Improved Sampling Algorithms for Lévy-Itô Diffusion Models

Vadim Popov, Assel Yermekova, Tasnima Sadekova, Artem Khrapov, Mikhail Kudinov
Huawei Noah's Ark Lab

ICLR 2025, Singapore

Lévy-Itô Diffusion Models: overview

- One of known drawbacks of common diffusion models is their poor performance on underrepresented classes in the scenario when training dataset is highly imbalanced
- Recently, it has been proposed [1] to replace Gaussian noise in diffusion models with α–stable distribution with heavier tails to allow for better diversity and mode coverage in the mentioned scenario

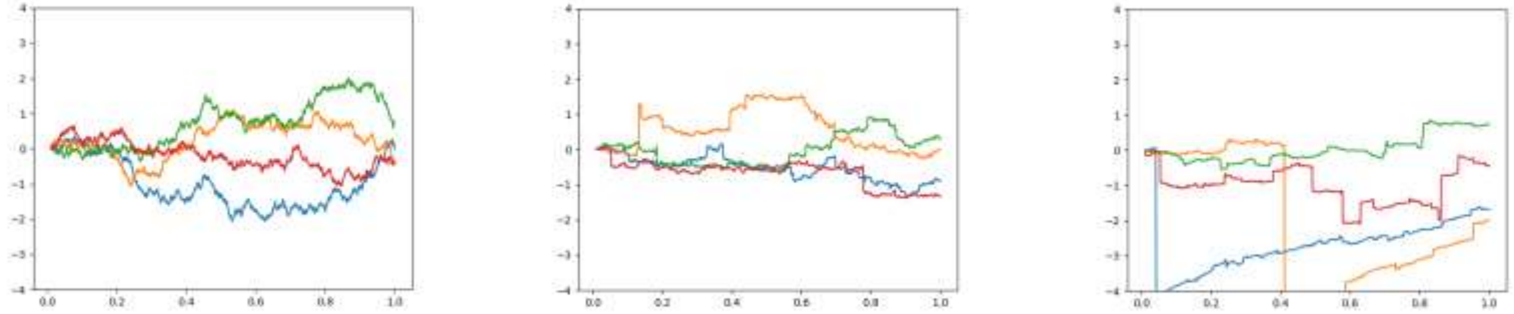
[1] Yoon et al. Score-based Generative Models with Lévy Processes, 2023

Lévy-Itô Diffusion Models: overview

- One of known drawbacks of common diffusion models is their poor performance on underrepresented classes in the scenario when training dataset is highly imbalanced
- Recently, it has been proposed [1] to replace Gaussian noise in diffusion models with α–stable distribution with heavier tails to allow for better diversity and mode coverage in the mentioned scenario
- Forward SDE now relied on a certain Lévy process and the prior is standard α–stable random variable:

$$dX_t = \mu(X_{t-}, t)dt + \sigma_t dL_t^\alpha$$

- Difference between Wiener process and α–stable Lévy processes for different α:



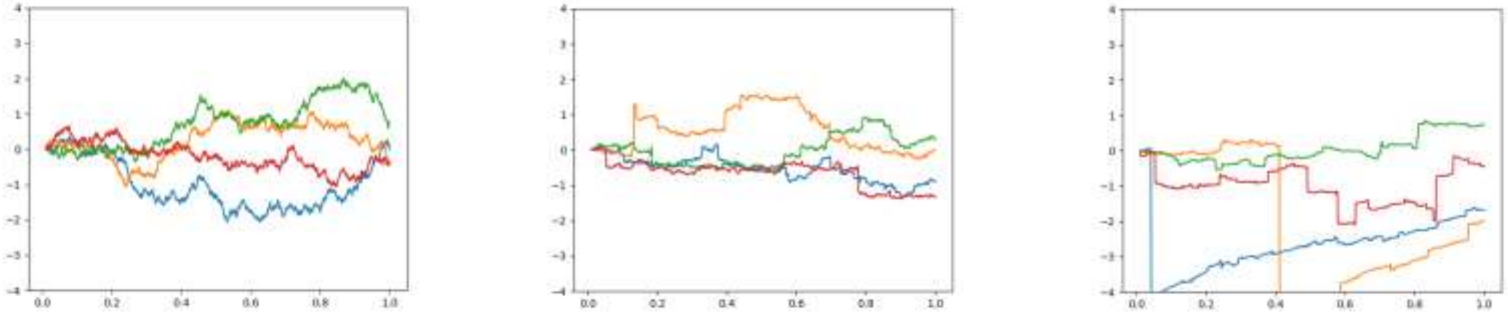
[1] *Yoon et al. Score-based Generative Models with Lévy Processes, 2023*

Lévy-Itô Diffusion Models: overview

- One of known drawbacks of common diffusion models is their poor performance on underrepresented classes in the scenario when training dataset is highly imbalanced
- Recently, it has been proposed [1] to replace Gaussian noise in diffusion models with α–stable distribution with heavier tails to allow for better diversity and mode coverage in the mentioned scenario
- Forward SDE now relied on a certain Lévy process and the prior is standard α–stable random variable:

$$dX_t = \mu(X_{t-}, t)dt + \sigma_t dL_t^\alpha$$

- Difference between Wiener process and α–stable Lévy processes for different α:



- Reverse SDE [1]:

$$d\bar{X}_t = (\mu(\bar{X}_{t+}, t) - \alpha\sigma_t^\alpha S_t^{(\alpha)}(\bar{X}_{t+}))dt + \sigma_t d\bar{L}_t^\alpha + d\bar{Z}_t$$

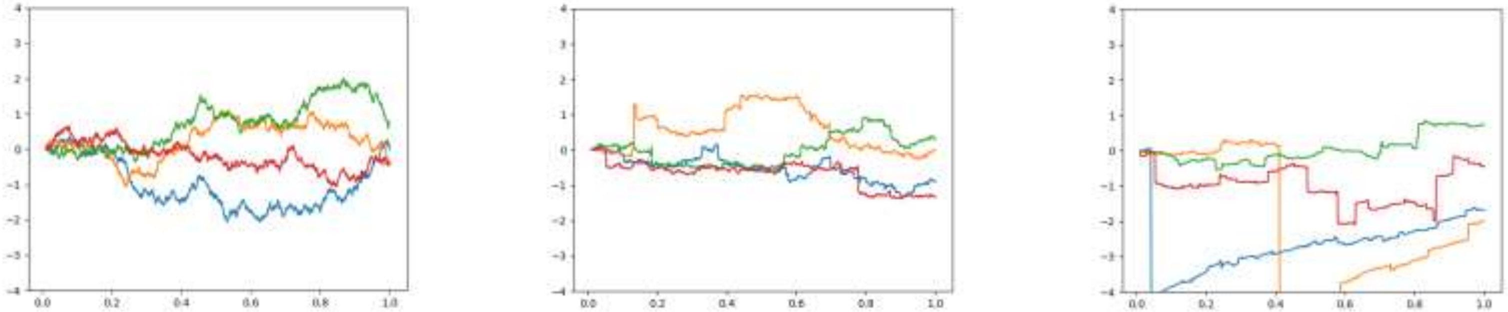
[1] Yoon et al. Score-based Generative Models with Lévy Processes, 2023

Lévy-Itô Diffusion Models: overview

- One of known drawbacks of common diffusion models is their poor performance on underrepresented classes in the scenario when training dataset is highly imbalanced
- Recently, it has been proposed [1] to replace Gaussian noise in diffusion models with α–stable distribution with heavier tails to allow for better diversity and mode coverage in the mentioned scenario
- Forward SDE now relied on a certain Lévy process and the prior is standard α–stable random variable:

$$dX_t = \mu(X_{t-}, t)dt + \sigma_t dL_t^\alpha$$

- Difference between Wiener process and α–stable Lévy processes for different α:



- Reverse SDE [1]:

$$d\bar{X}_t = (\mu(\bar{X}_{t+}, t) - \alpha\sigma_t^\alpha S_t^{(\alpha)}(\bar{X}_{t+}))dt + \sigma_t d\bar{L}_t^\alpha + d\bar{Z}_t$$

- The last term is usually skipped since it is a process of finite variation whose characteristic exponent is intractable:

$$\int_0^t \sigma_L^\alpha(s) \left[\int_0^1 \int_{\mathbb{R}^d} \left(e^{i\langle \mathbf{u}, \mathbf{y} \rangle} - 1 - i\langle \mathbf{u}, \mathbf{y} \rangle \cdot 1_{||\mathbf{y}||<1}(\mathbf{y}) \right) \frac{\langle \mathbf{y}, \nabla p_s(\mathbf{x} + u\mathbf{y}) \rangle}{p_s(\mathbf{x})} \nu(d\mathbf{y}) du \right] ds.$$

[1] Yoon et al. Score-based Generative Models with Lévy Processes, 2023

Lévy-Itô Diffusion Models: parametric family of SDEs

- We propose to use the following parametric family of reverse SDEs to sample from LIMs:

$$d\bar{X}_t = (\mu(\bar{X}_{t+}, t) - (1 + \eta_t)\sigma_t^\alpha S_t^{(\alpha)}(\bar{X}_{t+}))dt + \sigma_t \eta_t^{1/\alpha} d\bar{L}_t^\alpha$$

Lévy-Itô Diffusion Models: parametric family of SDEs

- We propose to use the following parametric family of reverse SDEs to sample from LIMs:

$$d\bar{X}_t = (\mu(\bar{X}_{t+}, t) - (1 + \eta_t)\sigma_t^\alpha S_t^{(\alpha)}(\bar{X}_{t+}))dt + \sigma_t \eta_t^{1/\alpha} d\bar{L}_t^\alpha$$

- We prove that solutions of these SDEs have the same marginal densities as that of the forward SDE

Lévy-Itô Diffusion Models: parametric family of SDEs

- We propose to use the following parametric family of reverse SDEs to sample from LIMs:

$$d\bar{X}_t = (\mu(\bar{X}_{t+}, t) - (1 + \eta_t)\sigma_t^{\alpha} S_t^{(\alpha)}(\bar{X}_{t+}))dt + \sigma_t \eta_t^{1/\alpha} d\bar{L}_t^{\alpha}$$

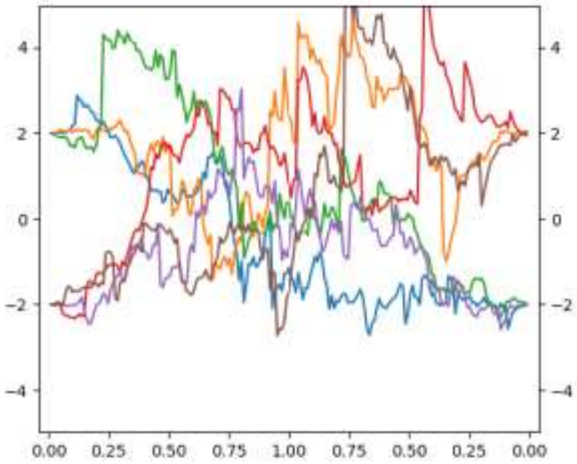
- We prove that solutions of these SDEs have the same marginal densities as that of the forward SDE
- These SDEs do not contain intractable terms, so allow for *exact* simulation

Lévy-Itô Diffusion Models: parametric family of SDEs

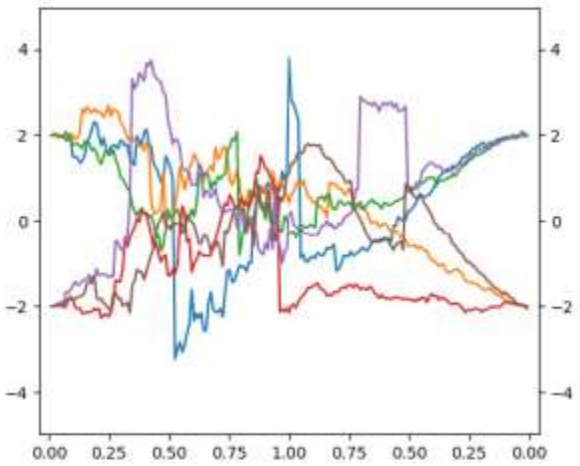
- We propose to use the following parametric family of reverse SDEs to sample from LIMs:

$$d\bar{X}_t = (\mu(\bar{X}_{t+}, t) - (1+\eta_t)\sigma_t^\alpha S_t^{(\alpha)}(\bar{X}_{t+}))dt + \sigma_t \eta_t^{1/\alpha} d\bar{L}_t^\alpha$$

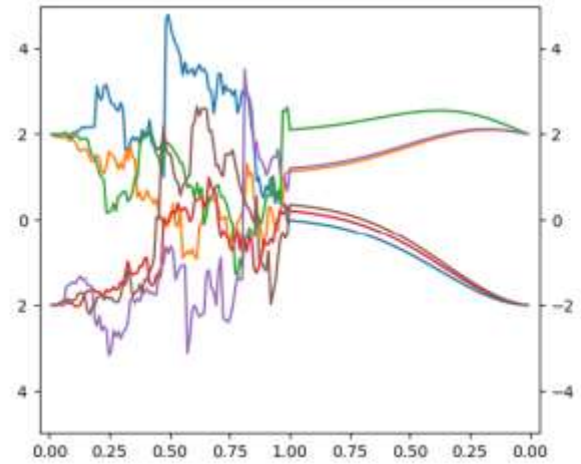
- We prove that solutions of these SDEs have the same marginal densities as that of the forward SDE
- These SDEs do not contain intractable terms, so allow for *exact* simulation
- Parameter η allows to control the amount of noise we want to add at inference



η=1



η=0.2



η=0

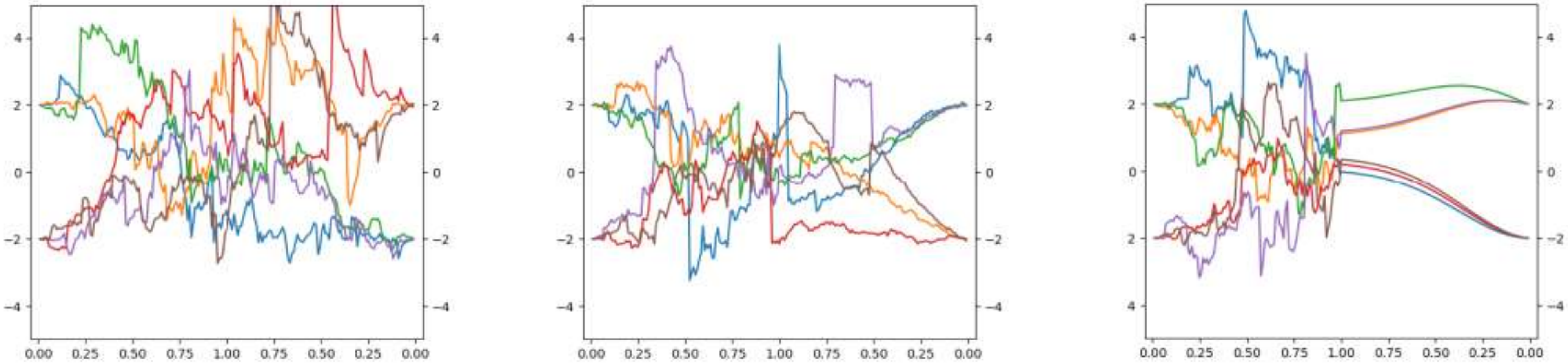
ICLR

Lévy-Itô Diffusion Models: parametric family of SDEs

- We propose to use the following parametric family of reverse SDEs to sample from LIMs:

$$d\bar{X}_t = (\mu(\bar{X}_{t+}, t) - (1 + \eta_t)\sigma_t^\alpha S_t^{(\alpha)}(\bar{X}_{t+}))dt + \sigma_t \eta_t^{1/\alpha} d\bar{L}_t^\alpha$$

- We prove that solutions of these SDEs have the same marginal densities as that of the forward SDE
- These SDEs do not contain intractable terms, so allow for *exact* simulation
- Parameter η allows to control the amount of noise we want to add at inference



η=1 η=0.2 η=0

- Parameter η has to be tuned at inference for optimal performance

Lévy-Itô Diffusion Models: results

- Unconditional image generation experiments on CIFAR10 demonstrate that for carefully chosen η we can obtain significant improvements in terms of FID compared to other known methods of sampling from LIMs:

	Euler-Maruyama		Exponential Integrator		
	N=20	N=50	N=20	N=50	N=500
SDE-A (LIM with $\alpha = 1.8$)	144.7	61.57	10.42	6.58	**2.64**
SDE-E (LIM with $\alpha = 1.8$)	**8.79**	**4.14**	**6.86**	**4.87**	3.36
ODE (LIM with $\alpha = 1.8$)	11.68	5.23	10.31	4.88	3.38
SDE-A (LIM with $\alpha = 1.5$)	109.5	22.68	9.05	4.25	**2.72**
SDE-E (LIM with $\alpha = 1.5$)	**7.86**	**4.37**	**6.27**	**4.09**	3.26
ODE (LIM with $\alpha = 1.5$)	9.95	5.94	10.50	5.14	3.35
SDE-A (LIM with $\alpha = 1.2$)	49.87	4.54	7.70	4.49	**3.28**
SDE-E (LIM with $\alpha = 1.2$)	**7.08**	**4.22**	**7.08**	**4.26**	3.74
ODE (LIM with $\alpha = 1.2$)	8.43	5.25	9.14	5.12	3.74

Lévy-Itô Diffusion Models: results

- Unconditional image generation experiments on CIFAR10 demonstrate that for carefully chosen η we can obtain significant improvements in terms of FID compared to other known methods of sampling from LIMs:

	Euler-Maruyama		Exponential Integrator		
	N=20	N=50	N=20	N=50	N=500
SDE-A (LIM with $\alpha = 1.8$)	144.7	61.57	10.42	6.58	**2.64**
SDE-E (LIM with $\alpha = 1.8$)	**8.79**	**4.14**	**6.86**	**4.87**	3.36
ODE (LIM with $\alpha = 1.8$)	11.68	5.23	10.31	4.88	3.38
SDE-A (LIM with $\alpha = 1.5$)	109.5	22.68	9.05	4.25	**2.72**
SDE-E (LIM with $\alpha = 1.5$)	**7.86**	**4.37**	**6.27**	**4.09**	3.26
ODE (LIM with $\alpha = 1.5$)	9.95	5.94	10.50	5.14	3.35
SDE-A (LIM with $\alpha = 1.2$)	49.87	4.54	7.70	4.49	**3.28**
SDE-E (LIM with $\alpha = 1.2$)	**7.08**	**4.22**	**7.08**	**4.26**	3.74
ODE (LIM with $\alpha = 1.2$)	8.43	5.25	9.14	5.12	3.74

- Additionally we have experiments in image generation and text-to-speech synthesis in imbalanced setting
- More details can be found in the paper

Thank you for your attention!