Choose your Diffusion

Efficient and flexible ways to accelerate ~~diffusion~~ (DM/CFM) in HEP

Cheng Jiang, Sitian Qian, Huilin Qu

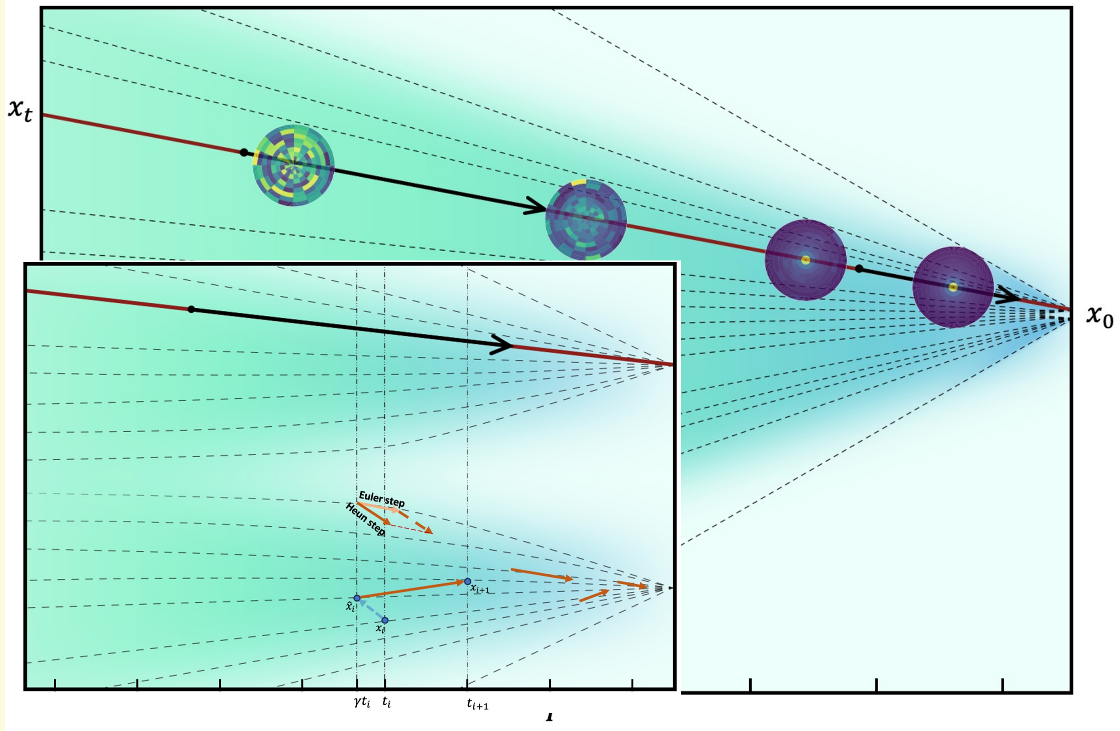
2404.18219

2401.13162

CERN
PEKING UNIVERSITY 1898
THE UNIVERSITY of EDINBURGH

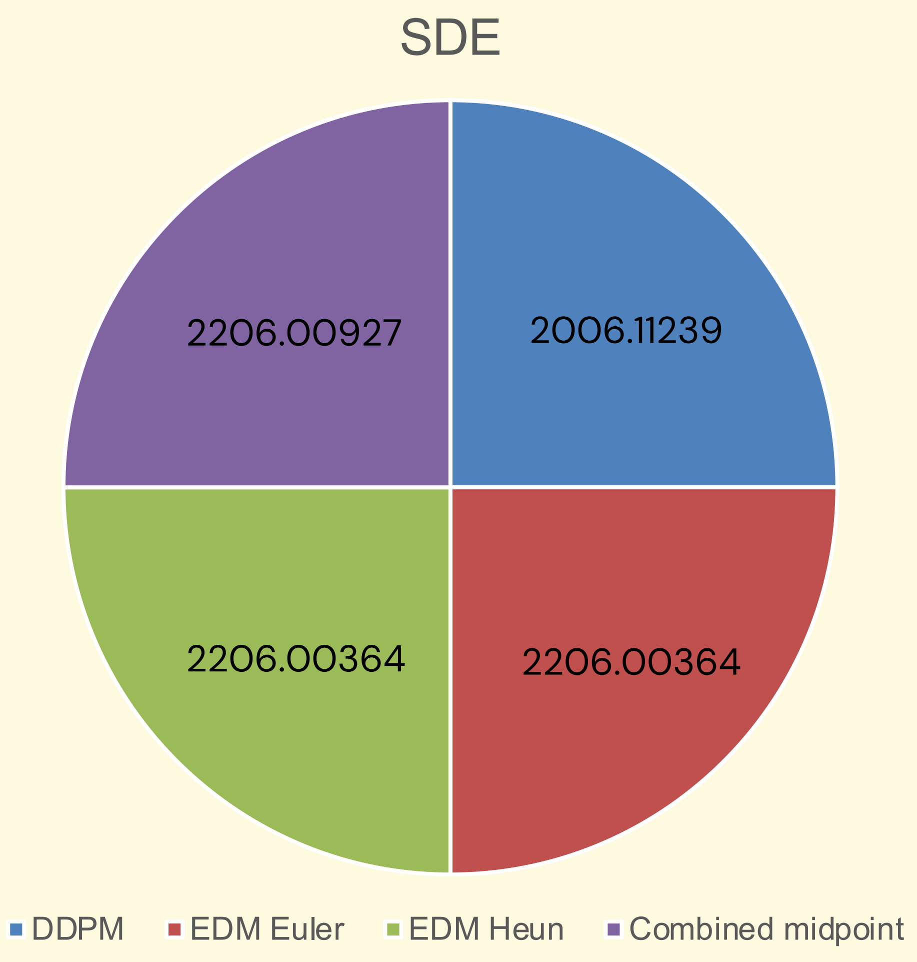
Motivation

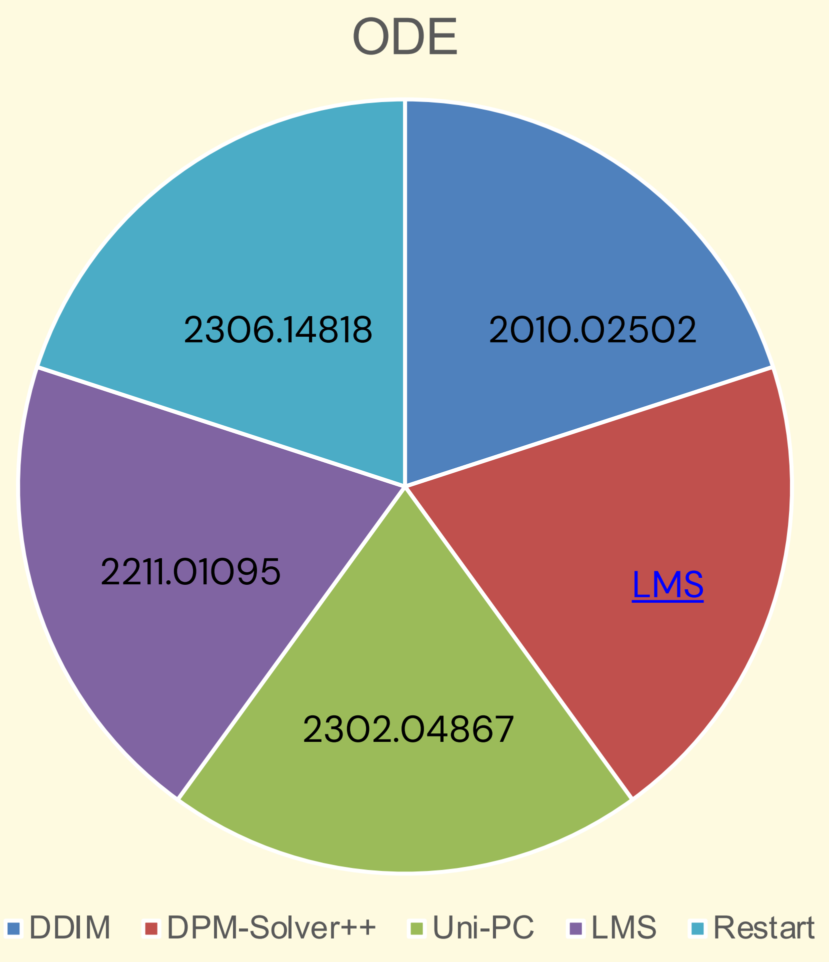
- The study focuses mostly on Score Matching, in which the score function is solved by different choices of SDE/ODE. How we could effectively accelerate the generative model, by replacing only parts of that.



- Backward process (training-free):

We have adopted **almost all mainstream samplers/schedulers** to do comprehensive comparisons **on both shower cells** (*CaloChallenge*) **and jet constituents** (*JetNet*)



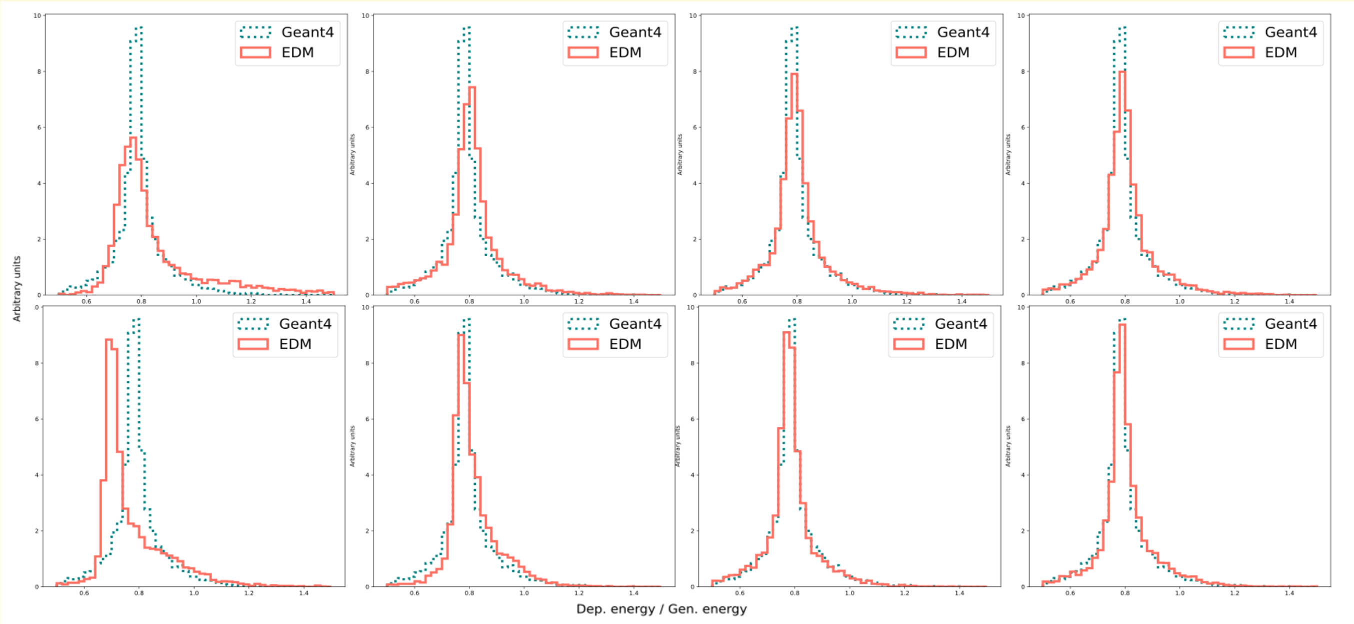


- Forward process (faster divergence):

Effective way to mitigate the challenging optimization: **Denoiser function with preconditioning parameters, weighted by min-Signal-to-Noise ratio (min-SNR)**

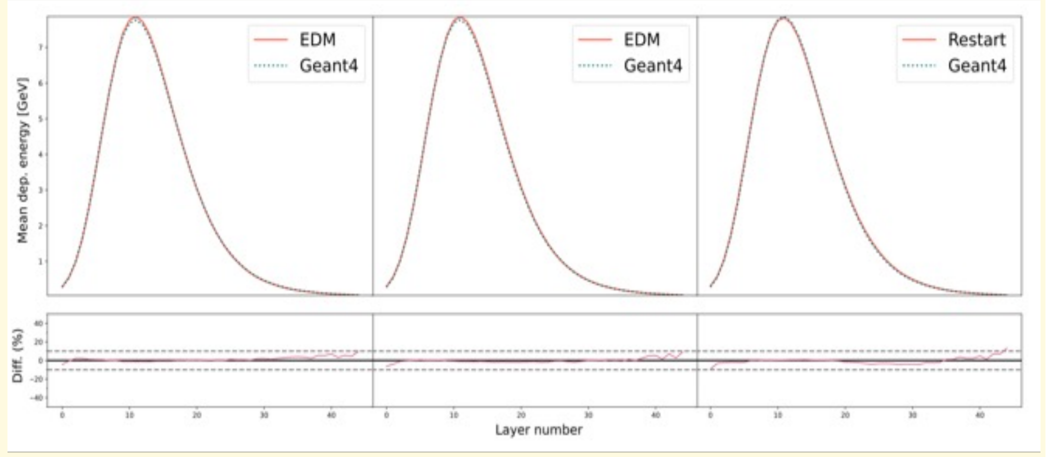
$$\mathcal{L} = \mathbb{E}_{t,\varepsilon}[w(t)||F(c_{in}x_t, t) - \frac{1}{c_{out}}(x_0 - c_{skip}x_t)||_2^2]$$

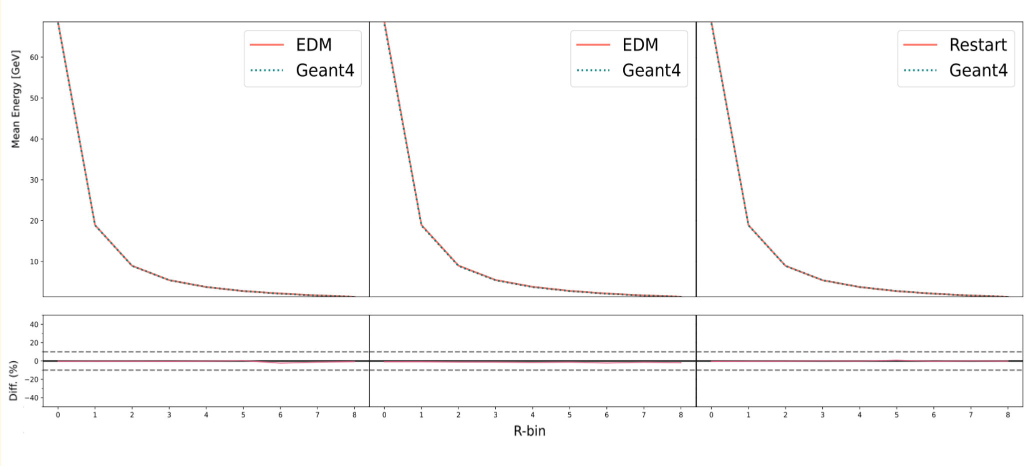
$$w(t) = {(t * \sigma_c)} / {(t^2 + \sigma_c^k)^2}$$



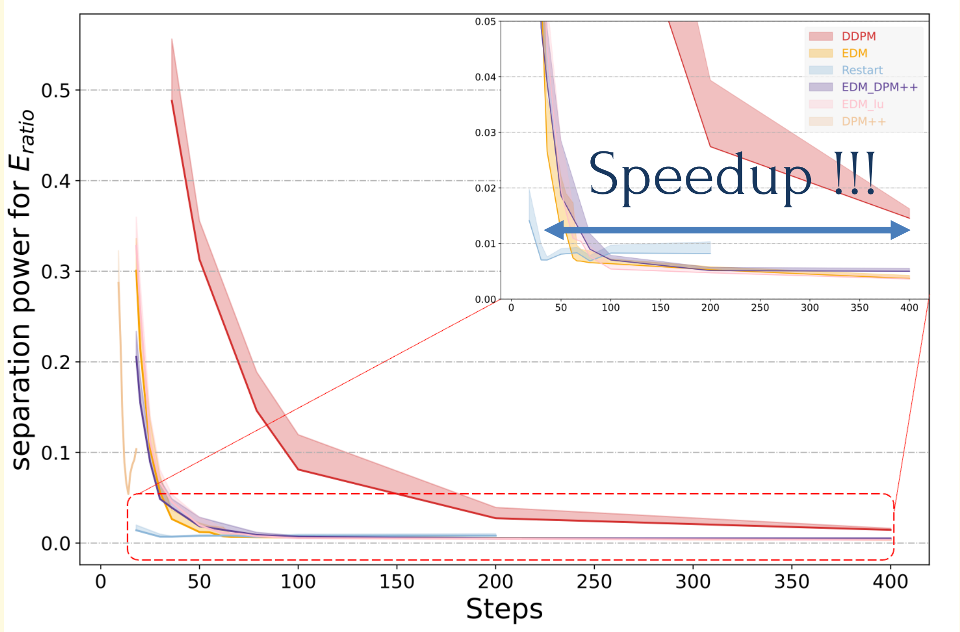
Results & More (Wed Loc #45)

Indistinguishable high-level features for shower from cell-level generations





Achieve O(10) acceleration with comparable performance for current benchmark models



	AUC	Separation	FPD
DDPM 79/200/400	55.3/53.2/52.6	0.0810/0.0344/*0.0155*	0.074/0.046/*0.043*
EDM 39/79/200	54.1/**52.0**/51.5	0.0256/0.0076/**0.0055**	0.035/**0.027**/0.023
EDM_DPM++ 79	52.3	0.0103	0.026
EDM_Lu 79	52.2	0.0086	0.026
Restart 18/36/79	55.2/**52.0**/51.8	0.0169/0.073/0.0057	0.059/**0.025**/0.022
LMS 36	53.8	0.0305	0.095
DPM++ 20	59.8	0.0534	0.146
Uni-PC 20	60.3	0.1304	0.152

How about replacing the backbone for the model? **Changing flow matching with Unet/Transformer backbone to GBDT, which latter has much faster training and inference time. Is this even possible?? YES!** (BUFF: BDT based-ultra-fast flow matching.) Few mins training, below millisecond generation time, could replace most flow-based model. E.g. Unfolding, huge improvement on correlation difference