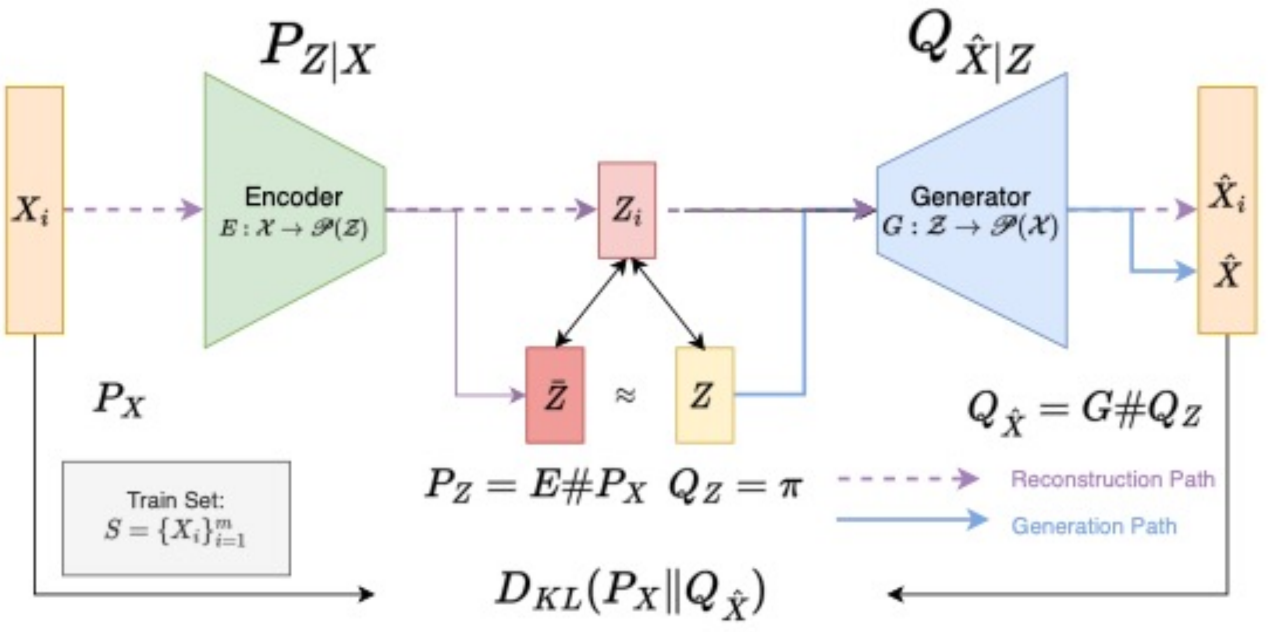
Generalization in VAE and Diffusion Models: A Unified Information-theoretic Analysis

Qi Chen, Jierui Zhu, Florian Shkurti

2025.03

General Theorem for Encoder-Generator Structure



$$|\mathrm{gen}_{P_X}^{\Delta_G}(E,G,\pi)| \leq \underbrace{\frac{\sqrt{2}R}{m}\sum_{i=1}^{m}\sqrt{\mathbb{E}_{X_i}[\mathbb{D}_{KL}(E(X_i)\|\pi)]}}_{\text{encoder related}} + \underbrace{\frac{\sqrt{2}R}{m}\sum_{i=1}^{m}\sqrt{I(\hat{X}_i; X_i|Z_i)}}_{\text{generator related}}.$$

where $gen_{P_X}^{\Delta_G}(E,G,\pi) \stackrel{\text{def}}{=} \mathbb{E}_{S\sim P_X^m}\left[\mathcal{L}_{P_X}^{\pi}(E,G) - \mathcal{L}_{\hat{P}_X}(E,G)\right]$.

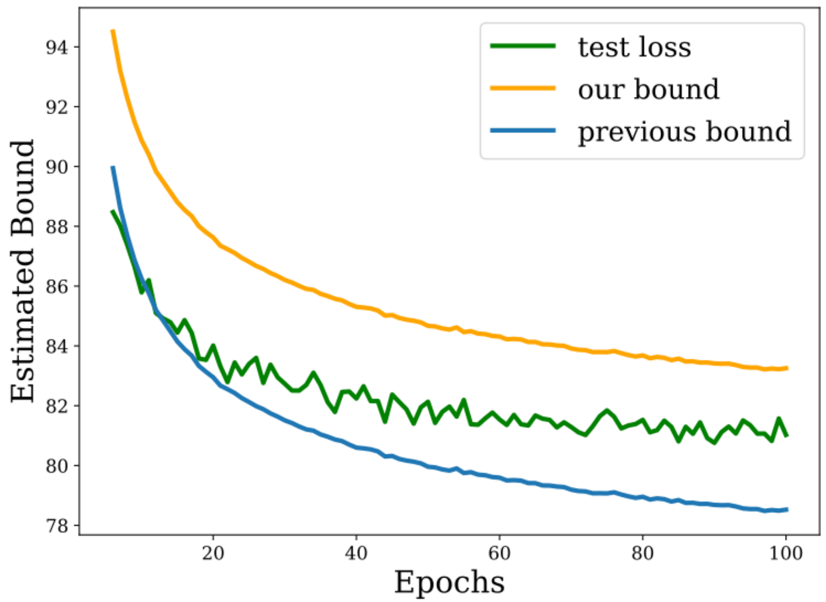
VAEs

$$\frac{1}{m}\sum_{i=1}^{m} I(\hat{X}_i; X_i|Z_i) \le \frac{1}{m}\mathbb{E}_S\left(\sum_{i=1}^{m}\frac{1}{m}\mathbb{E}_{Z_i\sim E_\phi(X_i)} D_{KL}(G_\theta(Z_i)\|G_{\bar{\theta}}(Z_i))\right)$$

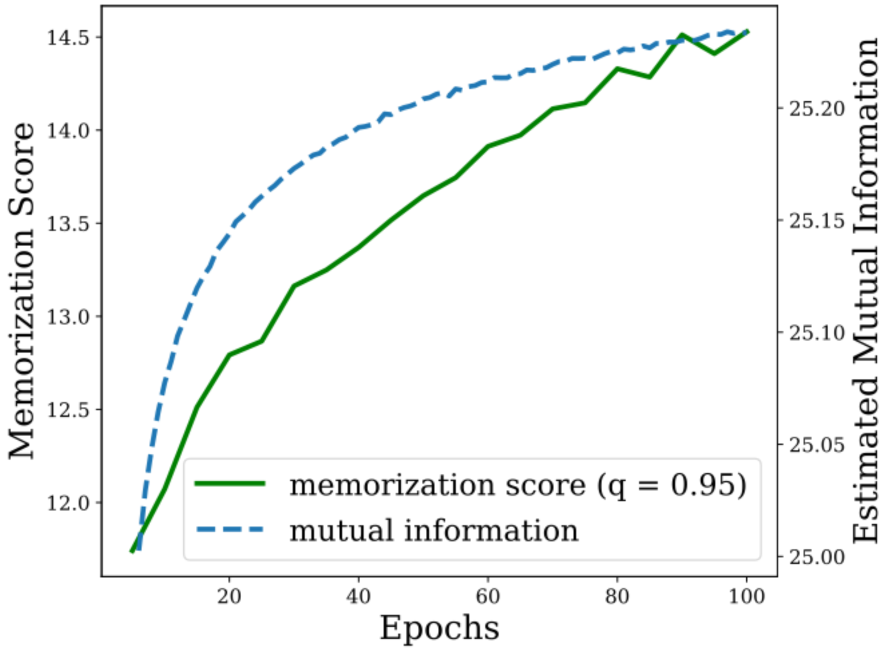
Bound:

$$\mathbb{D}_{W_1}(P_X\|Q^{\pi}_{G_\theta}) \le \overbrace{\mathbb{E}_S\mathcal{L}_{\hat{P}_X}(E_\phi, G_\theta) + \frac{\sqrt{2}R}{m}\sum_{i=1}^{m}\sqrt{\mathbb{E}_{X_i}[\mathbb{D}_{KL}(E_\phi(X_i)\|\pi)]}}^{\text{Common VAE Objective}} + I(\hat{X}_i, X_i|Z_i)$$

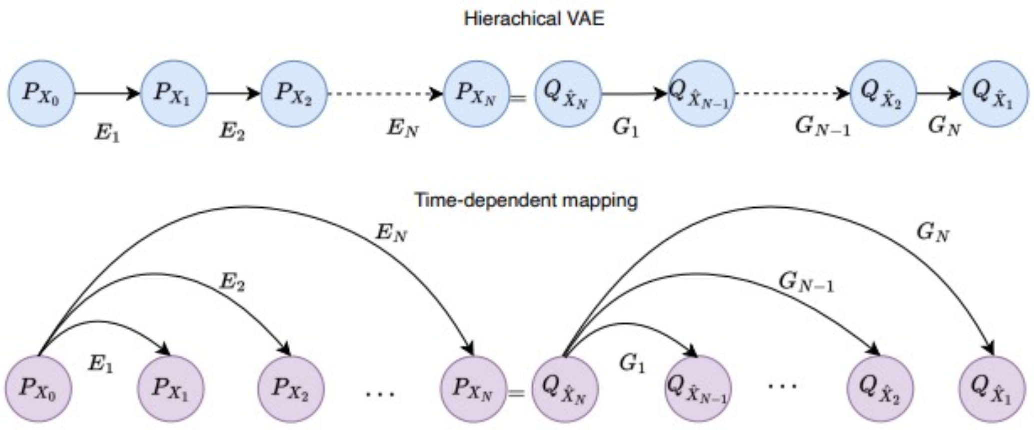
Necessity of the information term

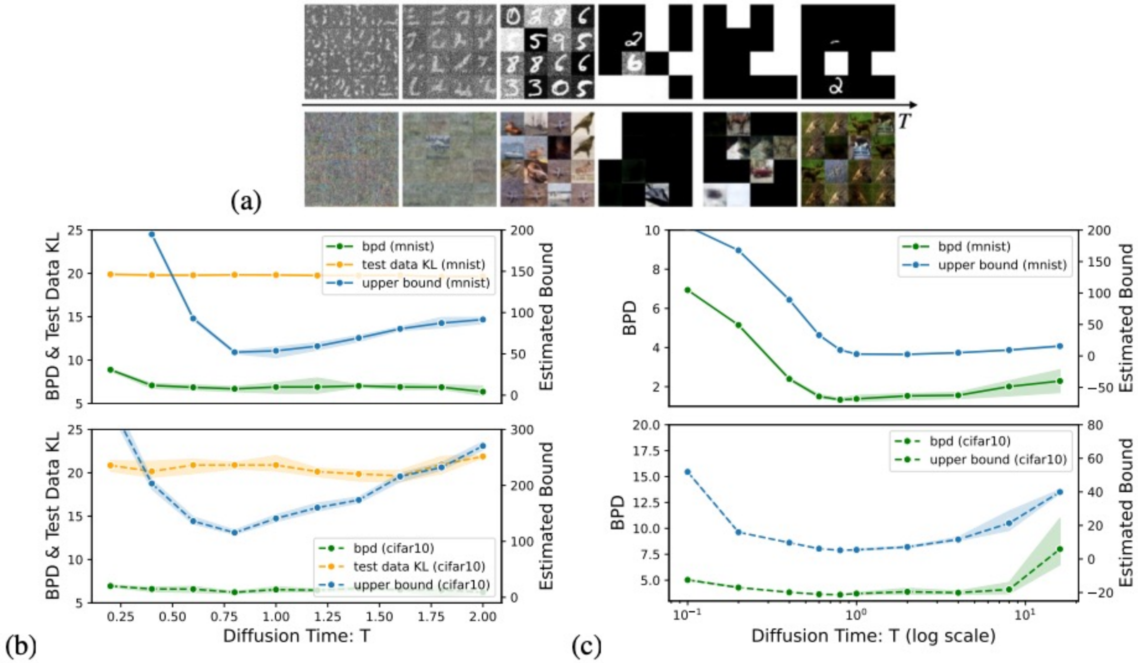


Bound Reflects Memorization in VAEs:



Diffusion Models





Trade-off on Diffusion Time T

Diffusion Models:

$$\mathbb{D}_{KL}(P_X\|Q^{\pi}_{G^{\theta}_{T}}) \le \mathbb{E}_S\Bigg(\underbrace{-\frac{1}{m}\sum_{i=1}^{m}\mathbb{D}_{KL}(E_T(X_i)\|E_T\#\hat{P}_X)}_{T_1}+\hat{\mathcal{L}}_{ESM}(\theta,\lambda(\cdot))\Bigg)$$

$$+\underbrace{\frac{\sqrt{2}R}{m}\sum_{i=1}^{m}\sqrt{\mathbb{E}_{X_i}[\mathbb{D}_{KL}(E_T(X_i)\|\pi)]}}_{T_2}+\underbrace{\frac{\sqrt{2}R}{m}\sum_{i=1}^{m}\sqrt{I(\hat{X}_0;X_i|\hat{X}_T)}}_{T_3}.$$

$$\le \frac{TL^2\sum_{k=1}^{N}\lambda^2(\frac{(k-1)T}{N})}{2mN}.$$

$$\mathbb{D}_{KL}(E_T(X_i)\|\pi)=\frac{1}{2}\left(e^{-\beta_T}X_i^TX_i-de^{-\beta_T}-d\log(1-e^{-\beta_T})\right)\xrightarrow{T\to\infty}0.$$

"Generalization in VAE and Diffusion Model: A Unified Information-theoretic Analysis" Qi Chen, Jierui Zhu, Florian Shkurti, ICLR 2025

Thanks!