Randomized midpoint for diffusion sampling

Stephen Huan

Carnegie Mellon University
https://cgdct.moe

SLMath lightning talk

Collaborators

http://arxiv.org/abs/2511.04844
https://github.com/stephen-huan/edm_rmd

Theorists: Matthew Zhang, Sitan Chen, Sinho Chewi
Empiricists: Stephen Huan, Jerry Huang, Nicholas Boffi



Matthew Zhang,
UToronto



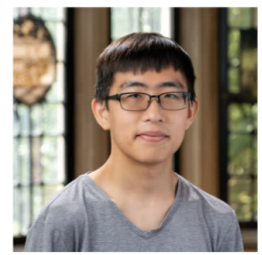
Jerry Huang,
CMU



Nick Boffi,
CMU



Sitan Chen,
Harvard



Sinho Chewi,
Yale

Randomized midpoint

Introduced by Shen and Lee (2019) for discretization of ULA

We port to diffusion model sampling (PF-ODE + SDE)

Expect to also apply to interpolants, etc.

The method

The crux of the method: $\dot{X}_t = b_t(X_t)$ has solution

$$X_h = X_0 + \int_0^h b_t(x_t)\,\mathrm{d}t$$

Unbiased estimator (Monte Carlo quadrature): $\tau \sim \mathrm{Unif}[0, h]$,

$$\hat{X}_h = \hat{X}_0 + h b_\tau(\hat{X}_\tau)$$

But we don't know $\hat{X}_\tau$! Approximate with (left) Euler

$$\hat{X^+_\tau} = \hat{X}_0 + \tau b_0(\hat{X}_0)$$
$$\hat{X}_h = \hat{X}_0 + h b_\tau(\hat{X^+_\tau})$$

Generalization to semi-linear SDEs

$$\mathrm{d}X_t = (\lambda(t)X_t + b_t(X_t))\,\mathrm{d}t + g(t)\,\mathrm{d}B_t$$

Performance

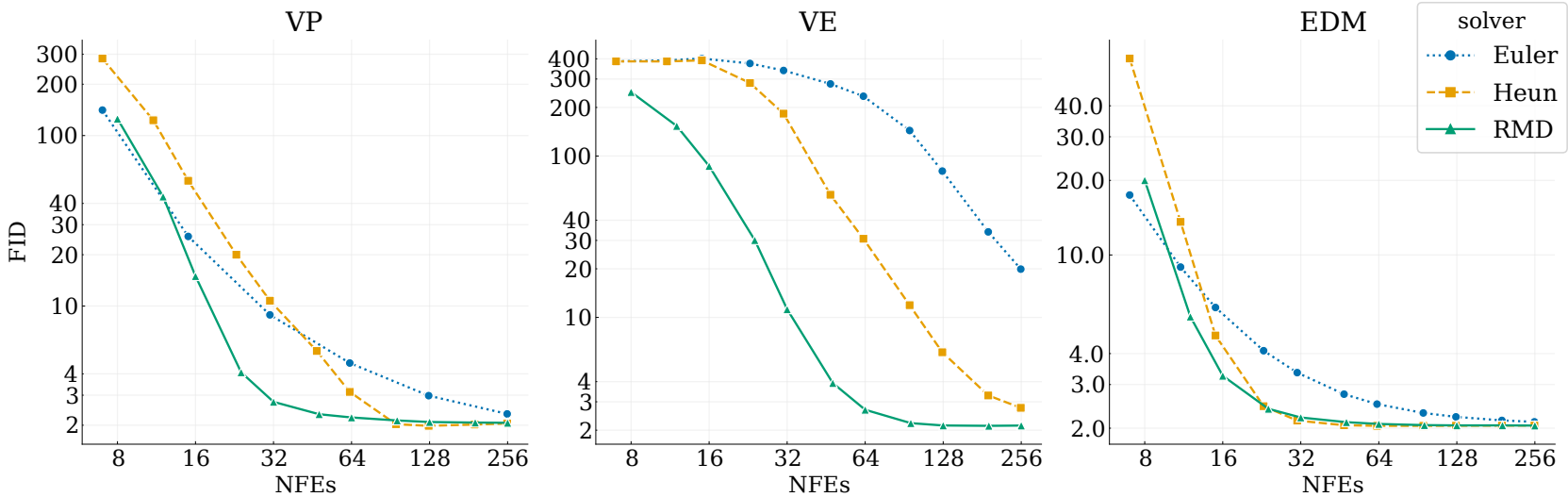


Figure: Fréchet inception distance (FID↓) with NFEs for ODE sampling

Quick notes on the theory

Assumptions: data has finite variance, ε^2 accurate score, estimated score has (time-varying) Lipschitz constant

Result: $\mathcal{O}(\sqrt{d}/\epsilon)$ steps in $W_2 + \mathsf{KL}$ (not early stopping)

Difficulties:

- Anticipating: midpoint "sees into the future"!
- Weak error ("bias") vs strong error ("variance")
- Standard Wasserstein analyses blow up exponentially in time

We use ✨shifted composition✨ to address these issues

Thank You!