y}}_t - \mathbf{1}_{K,K}\nabla_{\mathbf{y}} \log \overline{q}_t(\overline{\mathbf{Y}}_t)]\mathrm{d}t + \mathbf{1}\mathrm{d}\overline{B}_t$$

$$\nabla_{\mathbf{z}} \log p_t(\mathbf{Z}_t) = \ldots?$$

Optimal Score Model

We propose *augmented score matching* to train the score model $s_{\boldsymbol{\theta}}$:

$$\mathcal{L}(\boldsymbol{\theta}) := \mathbb{E}_t \left\{ \mathbb{E}_{(\mathbf{X}_0, \mathbf{Y}_t^{[K]})} \mathbb{E}_{(\mathbf{X}_t | \mathbf{Y}_t^{[K]}, \mathbf{X}_0)} \left[\| s_{\boldsymbol{\theta}}(\mathbf{X}_t - \sum_k \eta_t^k \mathbf{Y}_t^k, t) - \nabla_{\mathbf{x}} \log p_{0t}(\mathbf{X}_t | \mathbf{Y}_t^{[K]}, \mathbf{X}_0) \|_2^2 \right] \right\}.$$

Optimal Score Model

We propose *augmented score matching* to train the score model $s_{\boldsymbol{\theta}}$:

$$\mathcal{L}(\boldsymbol{\theta}) := \mathbb{E}_t \left\{ \mathbb{E}_{(\mathbf{X}_0, \mathbf{Y}_t^{[K]})} \mathbb{E}_{(\mathbf{X}_t | \mathbf{Y}_t^{[K]}, \mathbf{X}_0)} \left[\| s_{\boldsymbol{\theta}}(\mathbf{X}_t - \sum_k \eta_t^k \mathbf{Y}_t^k, t) - \nabla_{\mathbf{x}} \log p_{0t}(\mathbf{X}_t | \mathbf{Y}_t^{[K]}, \mathbf{X}_0) \|_2^2 \right] \right\}.$$

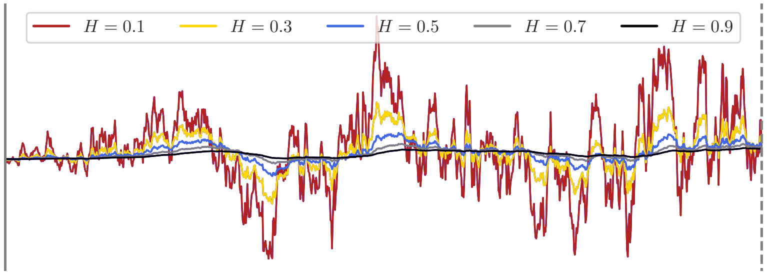
Assume that $s_{\boldsymbol{\theta}}$ is optimal w.r.t. the augmented score matching loss $\mathcal{L}$ in (12). The score model

$$S_\theta(\mathbf{Z}_t, t) := \left(s_{\boldsymbol{\theta}}(\mathbf{X}_t - \sum_k \eta_t^k \mathbf{Y}_t^k, t), -\eta_t^1 s_{\boldsymbol{\theta}}(\mathbf{X}_t - \sum_k \eta_t^k \mathbf{Y}_t^k, t), \ldots, -\eta_t^K s_{\boldsymbol{\theta}}(\mathbf{X}_t - \sum_k \eta \mathbf{Y}_t^k, t) \right)$$

yields the optimal $L^2(\mathbb{P})$ approximation of $\nabla_{\mathbf{z}} \log p_t(\mathbf{Z}_t)$.

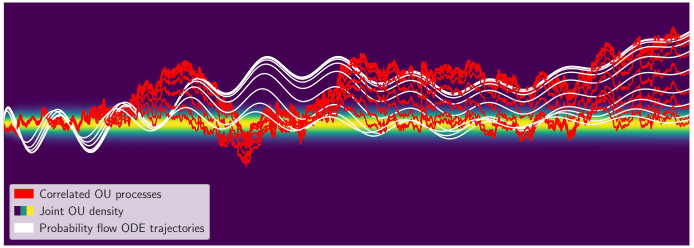
A Score-based Model Driven by MA-fBM - The Reverse Process

$$X_0 \xrightarrow{\; dX_t = \mu(t) X_t dt + g(t) d\hat{B}_t^H \;} X_T$$



$$X_0 \xleftarrow{\; d\overline{\mathbf{Z}}_t = \left[\overline{\mathbf{F}}(t)\overline{\mathbf{Z}}_t - \overline{\mathbf{G}}(t)\overline{\mathbf{G}}(t)^T \nabla_{\mathbf{z}} \log \overline{p}_t(\overline{\mathbf{Z}}_t)\right] dt + \overline{\mathbf{G}}(t) d\overline{B}_t \;} X_T$$

$$Y_0 \xrightarrow{\; d\mathbf{Y}_t = -\boldsymbol{\gamma}\mathbf{Y}_t dt + \mathbf{1} dB_t \;} Y_T$$



$$Y_0 \xleftarrow{\; d\overline{\mathbf{Y}}_t = [-\boldsymbol{\gamma}\overline{\mathbf{Y}}_t - \mathbf{1}_{K,K} \nabla_{\mathbf{y}} \log \overline{q}_t(\overline{\mathbf{Y}}_t)] dt + \mathbf{1} d\overline{B}_t \;} Y_T$$

Known guiding score function

$$\nabla_{\mathbf{z}} \log p_t(\mathbf{Z}_t) \approx S_{\boldsymbol{\theta}}(\mathbf{Z}_t, t) + \nabla_{\mathbf{z}} \log q_t(\mathbf{Y}_t^{[K]}), \quad \mathbf{Y}_t^{[K]} := (Y_t^1, \ldots, Y_t^K)$$

Gabriel Nobis

Fraunhofer HHI

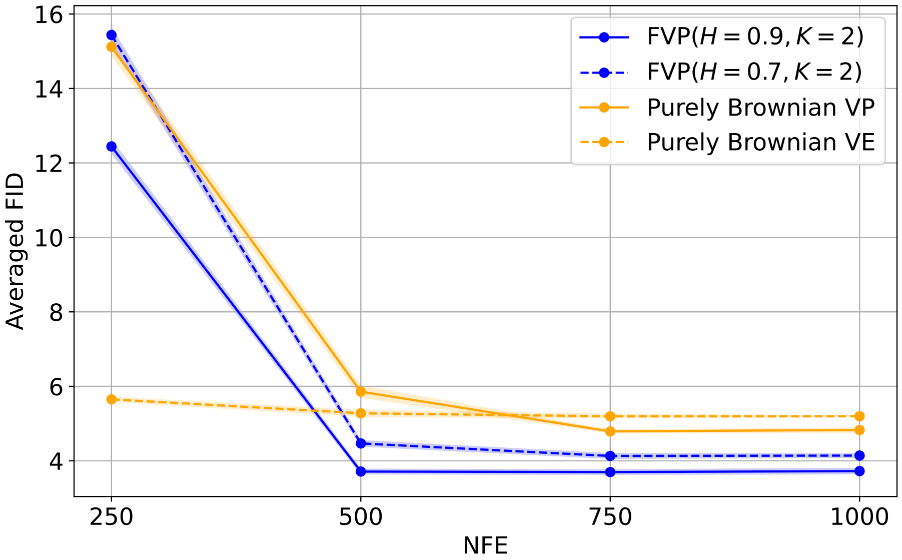
Results: Different Hurst Indices

MNIST	FID $\downarrow$	$\text{VS}_p \uparrow$
BM driven		
VE (retrained	10.82	24.20
VP (retrained)	1.44	23.64
MA-fBM driven		
$\text{FVP}(H = 0.9, K = 3)$	**0.72**	24.18
$\text{FVP}(H = 0.7, K = 3)$	0.86	24.39
$\text{FVP}(H = 0.9, K = 4)$	1.22	**24.76**

CIFAR10	FID $\downarrow$	$\text{VS}_p \uparrow$
BM driven		
VE (retrained)	5.20	3.42
VP (retrained)	4.85	3.28
MA-fBM driven		
$\text{FVP}(H = 0.9, K = 1)$	4.79	3.53
$\text{FVP}(H = 0.7, K = 2)$	4.17	3.35
$\text{FVP}(H = 0.9, K = 2)$	**3.77**	**3.60**

Conditional image generation on (LHS) MNIST and (RHS) CIFAR10.

Results: Number of Function Evaluations (NFEs)



Averaged FID on CIFAR10 over three rounds of sampling plotted across different NFEs.

Results: Class-wise performance

Metric	Dynamics	airplane	automobile	bird	cat	deer	dog	frog	horse	ship	truck
FID ↓	VP	15.29	12.06	14.08	18.08	10.68	16.92	16.48	12.49	10.74	10.57
	FVP$(H = 0.7, K = 2)$	14.67	9.55	**14.02**	16.97	11.05	17.14	16.43	10.97	**9.91**	8.81
	FVP$(H = 0.9, K = 2)$	**14.37**	**8.94**	14.18	**16.38**	**10.52**	**16.76**	**15.37**	**10.28**	10.04	**8.76**
Recall ↑	VP	0.6814	0.6186	0.6860	0.6466	0.7002	0.6730	0.6758	0.6392	0.6468	0.5982
	FVP$(H = 0.7K = 2)$	0.6838	0.6436	0.6870	0.6712	0.7140	0.6844	0.6922	0.6764	0.6550	0.6508
	FVP$(H = 0.9, K = 2)$	**0.7038**	**0.6614**	**0.7188**	**0.6842**	**0.7284**	**0.7096**	**0.7104**	**0.6806**	**0.6772**	**0.6852**

Class-wise image quality and class-wise distribution coverage on CIFAR10.

Conclusion

- We propose to use MA-fBM as the driving noise process for your diffusion models
- A score model that matches the dimensionality of the data suffices
- We achieve higher image quality, improved pixel-wise diversity and better distribution coverage

PAPER

Gabriel Nobis

gabriel__nobis

CODE