复旦大学

Convex Potential Mirror Langevin Algorithm for Efficient Sampling of Energy-Based Models

Zitao Yang[1], Amin Ullah[2], Shuai Li[3], Fuxin Li[4], Jun Li[1]

[1]Fudan University [2]Boeing Research and Technology
[3]University of Bonn [4]Oregon State University

yangzt21@m.fudan.edu.cn

NeurIPS 2025

Outline

1 Introduction
2 Background
3 Proposed Method: CPMLA
4 Theoretical Analysis
5 Experiments
6 Conclusion

Motivation: Challenges in EBM Sampling

- Energy-Based Models (EBMs) offer flexibility in modeling complex distributions.
- Applications: 3D recognition, image restoration, protein folding.

Motivation: Challenges in EBM Sampling

- Energy-Based Models (EBMs) offer flexibility in modeling complex distributions.
- Applications: 3D recognition, image restoration, protein folding.
- **Challenge:** Sampling relies on MCMC (e.g., Langevin Dynamics), which is often slow and can get stuck, especially in high dimensions or with complex energy landscapes.
- Existing methods to improve sampling still face issues like non-mixing.

Motivation: Challenges in EBM Sampling

- Energy-Based Models (EBMs) offer flexibility in modeling complex distributions.
- Applications: 3D recognition, image restoration, protein folding.
- **Challenge:** Sampling relies on MCMC (e.g., Langevin Dynamics), which is often slow and can get stuck, especially in high dimensions or with complex energy landscapes.
- Existing methods to improve sampling still face issues like non-mixing.
- **Goal:** Develop a more efficient and reliable sampling algorithm for EBMs.

Outline

1 Introduction

2 Background

3 Proposed Method: CPMLA

4 Theoretical Analysis

5 Experiments

6 Conclusion

Background: EBMs and Mirror Langevin Dynamics

Energy-Based Models (EBMs)

- Define probability density via an energy function $f_\theta(x)$: $p_\theta(x) = \frac{1}{Z(\theta)} \exp(f_\theta(x))$.
- Training involves MCMC sampling (e.g., Langevin) to approximate gradients.

Background: EBMs and Mirror Langevin Dynamics

Energy-Based Models (EBMs)

- Define probability density via an energy function $f_\theta(x)$: $p_\theta(x) = \frac{1}{Z(\theta)} \exp(f_\theta(x))$.
- Training involves MCMC sampling (e.g., Langevin) to approximate gradients.

Mirror Langevin Dynamics (MLD)

- Generalizes Langevin dynamics to non-Euclidean geometries using a *mirror map* ∇G derived from a convex potential G.
- Can accelerate convergence and adapt to data geometry.
- **Limitation:** Conventional MLD uses fixed mirror maps, which struggle with complex data manifolds.

Outline

1 Introduction

2 Background

3 Proposed Method: CPMLA

4 Theoretical Analysis

5 Experiments

6 Conclusion

CPMLA: Convex Potential Mirror Langevin Algorithm

Key Idea: Use a *dynamic*, learnable mirror map within MLD for EBM sampling.

- Mirror Map ∇G_{ϑ}: Implemented using Convex Potential Flow (CP-Flow).
- CP-Flow uses Input Convex Neural Networks (ICNNs) to guarantee convexity.
- Learns optimal transport map, capturing data geometry.

CPMLA: Convex Potential Mirror Langevin Algorithm

Key Idea: Use a *dynamic*, learnable mirror map within MLD for EBM sampling.

- Mirror Map ∇G_{ϑ}: Implemented using Convex Potential Flow (CP-Flow).
- CP-Flow uses Input Convex Neural Networks (ICNNs) to guarantee convexity.
- Learns optimal transport map, capturing data geometry.
- **Cooperative Learning:** Jointly train the EBM (f_{θ}) and the dynamic mirror map (∇G_{ϑ}).
  - EBM learns energy function using samples from CPMLA.
  - CP-Flow learns geometry using real data and EBM samples.
  - Creates a virtuous cycle of improvement.

CPMLA: Algorithm Overview

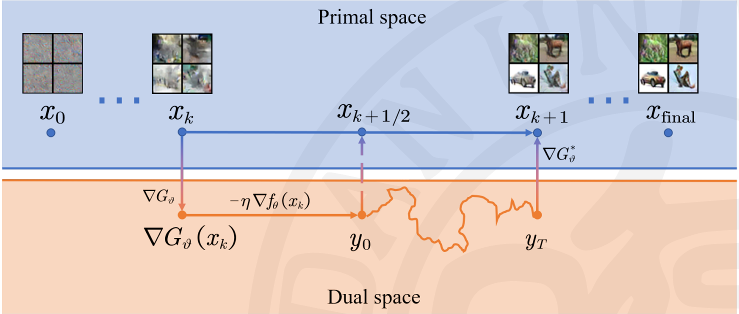


Figure: Overview of CPMLA sampling: Alternating updates between primal (image) and dual (geometry) spaces using the dynamic mirror map ∇G_ϑ and EBM gradients ∇f_θ.

Outline

1 Introduction

2 Background

3 Proposed Method: CPMLA

4 Theoretical Analysis

5 Experiments

6 Conclusion

Theoretical Analysis: Convergence Guarantee

We provide the *first* convergence analysis for MLD with deep neural network mirror maps.

Theorem (Convergence of CPMLA (Informal))

Under standard assumptions (Mirror LSI, Self-Concordance, Relative Lipschitz/Smoothness), CPMLA achieves:

- ***Exponential Convergence:*** *The distribution of samples ρ_t converges exponentially fast to the target distribution p_{data} in Total Variation distance.*
- ***Vanishing Bias:*** *The bias approaches zero as the step size $h \to 0$.*
- *Accounts for approximation errors from both the EBM (δ_2) and CP-Flow (δ_3) networks.*

Supports the efficiency and reliability of CPMLA in practice.

Outline

1 Introduction

2 Background

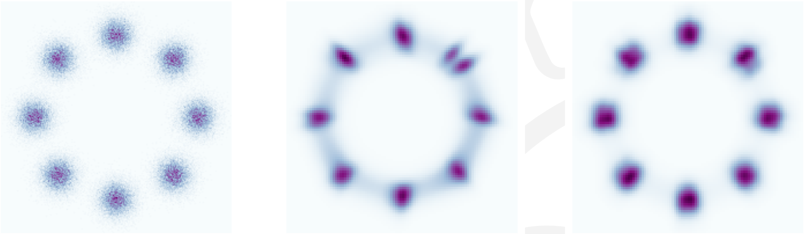
3 Proposed Method: CPMLA

4 Theoretical Analysis

5 Experiments

6 Conclusion

Experiments: Toy Example - Eight Gaussians



(a) Data (b) CP-Flow (10 it.) (c) CPMLA (3 it.)

Figure: CPMLA efficiently captures the target distribution, matching CP-Flow's quality in fewer iterations.

Demonstrates faster convergence on synthetic data.

Experiments: Image Generation Quality (CIFAR-10)

Evaluated on CIFAR-10, SVHN, CelebA (32x32).

Model type	**Models**	**FID**$\downarrow$
Flow+EBM	EBM-FCE	37.30
	CoopFlow (T=30)	21.16
CPMLA (Ours)	**CPMLAprt (T=20)**	**20.85**
	CPMLA (T=30)	21.09
Score-Based	NCSN++	2.20
GAN	StyleGAN2-ADA	2.92

Note: Lower FID is better. CPMLA outperforms prior EBM+Flow methods. Comparison to other model classes (Score, GAN) shows EBMs becoming more competitive.

(a) CIFAR-10 (b) SVHN (c) CelebA

Experiments: Efficiency vs. CoopFlow (CIFAR-10)

CPMLA demonstrates superior efficiency compared to CoopFlow:

Inference Speed & Quality:

- Achieves lower FID (20.85 vs 21.16).
- Uses fewer MCMC steps (T=20 vs T=30).
- Requires less wall-clock time (15.92s vs 16.84s / 1k images).

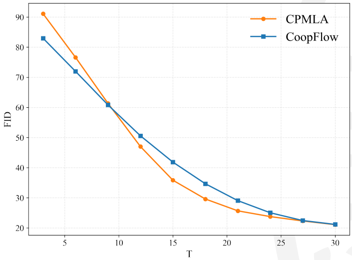


Figure: Faster FID improvement per step.

Parameter Efficiency:

- EBM part size identical (17.13M).
- **Flow part:** CPMLA **0.26M** vs. CoopFlow **28.78M**.
- Better results with ∼**0.9%** flow params.

Training Efficiency:

- Faster overall training due to memory efficiency (larger batches possible).
- **10.5h** vs 38h est. (CIFAR-10, 24GB VRAM).

Experiments: Image Inpainting (CelebA)

Demonstrates applicability to conditional generation tasks.

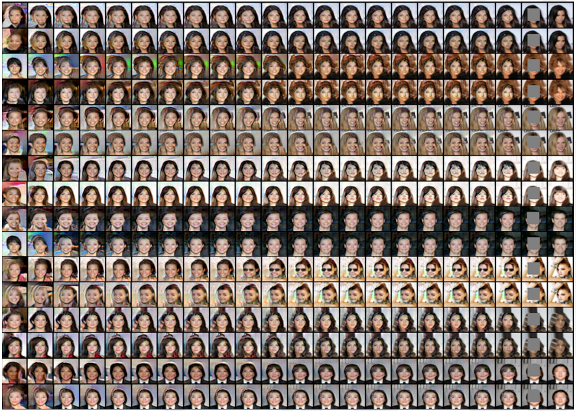


Figure: CPMLA successfully inpaints masked regions on CelebA images over optimization iterations (left to right), compared to masked and original images (last two columns).

Outline

1 Introduction

2 Background

3 Proposed Method: CPMLA

4 Theoretical Analysis

5 Experiments

6 Conclusion

Conclusion

- Introduced **CPMLA**, a novel algorithm designed for **efficient EBM sampling** using a dynamic mirror map based on CP-Flow.
- **Key Features for Efficiency:**
  - Adapts sampling geometry to the data manifold, potentially speeding up exploration.
  - Achieves exponential convergence with vanishing bias (theoretically proven for DNNs), supporting its reliability.
- **Results:** Demonstrated significant improvements in sampling efficiency (time, parameters, training throughput) compared to related EBM methods on benchmark datasets. Showcased strong performance in image generation, reconstruction, and inpainting.
- CPMLA offers a principled and highly performant approach to EBM sampling, enhancing the practicality of EBMs for complex generative tasks.

Thank You

For further questions or discussion, please contact:

yangzt21@m.fudan.edu.cn