# Fractional Langevin Dynamics for Combinatorial Optimization via Polynomial-Time Escape

Shiyue Wang+, Ziao Guo+, Changhong Lu, Junchi Yan*


## Abstract

**Langevin dynamics (LD)** and its discrete proposal have been widely applied in the field of **Combinatorial Optimization (CO).** Both **sampling-based** and **data-driven** approaches have benefited significantly from these methods. However, LD's reliance on Gaussian noise limits its ability to escape narrow local optima, requires costly parallel chains, and performs poorly in rugged landscapes or with non-strict constraints. These challenges have impeded the development of more advanced approaches. To address these issues, we introduce **fractional Langevin dynamics (FLD)** for CO, replacing Gaussian noise with α-stable Lévy noise. FLD can escape from local optima more readily via Lévy flights, and in multiple-peak CO problems with high potential barriers it exhibits a **polynomial escape time** that outperforms the exponential escape time of LD. Moreover, FLD coincides with LD when α=2, and by tuning α it can be adapted to a wider range of complex scenarios in the CO field. We provide theoretical proof that our method offers enhanced exploration capabilities and improved convergence. Experimental results on the Maximum Independent Set, Maximum Clique, and Maximum Cut problems demonstrate that incorporating FLD advances both sampling-based and data-driven approaches, achieving **state-of-the-art (SOTA)** performance in most of the experiments.


## Contributions

- We incorporate **symmetric α-stable ($\mathcal{S}\alpha\mathcal{S}$) noise** with truncation into FLD: unlike Gaussian perturbations, enabling instantaneous energy-barrier leaps that facilitate escape from local minima.
- We propose the $\mathcal{S}\alpha\mathcal{S}$-noise FLD sampling process and present both **explicit- and implicit-gradient formulations** to advance both sampling-based and data-driven approaches.
- We adopt the mean escape time as our convergence metric, and derive theoretical upper bounds in the discrete setting, showing a **polynomial-time bound** for FLD versus an exponential bound for LD.
- Our methods **outperform** existing sampling-based and data-driven methods.

## Sampling

**Theorem 3.** *Let $\gamma$ be uniformly distributed on $(-\frac{\pi}{2}, \frac{\pi}{2})$ and $W$ be an independent exponential random variable with mean 1. The $\mathcal{S}\alpha\mathcal{S}$ sampling is:*

$$Z = \begin{cases} \frac{\sin \alpha\gamma}{(\cos\gamma)^{1/\alpha}} \left(\frac{\cos(\gamma - \alpha\gamma))}{W}\right)^{(1-\alpha)/\alpha} \triangleq \mathcal{S}\alpha\mathcal{S}(1),\ \alpha \neq 1 \\ \frac{\pi}{2}\tan\gamma \triangleq \mathcal{S}1\mathcal{S}(1),\ \alpha = 1 \end{cases} \tag{18}$$

$$x_{n+1} = x_n - \frac{\eta_{n+1} c_\alpha}{\tau} \nabla H(x) + \eta_{n+1}^{1/\alpha} z_{n+1} \tag{19}$$

## Convergency

LD:

$$\lambda_1 \geq \frac{(\Phi_2(\mathcal{B}))^2}{2}$$

FLD:

$$\lambda_1^\alpha \geq C_{N,\alpha}\Phi_\alpha(\mathcal{B}) = \frac{\alpha 2^{\alpha-1}\Gamma(\frac{N+\alpha}{2})}{\pi^{N/2}\Gamma(1-\frac{\alpha}{2})}\Phi_\alpha(\mathcal{B})$$

## Results

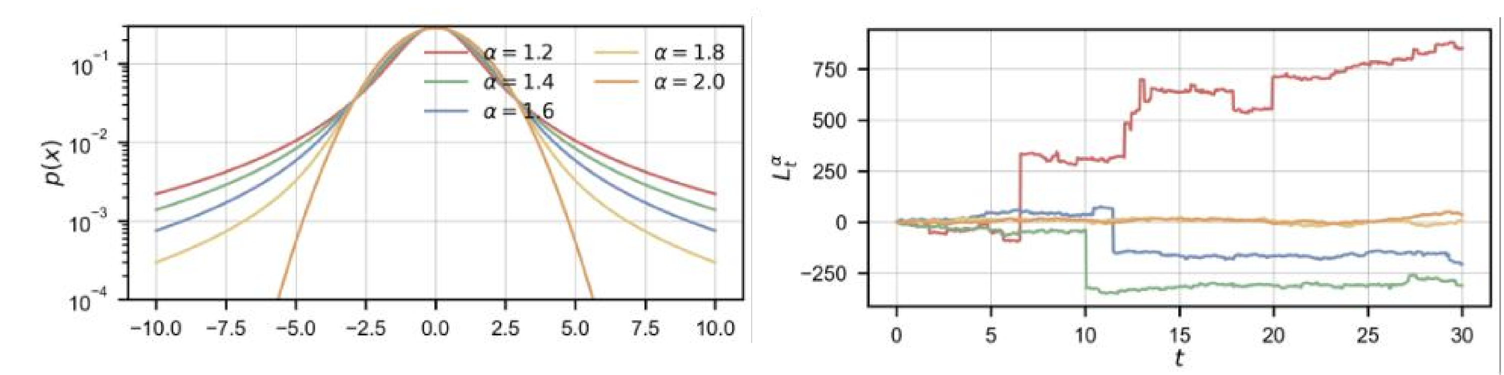
Table 1: Results of compared methods for MIS problem.

| MIS | | RB-[200–300] | | RB-[800–1200] | | ER-[700–800] | | ER-[9000–11000] | |
|---|---|---|---|---|---|---|---|---|---|
| Method | Type | Size ↑ | Time ↓ | Size ↑ | Time ↓ | Size ↑ | Time ↓ | Size ↑ | Time ↓ |
| Gurobi | OR | 19.98 | 47.57 m | 40.90 | 2.17 h | 41.38 | 50.00 m | — | — |
| KaMIS | OR | 20.10 | 1.40 h | 43.15 | 2.05 h | 44.87 | 52.13 m | 381.31 | 7.60 h |
| DGL | SL | 17.36 | 12.78 m | 34.50 | 23.90 m | 37.26 | 22.71 m | — | — |
| INTEL | SL | 18.47 | 13.07 m | 34.47 | 20.28 m | 34.86 | 6.06 m | 284.63 | 5.02 m |
| DIFUSCO | SL | 18.52 | 16.05 m | — | — | 41.12 | 26.67 m | — | — |
| LTFT | UL | 19.18 | 32 s | 37.48 | 4.37 m | — | — | — | — |
| DiffUCO | UL | 19.24 | 54 s | 38.87 | 4.95 m | — | — | — | — |
| SDDS | UL | 19.62 | 20 s | 39.99 | 6.35 m | — | — | — | — |
| PPO | RL | 19.01 | 1.28 m | 32.32 | 7.55 m | — | — | — | — |
| DIMES | RL | — | — | — | — | 42.06 | 12.01 m | 332.80 | 12.72 m |
| RLNN | PRL | 19.52 | 1.64 m | 38.46 | 6.24 m | 43.34 | 1.37 m | 363.34 | 11.76 m |
| iSCO | H | 19.29 | 2.71 m | 36.96 | 11.26 m | 42.18 | 1.45 m | 365.37 | 1.10 h |
| RLSA | H | 19.97 | 35 s | 40.19 | 1.85 m | 44.10 | 20 s | 375.31 | 1.66 m |
| FLD-IG | PRL | 19.72 | 1.08 m | 39.56 | 6.31 m | 43.50 | 1.35 m | 365.03 | 11.41 m |
| FLD-EG | H | **20.02** | 38 s | **40.25** | 1.93 m | **44.37** | 19 s | **377.50** | 1.12 m |

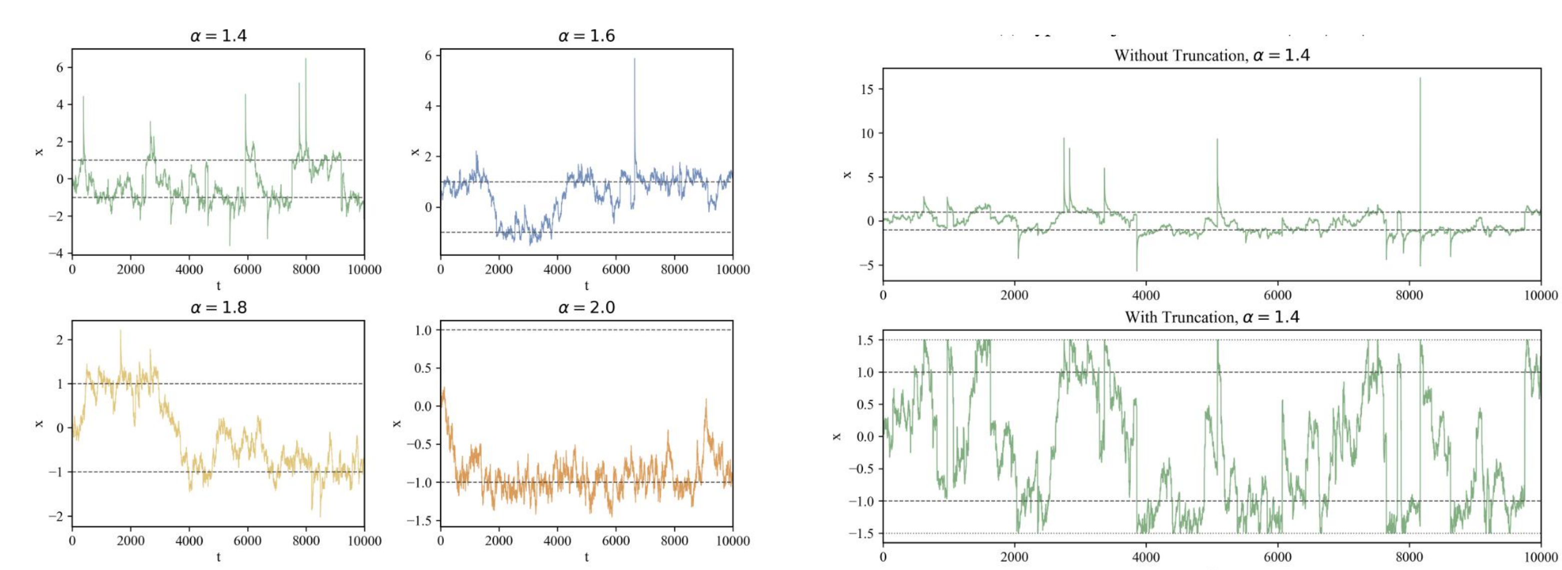
Table 2: Results of compared methods for MaxCl and MaxCut problems.

| MaxCl | | RB-[200–300] | | RB-[800–1200] | | MaxCut | | BA-[200–300] | | BA-[800–1200] | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Method | Type | Size ↑ | Time ↓ | Size ↑ | Time ↓ | Method | Type | Size ↑ | Time ↓ | Size ↑ | Time ↓ |
| Gurobi | OR | 19.05 | 1.92 m | 33.89 | 19.67 m | Gurobi | OR | 730.87 | 8.50 m | 2944.38 | 1.28 h |
| ERDOES | UL | 12.02 | 41 s | 25.43 | 2.27 m | ERDOES | UL | 693.45 | 46 s | 2870.34 | 2.82 m |
| LTFT | UL | 16.24 | 42 s | 31.42 | 4.83 m | LTFT | UL | 704.30 | 2.95 m | 2864.61 | 21.33 m |
| DiffUCO | UL | 16.22 | 1.00 m | — | — | DiffUCO | UL | 727.32 | 1.00 m | 2947.53 | 3.78 m |
| SDDS | UL | 18.90 | 38 s | — | — | SDDS | UL | 731.93 | 14 s | **2971.62** | 1.08 m |
| RLNN | PRL | 18.13 | 1.36 m | 35.23 | 7.83 m | RLNN | PRL | 729.00 | 1.58 m | 2907.18 | 3.67 m |
| Greedy | H | 13.53 | 25 s | 26.71 | 25 s | Greedy | H | 688.31 | 13 s | 2786.00 | 3.12 m |
| MFA | H | 14.82 | 27 s | 27.94 | 2.32 m | MFA | H | 704.03 | 1.60 m | 2833.86 | 7.27 m |
| iSCO | H | 18.96 | 54 s | 40.35 | 11.37 m | iSCO | H | 728.24 | 1.67 m | 2919.97 | 4.18 m |
| RLSA | H | **18.97** | 23 s | 40.53 | 1.27 m | RLSA | H | 733.54 | 27 s | 2955.81 | 1.45 m |
| FLD-IG | PRL | 18.52 | 1.40 m | 37.40 | 6.89 m | FLD-IG | PRL | 733.48 | 1.57 m | 2922.54 | 3.07 m |
| FLD-EG | H | **18.97** | 20 s | **40.63** | 1.91 m | FLD-EG | H | **734.18** | 25 s | 2960.13 | 1.70 m |

## Pdf of $\mathcal{S}\alpha\mathcal{S}$ Distribution



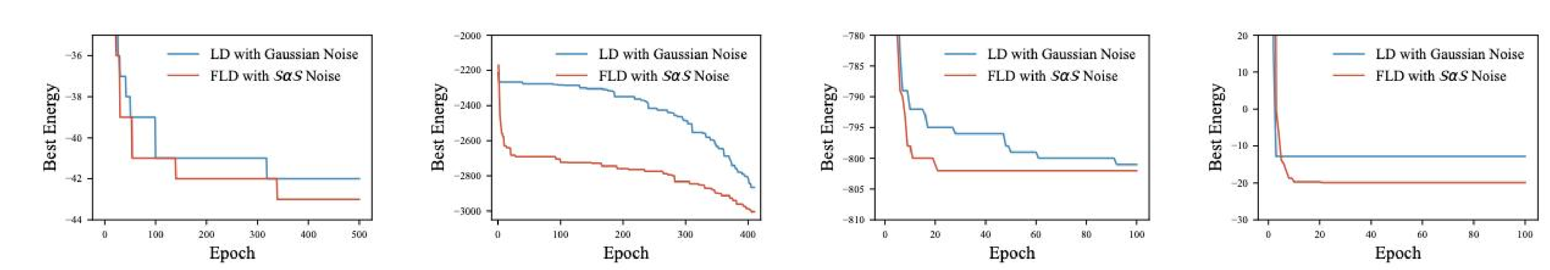

## Ablation Study




(a) Typical trajectories for α = 1.4, 1.6, 1.8, 2.0.

(b) Trajectories w/ and w/o truncation for α = 1.4.

Figure 2: Sampling trajectories of the FLD-based SDE.




(a) ER-[700-800]. (b) BA-[800-1200]. (c) BA-[200-300]. (d) RB-[200-300], MaxCl.

Figure 3: Ablation for our methods (FLD with $\mathcal{S}\alpha\mathcal{S}$) and LD sampling process method with Gaussian noise. The staircase curves show how the "Best Energy" evolves as the number of iterations ("epochs"), where the "Best Energy" is the minimum energy function value between last and current epoch.