Spontaneous Symmetry Breaking in Generative Diffusion Models

Gabriel Raya & Luca Ambrogioni

Accepted at NeurIPS 2023

[1]Song et al. arxiv:2011.13456 (2021)
[2]Ho et al. arxiv:2210.02303 (2022)
[3] Anand et al. arxiv:2205.15019 (2022)
[4] Zeng et al. arxiv:2210.06978 (2022)

Generative diffusion models

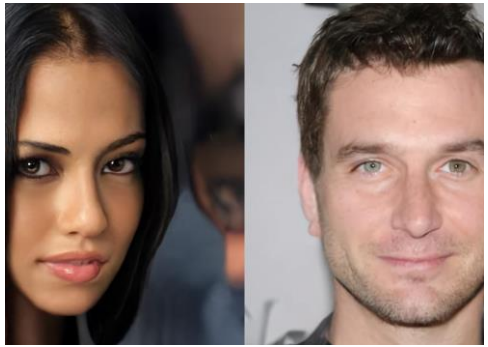
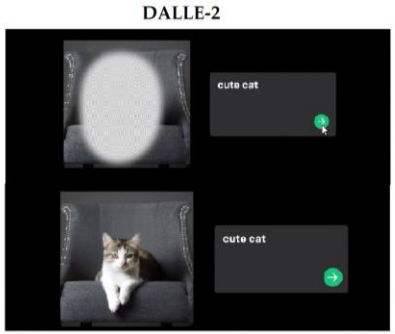
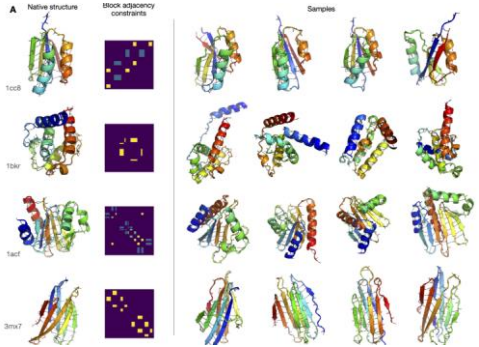


Image generation[1]

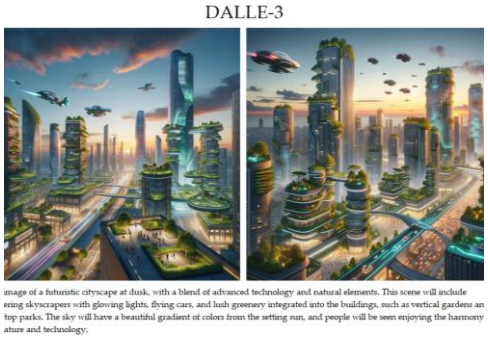


Inpainting

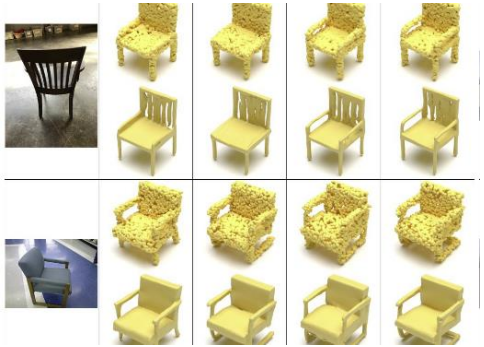


Protein structure generation[3]

Video Generation[2]

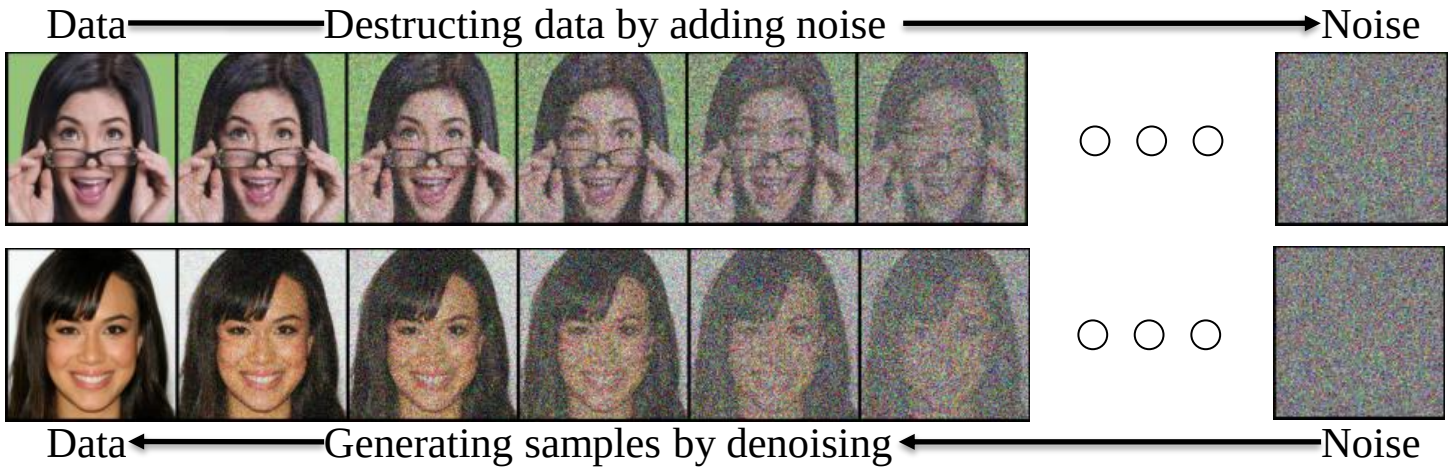


Conditional generation (text-image)



Point Cloud generation[4]

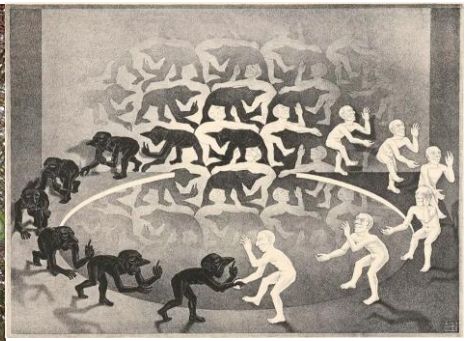
Generating samples by denoising



Generation as denoising → Generation as spontaneous symmetry breaking

Symmetry shapes our understanding of the laws of nature

Symmetries plays a central role in the study of any physical system → conservation laws



$$\frac{d}{dt}p_x = -\frac{\partial H}{\partial x} = 0$$

Translational symmetry x-direction

Symmetries are everywhere!

Taken from : https://www.dpreview.com/challenges/Entry.aspx?ID=1166560 | https://www.ccdc.cam.ac.uk/ | https://www.nationalgeographic.es/animales/mariposa-monarca

Symmetry breaking → pattern formation in nature

The emergence of new particles and excitations, and the rigidity of collective states of matter.

Lower symmetry

Higher symmetry

Fixed orientation in space ← Rotational, translational symmetry

Order state ← Disorder state

"The secret of nature is symmetry, but much of the texture of the world is due to mechanisms of symmetry breaking." Gross (1996).

Spontaneous symmetry breaking – Mexican hat potential

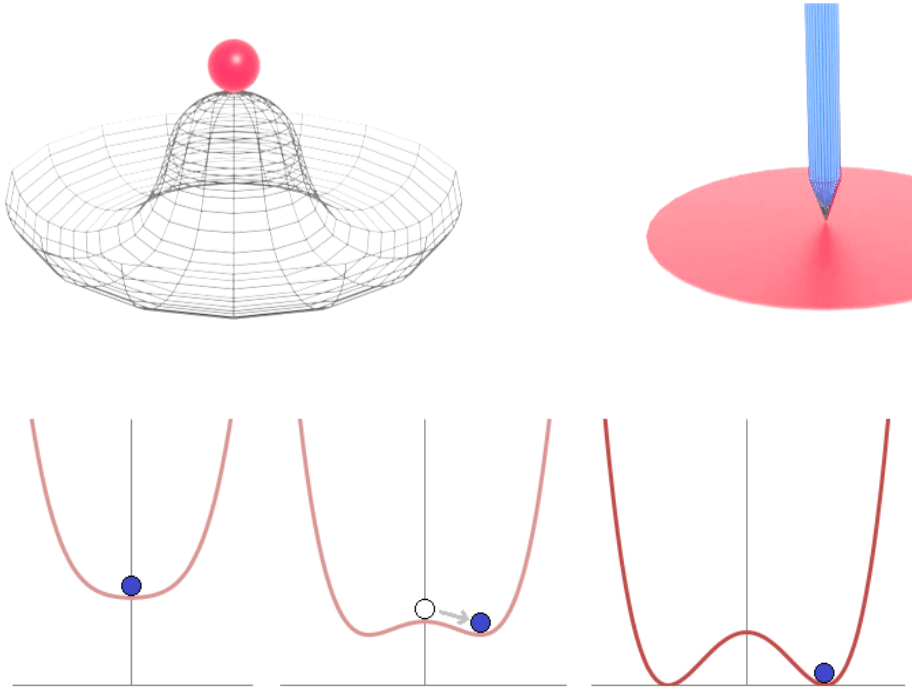
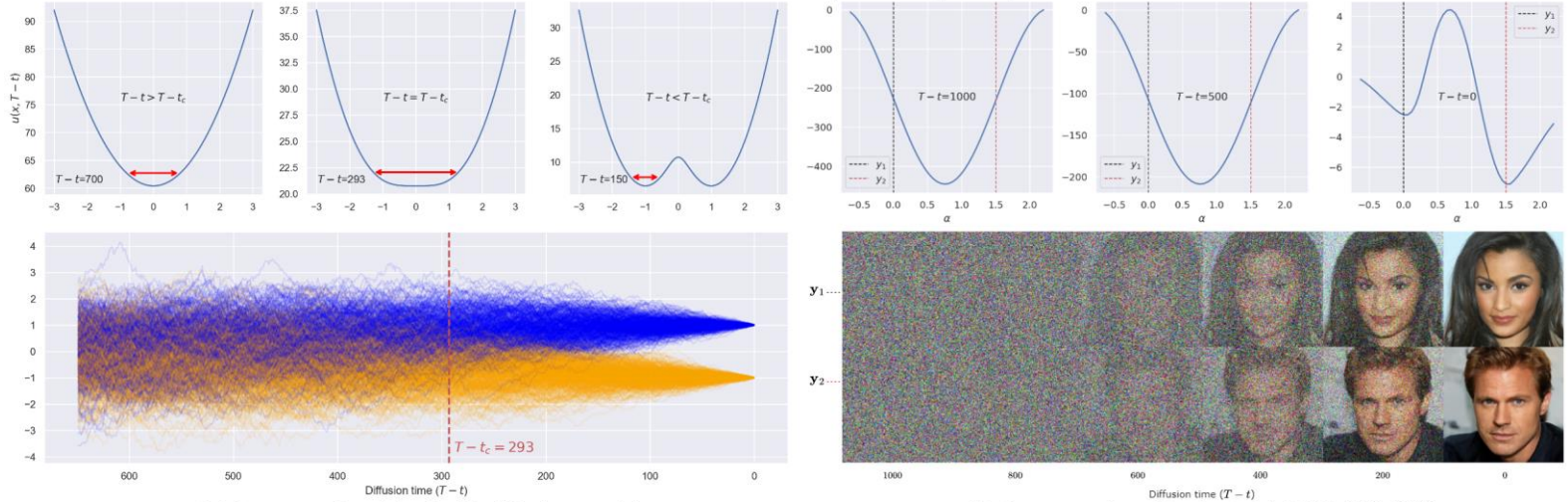


Image : Wikimedia Commons

Diffusion models undergoes spontaneous symmetry breaking

$$d\mathbf{X}_t = -\nabla_{\mathbf{x}} u(\mathbf{X}_t, T-t)\mathrm{d}t + g(T-t)\mathrm{d}\mathbf{W}_t$$



(a) Symmetry breaking in 1D diffusion model

(b) Symmetry breaking in CelebA HQ 256x256

Generative capabilities of diffusion models are the result of a **phase transition**

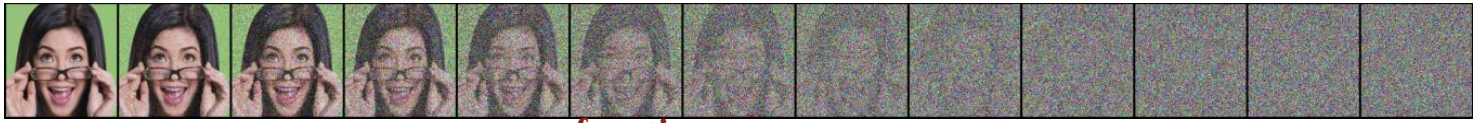
Diversity comes from this window of instability (criticality)

Continuous diffusion models

Expressing the generative backward equation in terms of a potential energy function

Forward SDE (data → noise)

$$\boldsymbol{Y}_0 \longrightarrow d\boldsymbol{Y}_s = f(\boldsymbol{Y}_s, s)dt + g(s)d\boldsymbol{W}_s \longrightarrow \boldsymbol{Y}_T$$



score function

$$\boldsymbol{X}_0 \longleftarrow d\boldsymbol{X}_t = [g^2(T-t)\nabla_{\boldsymbol{x}} \log p(\boldsymbol{X}_t, T-t) - f\,(\boldsymbol{X}_t, T-t)]dt + g(T-t)d\boldsymbol{W}_t \longleftarrow \boldsymbol{X}_T$$

Reverse SDE (data ← noise)

$-\nabla_x u(X_t, T-t)$ ← Potential Energy Function

Transition kernel

$k(\boldsymbol{y}, s; \boldsymbol{y}_0, 0) = \mathcal{N}(\boldsymbol{y}; \theta_s y_0, (1-\theta_s^2)I)$

$\theta_s = e^{-\frac{1}{2}\int_0^s \beta(\tau)d\tau}$

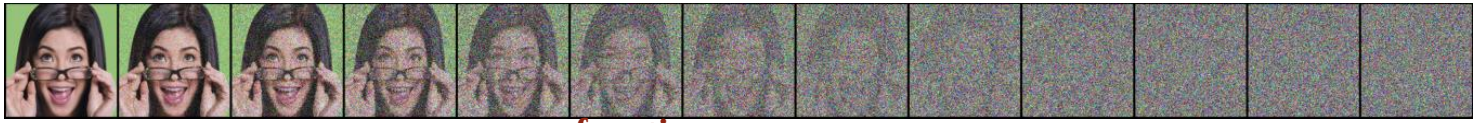
$$u(\boldsymbol{X}_t, T-t) = -g^2(T-t)\log p(\boldsymbol{X}_t, T-t) + \int_0^x f\,(z, T-t)dz$$

OU Process (VP-SDE)

Expressing the generative backward equation in terms of a potential energy function

Forward SDE (data → noise)

$$\boldsymbol{Y}_0 \longrightarrow d\boldsymbol{Y}_s = -\frac{1}{2}\beta(s)\boldsymbol{Y}_s dt + \sqrt{\beta(s)}d\boldsymbol{W}_s \longrightarrow \boldsymbol{Y}_T$$



score function

$$\boldsymbol{X}_0 \longleftarrow d\boldsymbol{X}_t = \left[\beta(T-t)\nabla_{\boldsymbol{x}} \log p(\boldsymbol{X}_t, T-t) + \frac{1}{2}\beta(T-t)\boldsymbol{X}_t\right] dt + \sqrt{\beta(T-t)}d\boldsymbol{W}_t \longleftarrow \boldsymbol{X}_T$$

Reverse SDE (data ← noise)

$-\nabla_x u(X_t, T-t)$ ← Potential Energy Function

$$u(\boldsymbol{X}_t, T-t) = -\beta(T-t)\log p(\boldsymbol{X}_t, T-t) + -\frac{1}{4}\beta(T-t)\boldsymbol{X}_t^2$$

Theoretical analysis

Spontaneous symmetry breaking in a one-dimensional diffusion model

Spontaneous symmetry breaking in one-dimensional diffusion models

Let's consider a dataset consisting of two points {-1, 1} sampled with equal probability.

- The distribution at time $s = 0$ is : $p(y, 0) = \frac{1}{2}\,\delta(x - a) + \frac{1}{2}\,\delta(x - b)$
- The symmetry group that preserves the potential : identity and $g(x) = -x$

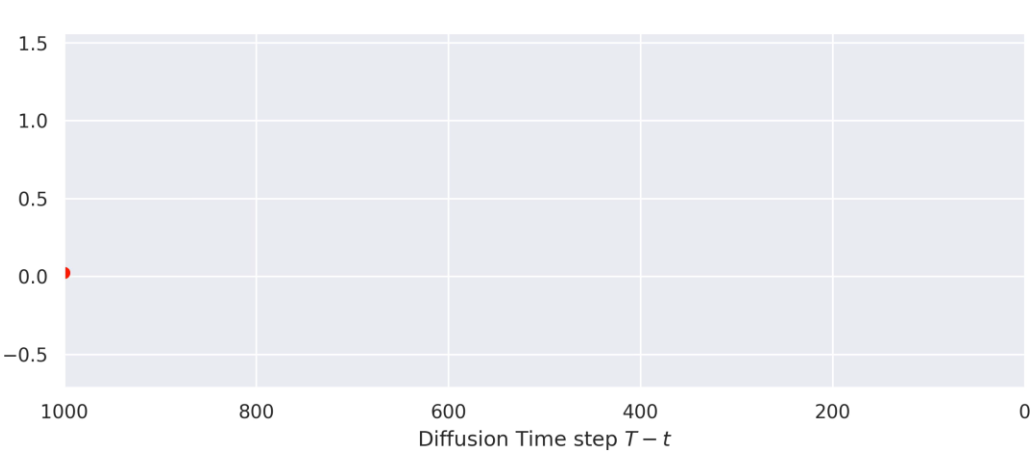
Potential Energy Function

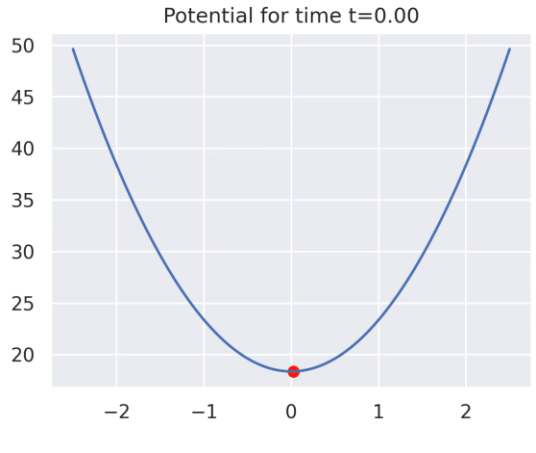
$$u(\boldsymbol{X}_t, T - t) = -\beta(T - t)\log p(\boldsymbol{X}_t, T - t) - \frac{1}{4}\beta(T - t)\boldsymbol{X}_t^2$$

$$u(\boldsymbol{X}_t, T - t) = -\beta(T - t)\log\left(\frac{1}{2\sqrt{2\pi(1 - \theta_s^2)}}\left(e^{-\frac{(x-\theta_s)^2}{2(1-\theta_2)^2}} + e^{-\frac{(x+\theta_s)^2}{2(1-\theta_2)^2}}\right)\right) - \frac{1}{4}\beta(T - t)X_t^2$$

Generative dynamics and the energy landscape

$$u(\boldsymbol{X}_t, T-t) = -\beta(T-t)\log\left(\frac{1}{2\sqrt{2\pi(1-\theta_s^2)}}\left(e^{-\frac{(x-\theta_s)^2}{2(1-\theta_2)^2}} + e^{-\frac{(x+\theta_s)^2}{2(1-\theta_2)^2}}\right)\right) - \frac{1}{4}\beta(T-t)X_t^2$$





The evolution of a generated particle through time and across the energy landscape.

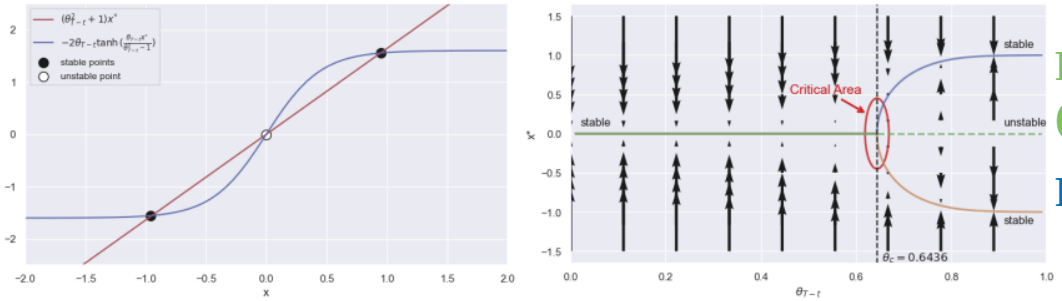
Phase 1: Equal potentiality (mean reverting)

Phase 2: Denoising phase

Stability analysis – Stable equilibrium states

All fixed points can be found by solving the self-consistency equation:

$$(\theta_{T-t} + 1)x^* = -2\theta_{T-t} \tanh\left(\frac{\theta_{T-t}x^*}{\theta_{T-t} - 1}\right)$$



Phase 1: Equal potentiality (mean reverting)

Phase 2: Denoising phase

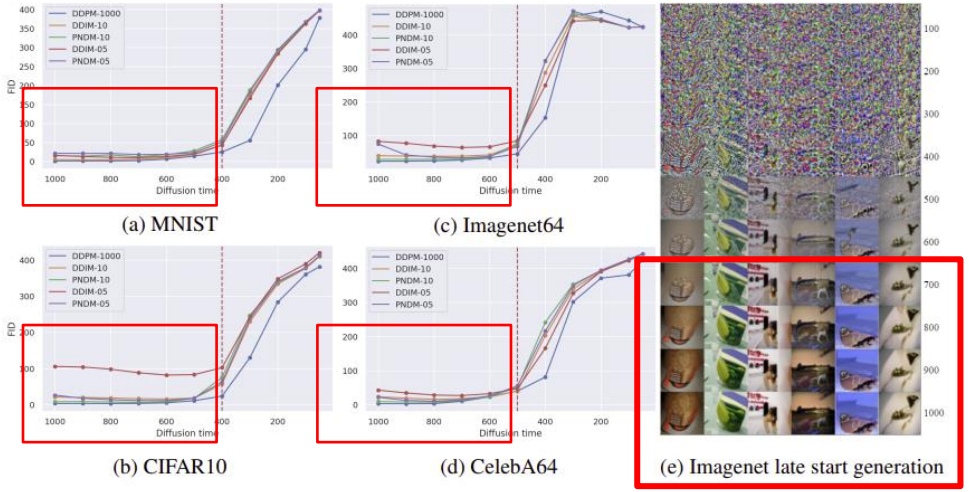
Figure 3: Bifurcation analysis of the generative dynamics of a one-dimensional diffusion model. (a) Geometric visualization of bifurcation of fixed points through the intersection of a straight line and a hyperbolic tangent at a value $\theta_{T-t} > \theta_c$. (b) Bifurcation diagram obtained by numerically solving the self-consistency equation Eq. 10, demonstrating the bifurcation at the critical value θ_c. The blue, orange and green lines denote the three paths of fixed-points. The vector field is given by the drift term (i.e. the gradient of the potential) in the generative SDE.

Each of the branched stable paths only preserve a sub-group of the overall symmetry

Empirical analysis

Spontaneous symmetry breaking in trained diffusion models

Analyzing the impact of a late start initialization on FID performance



(a) MNIST (c) Imagenet64

(b) CIFAR10 (d) CelebA64 (e) Imagenet late start generation

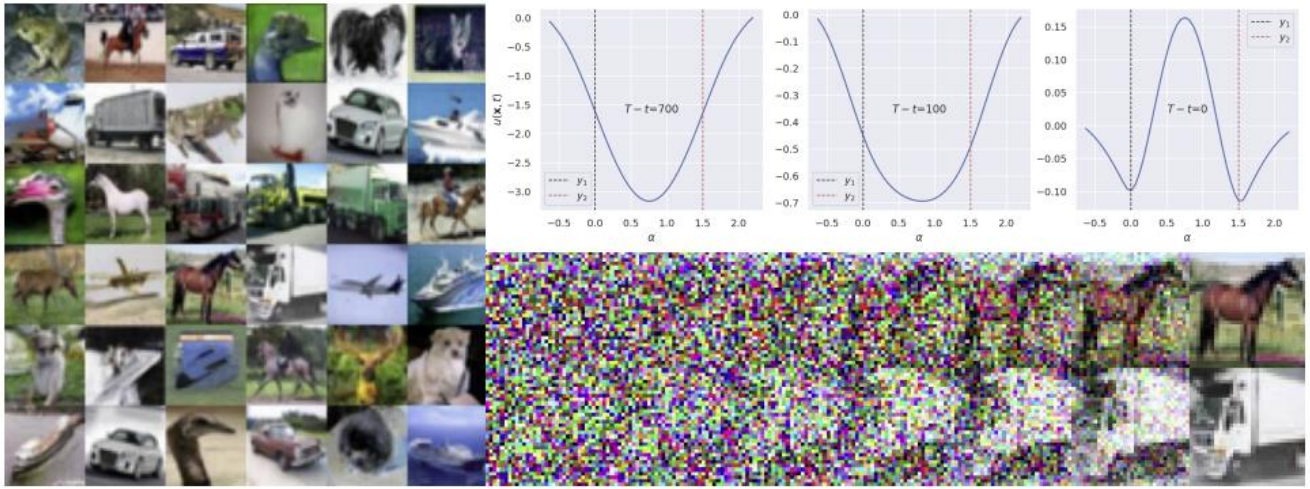
Figure 4: Analysis of the model's performance, as measured by FID scores, for different starting times using three different sampling methods: the normal DDPM sampler with decreasing time steps from $T = 1000$ to 0, and fast sampler DDIM and PSDM for 10 and 5 denoising steps. The vertical line corresponds to the maximum of the second derivative of the FID curve, which offers a rough estimate of the first bifurcation time. (e) Illustrates samples generation on Imagenet64, while progressively varying the starting time from 1000 to 100.

Exploratory analysis of FID scores for different diffusion time starts

There is an early phase that remains unaffected!

There is a sharp degradation in performance after a critical time, accompanied by an apparent discontinuity in the second derivative of the FID curve.

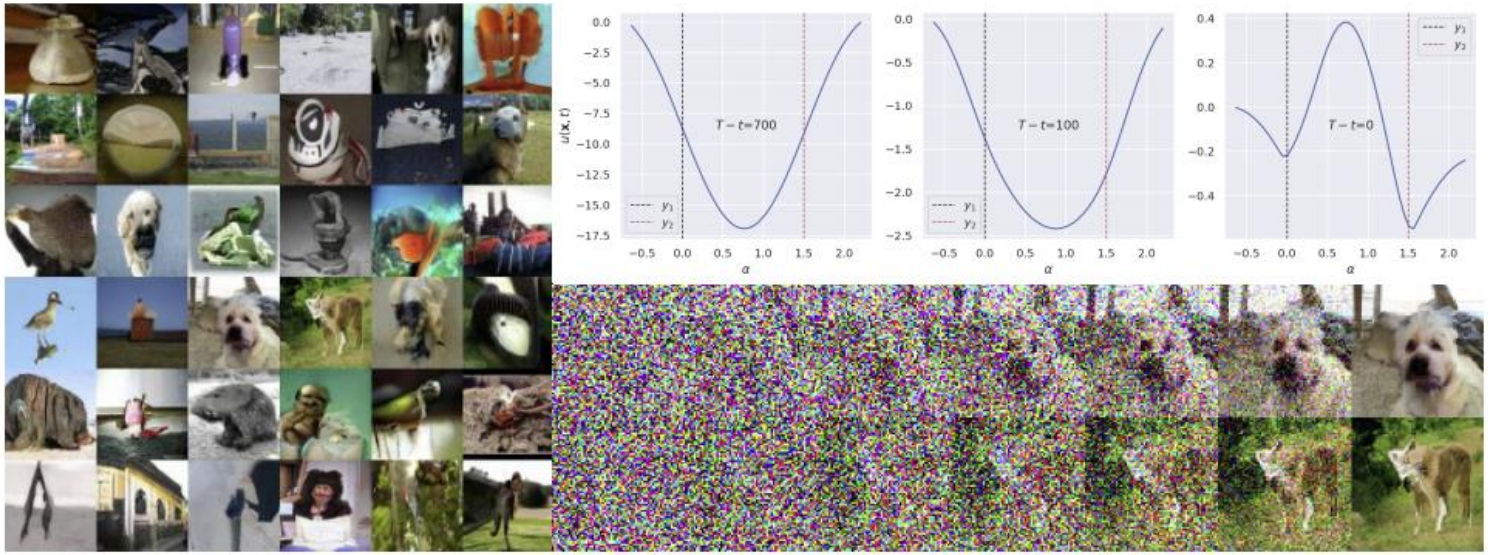
Empirical analysis of the potential function in trained diffusion models



(a) (b)

Figure 15: Symmetry breaking in CIFAR10: (a) Generated samples; (b) Time-varied 1D potential sections (top figure) from a trained diffusion model along circular paths between two samples (bottom figure), averaged over 20 generated samples from (a).

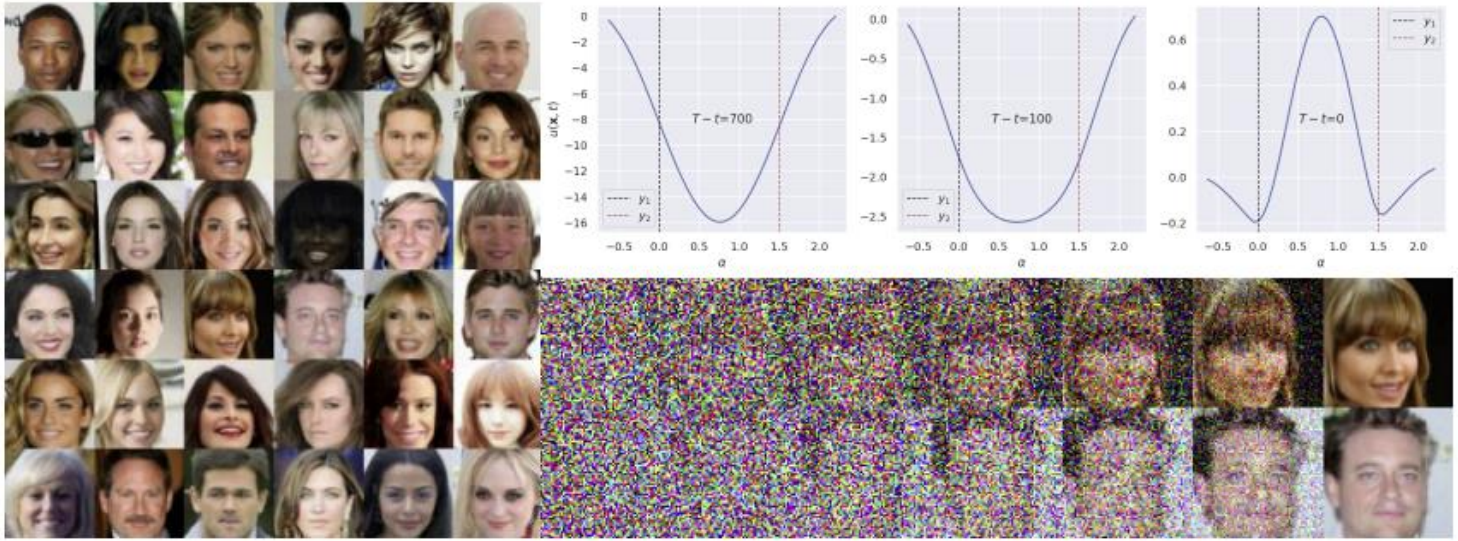
Empirical analysis of the potential function in trained diffusion models



(a) (b)

Figure 16: Symmetry breaking in Imagenet64.

Empirical analysis of the potential function in trained diffusion models



(a) (b)

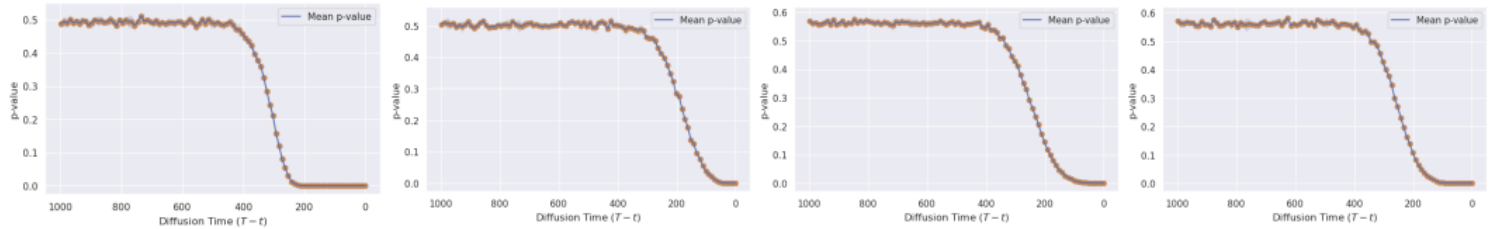
Figure 17: Symmetry breaking in CelebA64.

Improving the generative performance of fast samplers

Through Gaussian late start initialization (gls)

Gaussian approximation

- The early dynamics are approximately linear and mean reverting



(a) MNIST (b) CIFAR10 (c) Imagenet64 (d) CelebA64

Figure 19: The Shapiro-Wilk test assesses the normality of data over time, evaluated over 500 perturbed samples. It helps determine if the data closely follows a Multivariate Gaussian distribution up to a specific critical time.

- $\boldsymbol{x}_{start}$ no longer $N(\mathbf{0}, \boldsymbol{I})$
- We initialize the samplers just before the onset of the instability with a Gaussian approximation.
- We refer to this method as “Gaussian late start” (**gls**)

Visual results on CelebA64 with our Gaussian late start initialization

(a) DDPM n=5, FID=48.38 (c) gls-DDPM n=5, FID=31.24

(b) DDPM n=10, FID=36.66 (d) gls-DDPM n=10, FID=23.79

Figure 5: Comparison of stochastic DDPM samplers on CelebA64 with varying denoising steps. Subfigures (a) and (c) represent the generative model performance for 5 denoising steps, while (b) and (d) showcase the results for 10 denoising steps. The DDPM was initialized with the common standard initialization point $s_{start} = 800$ for 5 steps and $s_{start} = 900$ for 10 steps. Notably, our Gaussian late start initialization (gls-DDPM) with $s_{start} = 400$ for both 5 and 10 denoising steps demonstrates significant improvements in FID scores and diversity, leveraging spontaneous symmetry breaking in diffusion models.

Results using Gaussian late start initialization

Dataset	n	gls-DDPM	DDPM
MNIST	10	**4.21**	6.75
	5	**6.95**	13.25
	3	**11.92**	42.63
CIFAR10	10	**28.77**	43.35
	5	**42.46**	84.82
	3	**57.03**	146.95
CelebA32	10	**11.05**	26.79
	5	**14.79**	40.92
	3	**18.93**	59.75
Imagenet64	10	**57.31**	65.68
	5	**75.11**	99.99
	3	**91.69**	145.71
CelebA64	10	**23.79**	36.66
	5	**31.24**	48.38
	3	**37.05**	62.18

(a) DDPM

Dataset	n	gls-DDIM	DDIM
MNIST	10	**2.44**	4.46
	5	**6.95**	13.25
	3	**11.92**	42.63
CIFAR10	10	**15.98**	19.79
	5	**26.36**	44.61
	3	**42.31**	109.37
CelebA32	10	**7.27**	11.37
	5	**10.83**	23.45
	3	**16.24**	45.34
Imagenet64	10	**36.25**	38.21
	5	**52.11**	68.21
	3	**76.92**	126.3
CelebA64	10	**15.82**	19.37
	5	**22.06**	28.51
	3	**29.96**	50.304

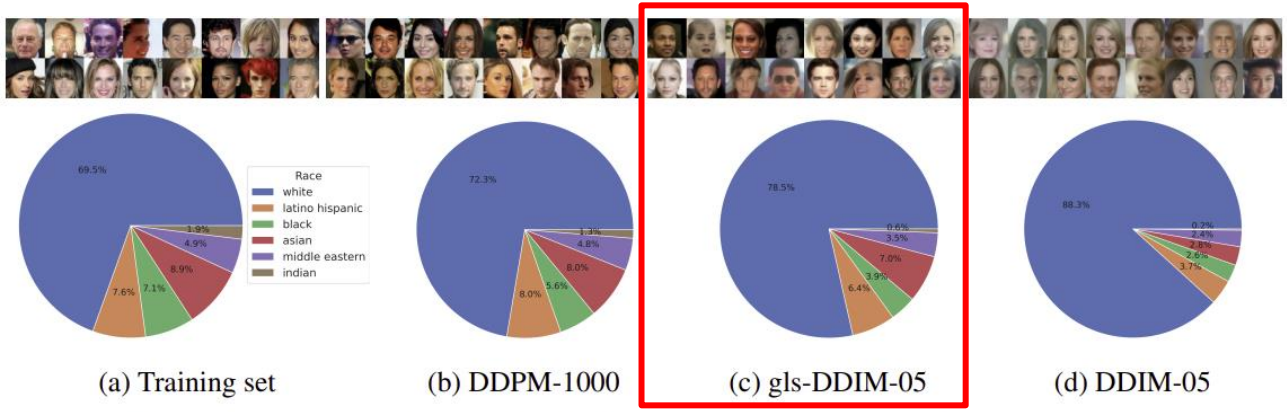
(b) DDIM

Dataset	n	gls-PNDM	PNDM
MNIST	10	**5.02**	14.36
	5	**5.11**	21.22
	3	**38.23**	154.89
CIFAR10	10	**5.90**	8.35
	5	**9.55**	13.77
	3	**34.20**	103.11
CelebA32	10	**2.88**	4.92
	5	**4.2**	6.61
	3	**28.60**	235.87
Imagenet64	10	**27.9**	28.27
	5	**33.35**	34.86
	3	**50.92**	70.58
CelebA64	10	**6.80**	8.03
	5	**9.26**	10.26
	3	**51.72**	171.75

(c) PNDM

Table 1: Summary of findings regarding image generation quality, as measured by FID scores. The performance of the stochastic DDPM sampler (a) is compared to the deterministic DDIM (b) and PNDM (c) samplers in the vanilla case, as well as our Gaussian late start initialization scheme denoted as "gls". Results are presented for 3, 5, and 10 denoising steps (denoted as "n") across diverse datasets.

Diversity analysis



(a) Training set (b) DDPM-1000 (c) gls-DDIM-05 (d) DDIM-05

Figure 7: "Race" diversity analysis on CelebA64 over 50,000 generated samples by (c) gls-DDIM and (d) DDIM samplers with 5 denoising steps. Results obtained on (a) training set and (b) DDPM using 1000 denoising steps are provided for reference. Corresponding samples obtained by each set are shown on top of the pie charts.

High generative diversity relies on sampling from the narrow temporal window around the critical time, since small perturbations during that window are amplified by the instability.

Summary

- We show that the generative dynamics of diffusion models exhibit a spontaneous symmetry-breaking phenomenon, resulting in two distinct generative phases:
  - **Phase 1**: A linear steady-state dynamics centered around a fixed point.
  - **Phase 2**: An attractor dynamics that guides the model towards the data manifold.
- The **period of instability** during this transition contributes to the **diversity** of generated samples (critical phenomena).
- We propose a Gaussian late initialization scheme that enhances model performance, resulting in up to a 3x FID improvement on fast samplers.
  - The early phase does not significantly contribute to the model's performance.

Thank you for your attention!

g.raya@jads.nl, l.ambrogioni@donders.ru.nl

SCAN ME

For more information