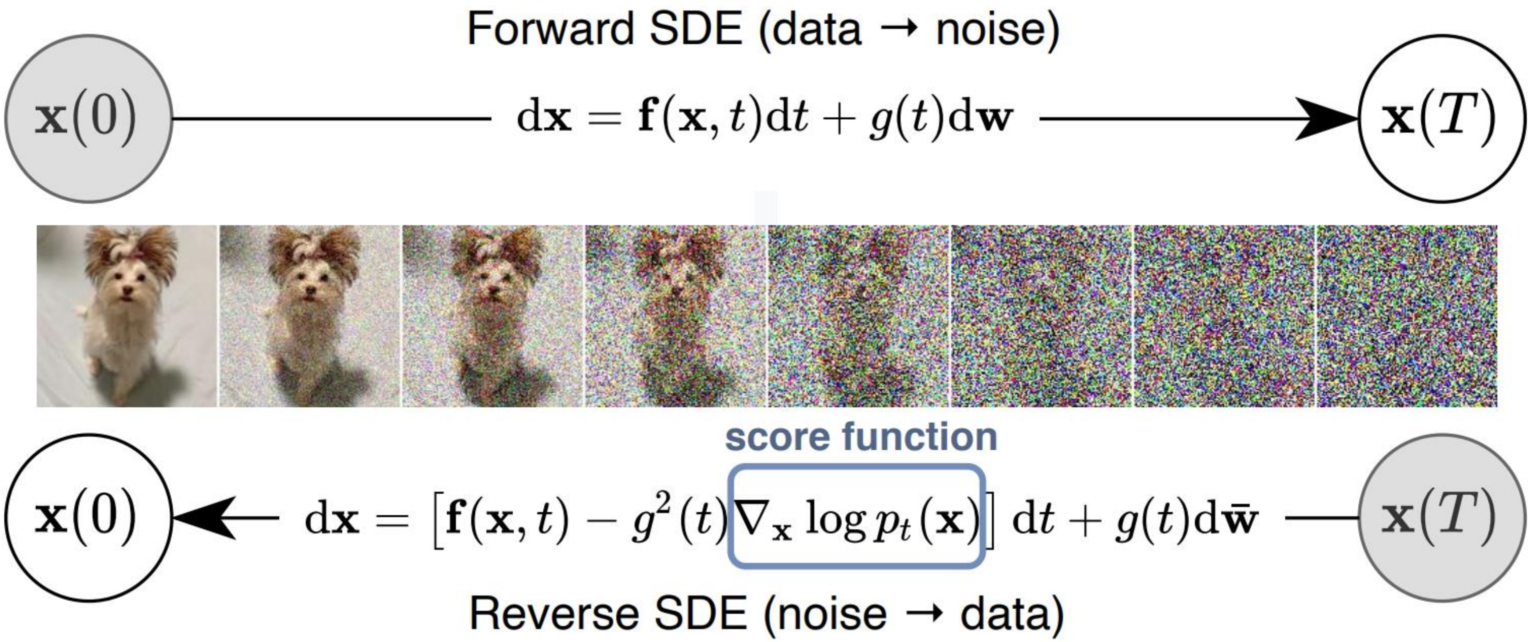
Smoothed Energy Guidance: Guiding Diffusion Models with Reduced Energy Curvature of Attention

Susung Hong

University of Washington

Introduction



Song, Yang, et al. "Score-based generative modeling through stochastic differential equations." ICLR 2021.

Introduction

Classifer-Free Guidance [Ho and Salimans, 2022]

- Guidance without a classifier
- Reverse SDE

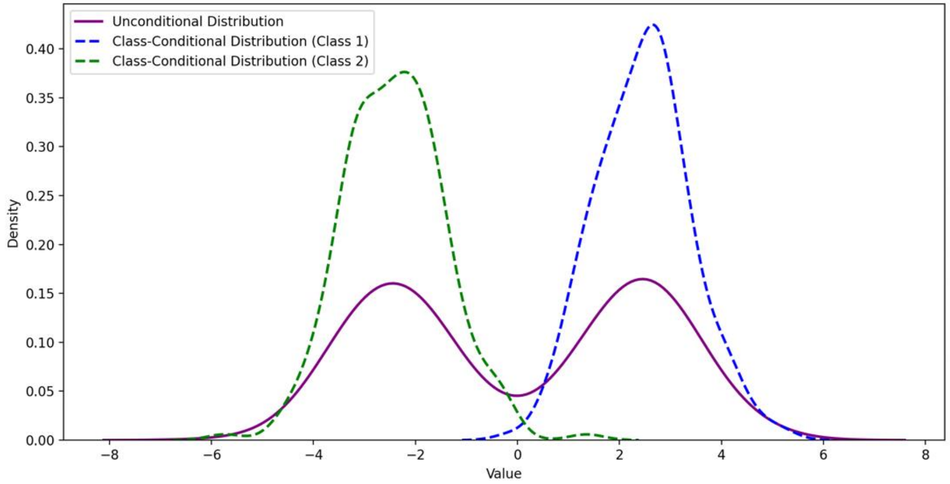
$$d\mathbf{x} = [\mathbf{f}(\mathbf{x}, t) - g(t)^2(\gamma_{\text{cfg}}\mathbf{s}_\theta(\mathbf{x}, t, c) - (\gamma_{\text{cfg}} - 1)\mathbf{s}_\theta(\mathbf{x}, t))]dt + g(t)d\bar{\mathbf{w}}$$

Ho, Jonathan, and Tim Salimans. "Classifier-free diffusion guidance." arXiv preprint arXiv:2207.12598 (2022).

Introduction

Classifer-Free Guidance [Ho and Salimans, 2022]

- Moves samples towards **“sharper”** class-conditional distributions



Ho, Jonathan, and Tim Salimans. "Classifier-free diffusion guidance." arXiv preprint arXiv:2207.12598 (2022).

Main Idea

We propose using the energy of the attention weights, inspired by the modern (continuous) **Hopfield networks** (Definition 2.1)

$$E(\mathbf{A}) := \sum_{i=1}^{H} \sum_{j=1}^{W} E'(\mathbf{a}_{:(i,j)})$$

$$E'(\mathbf{a}) := -\mathrm{lse}\left(\mathbf{a}\right) = -\log\left(\sum_{k=1}^{H}\sum_{l=1}^{W} e^{a_{(k,l)}}\right)$$

Hong, Susung. "Smoothed Energy Guidance: Guiding Diffusion Models with Reduced Energy Curvature of Attention." NeurIPS 2024.

Main Idea

Blurring attention weights

- Preserves the mean and decreases the variance (Lemma 3.1)
- Increases the value of the energy function (Lemma 3.2)
- Results in decreasing the Gaussian curvature of the energy landscape (Theorem 3.1)

Main Idea

Blurring attention weights

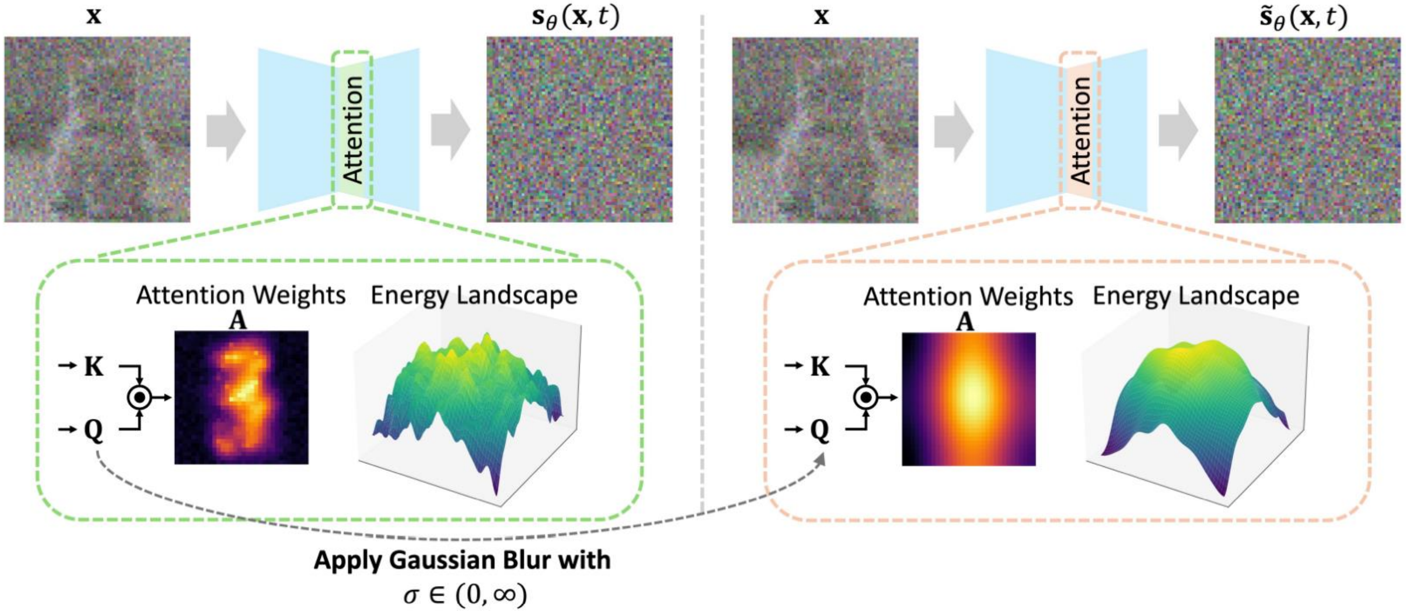
- Is equivalent to blurring queries (Proposition 3.1)

$$G * (\mathbf{Q}\mathbf{K}^\top) = \mathbf{B}(\mathbf{Q}\mathbf{K}^\top)$$

$$\mathbf{B}(\mathbf{Q}\mathbf{K}^\top) = (\mathbf{B}\mathbf{Q})\mathbf{K}^\top = (G * \mathbf{Q})\mathbf{K}^\top$$

Main Idea

Smoothed Energy Guidance



Hong, Susung. "Smoothed Energy Guidance: Guiding Diffusion Models with Reduced Energy Curvature of Attention." NeurIPS 2024.

Main Idea

Smoothed Energy Guidance

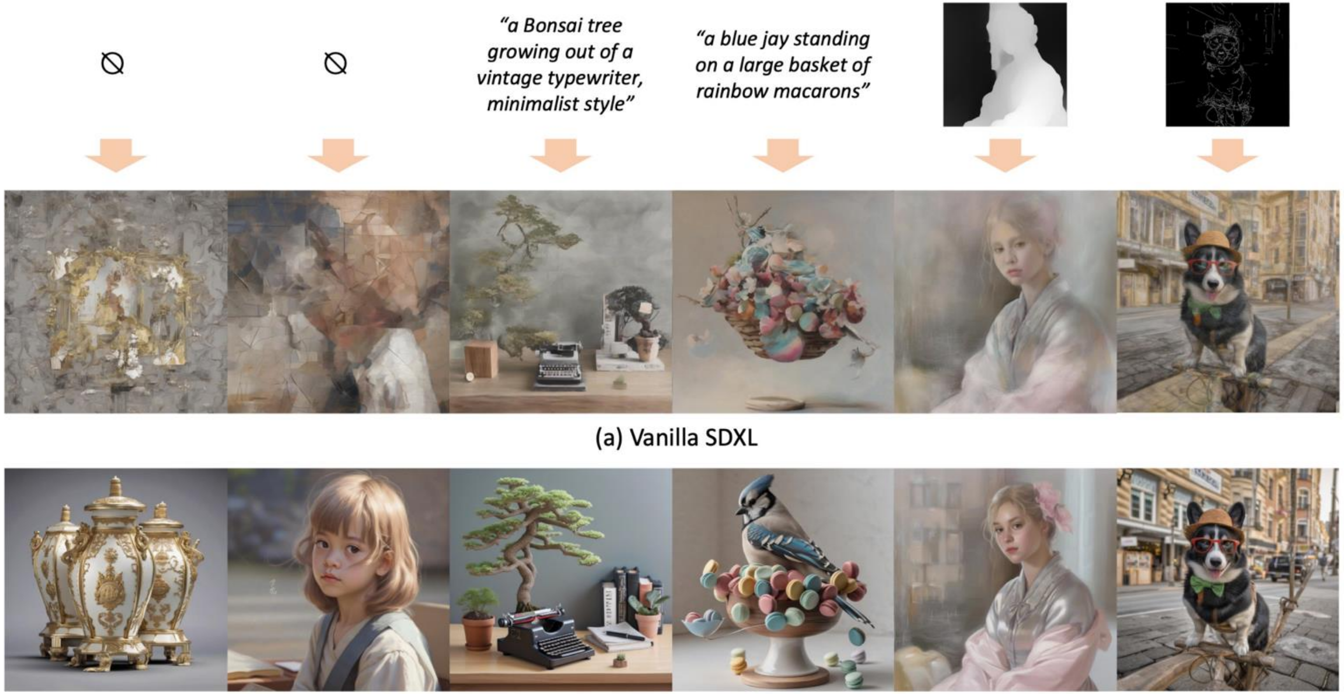
- We propose training- and condition-free guidance

$$d\mathbf{x} = [\mathbf{f}(\mathbf{x}, t) - g(t)^2(\gamma_{\text{seg}}\mathbf{s}_\theta(\mathbf{x}, t) - (\gamma_{\text{seg}} - 1)\tilde{\mathbf{s}}_\theta(\mathbf{x}, t))]dt + g(t)d\bar{\mathbf{w}}$$

- We can control the effect with $\sigma \in (0, \infty)$ of the Gaussian blur

Hong, Susung. "Smoothed Energy Guidance: Guiding Diffusion Models with Reduced Energy Curvature of Attention." NeurIPS 2024.

Experiments



(a) Vanilla SDXL

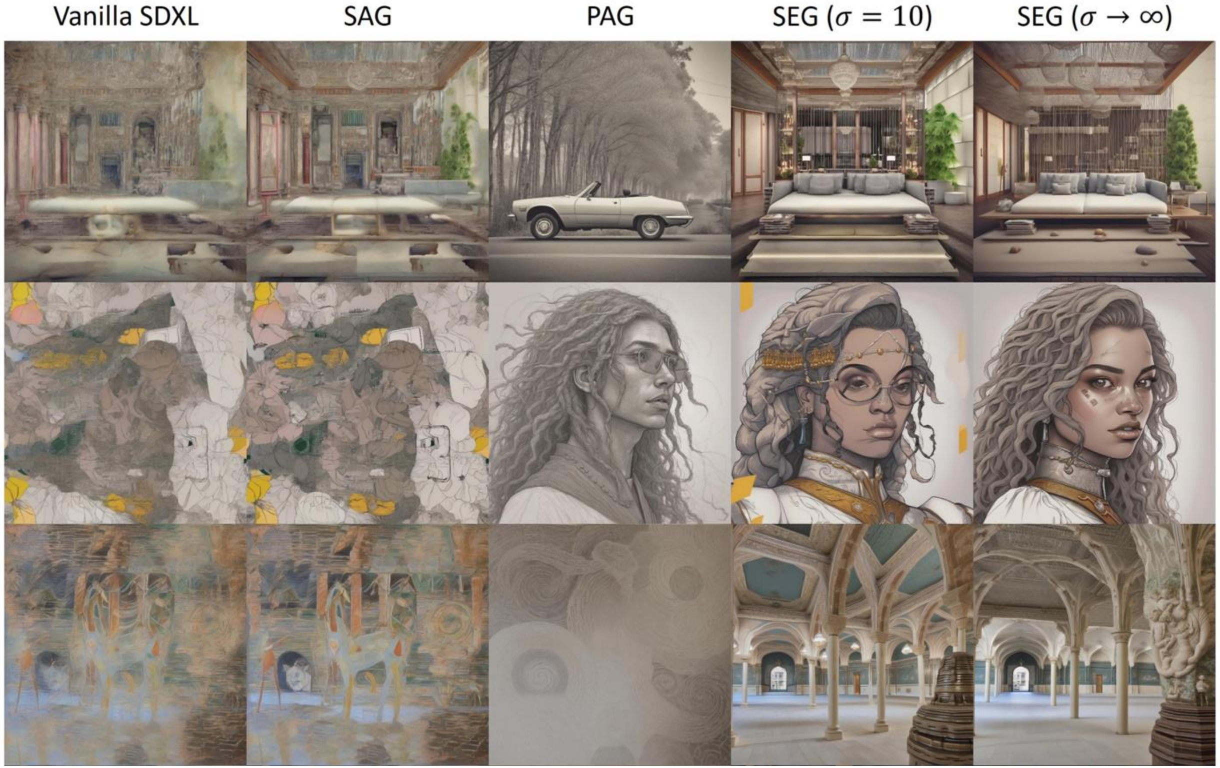
(b) SDXL + Smoothed Energy Guidance (Ours)

Hong, Susung. "Smoothed Energy Guidance: Guiding Diffusion Models with Reduced Energy Curvature of Attention." NeurIPS 2024.

Experiments

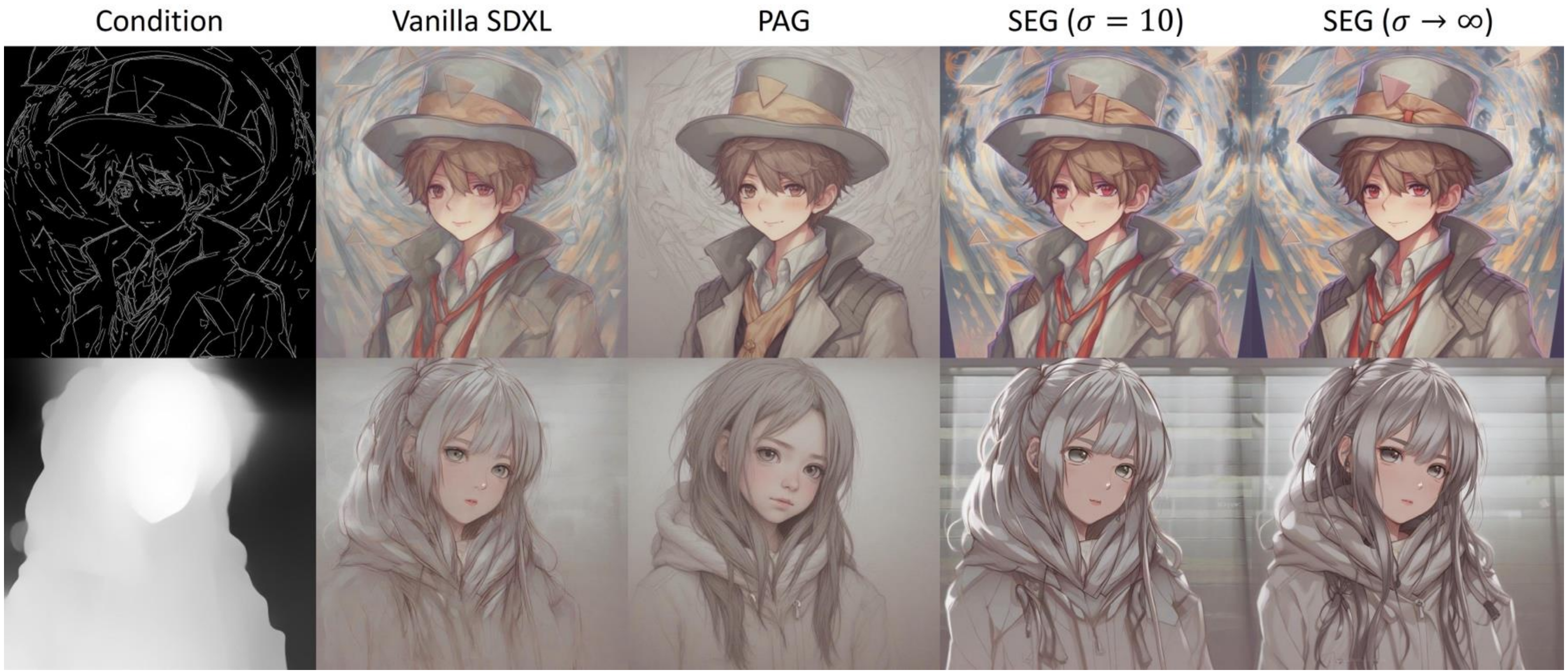
Hong, Susung. "Smoothed Energy Guidance: Guiding Diffusion Models with Reduced Energy Curvature of Attention." NeurIPS 2024.

Experiments



Hong, Susung. "Smoothed Energy Guidance: Guiding Diffusion Models with Reduced Energy Curvature of Attention." NeurIPS 2024.
Ahn, Donghoon, et al. "Self-Rectifying Diffusion Sampling with Perturbed-Attention Guidance." ECCV 2024.
Hong, Susung, et al. "Improving sample quality of diffusion models using self-attention guidance." ICCV 2023.

Experiments

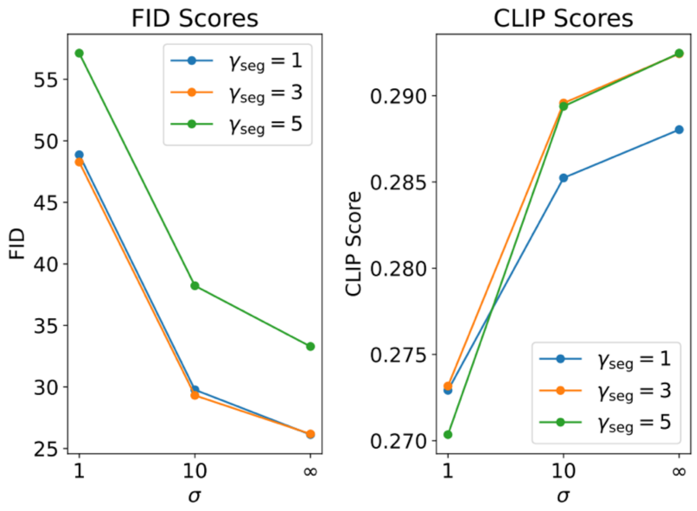


Hong, Susung. "Smoothed Energy Guidance: Guiding Diffusion Models with Reduced Energy Curvature of Attention." NeurIPS 2024.

Experiments

Metric	Vanilla SDXL [35]	SAG [17]	PAG [1]	SEG $\sigma = 10$	SEG $\sigma \to \infty$
FID↓	129.496	106.683	105.271	<u>95.316</u>	**88.215**
$\text{LPIPS}_{\text{vgg}}$ ↓	-	0.706	0.542	**0.522**	<u>0.536</u>
$\text{LPIPS}_{\text{alex}}$ ↓	-	0.644	0.472	**0.454**	<u>0.472</u>

Metric	Vanilla SDXL [35]	SEG				
		1	2	5	10	∞
FID↓	53.423	48.284	41.784	33.819	<u>29.325</u>	**26.169**
CLIP Score↑	0.271	0.273	0.278	0.285	<u>0.290</u>	**0.292**
$\text{LPIPS}_{\text{vgg}}$ ↓	-	**0.361**	<u>0.410</u>	0.449	0.472	0.493
$\text{LPIPS}_{\text{alex}}$ ↓	-	**0.295**	<u>0.347</u>	0.390	0.416	0.440



Hong, Susung. "Smoothed Energy Guidance: Guiding Diffusion Models with Reduced Energy Curvature of Attention." NeurIPS 2024.