NEURAL INFORMATION PROCESSING SYSTEMS

Berkeley UNIVERSITY OF CALIFORNIA

Immiscible Diffusion

Accelerating Diffusion Training with Noise Assignment

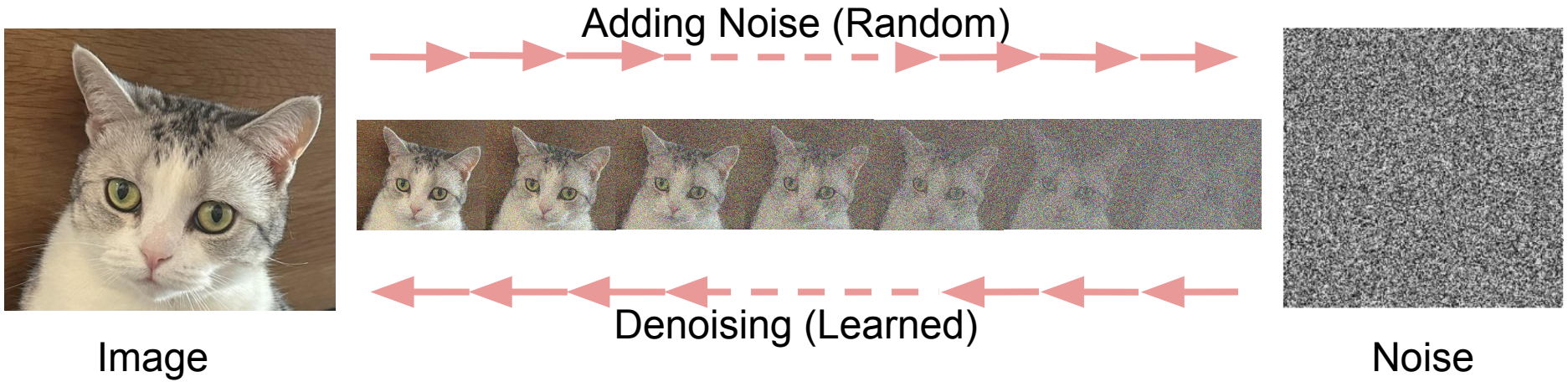
Yiheng Li, Heyang Jiang, Akio Kodaira
Masayoshi Tomizuka, Kurt Keutzer, Chenfeng Xu

2024/12

BAIR
BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH

MSC laboratory

Learning Step-by-step **Denoising** for Image Generation



Learning Step-by-step **Distribution Transfer** for Image Generation

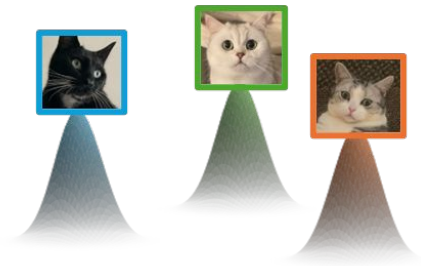
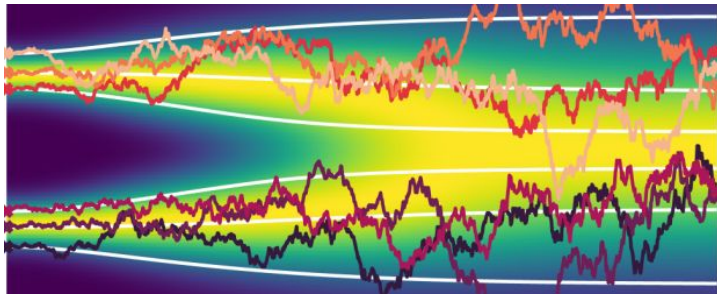
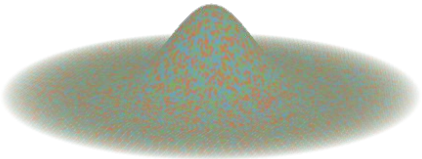


Image Distribution (One-hots)



Adding Noise

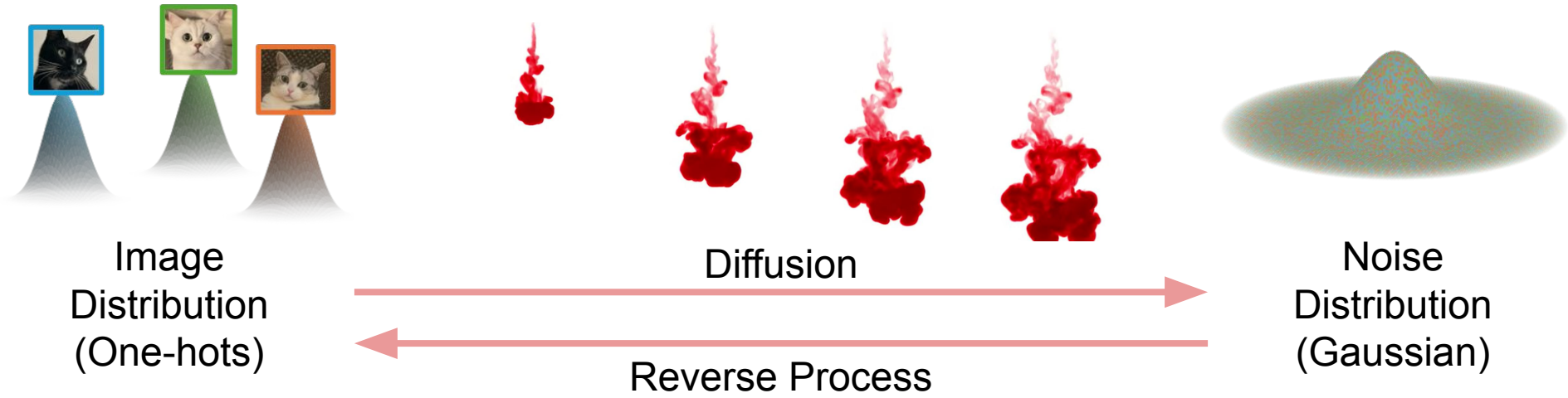


Noise Distribution (Gaussian)

Denoising

https://yang-song.net/blog/2021/score/

Physics Understanding: Data (Particle) Diffusion & Reverse Process

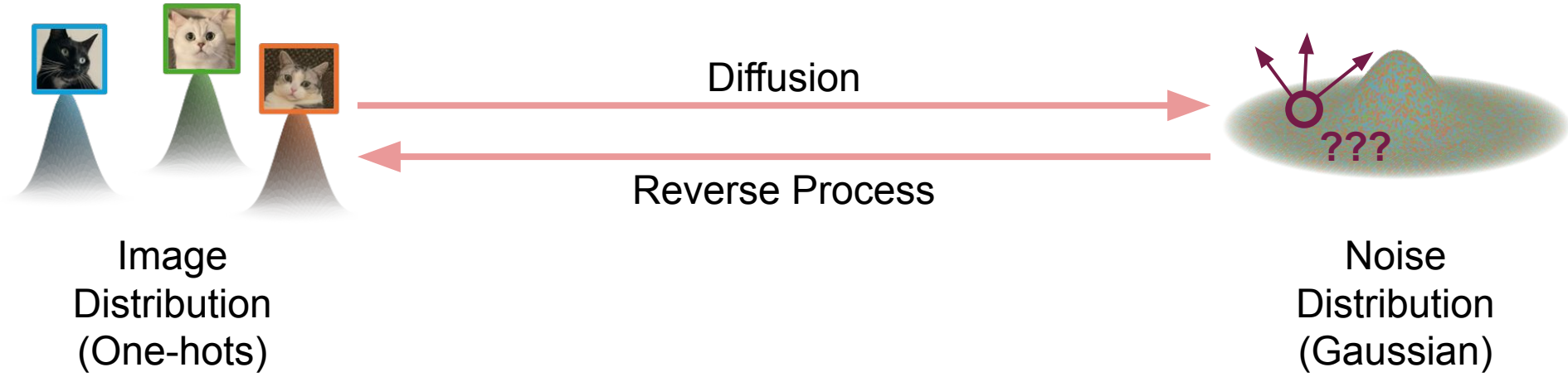


https://yang-song.net/blog/2021/score/

Miscibility

Mixing during Diffusion in Physics

Miscibility Challenges the De-noising



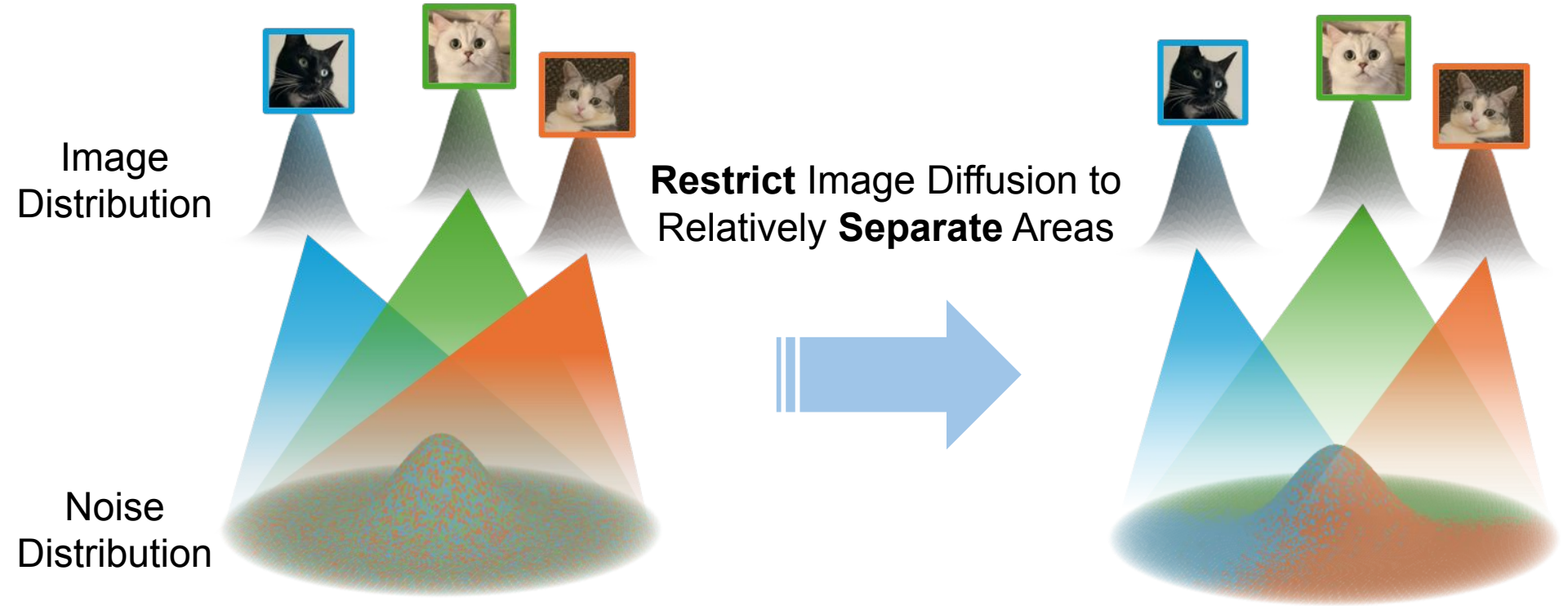
The last diffusion step provides **trivial information**.

Immiscibility

Restricted Diffusion of Each Solvent

Caused by Intermolecular Forces

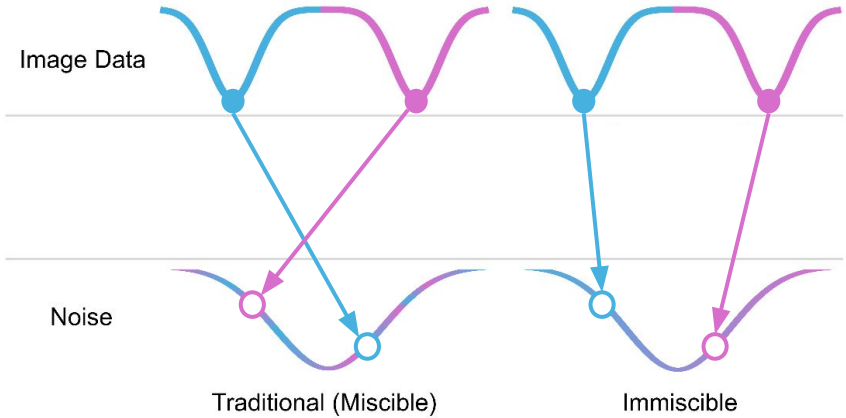
Immiscible Diffusion



Immiscible Diffusion Implementation: **Assignment**

Method

Assign Noise to Nearby Image



Performance

Efficient Execution & Little Image-Noise Average Distance Change

Batch Size	128	256	512	1024
Execution Time (ms)	5.4	6.7	8.8	22.8
ΔAve. Dist. (image, noise)	-1.93%	-2.16%	-2.32%	-2.44%

Immiscible Diffusion Solves the Denoising Challenge!

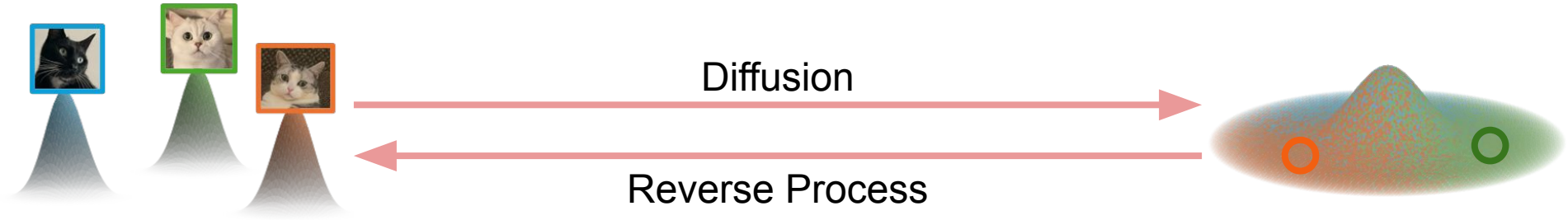
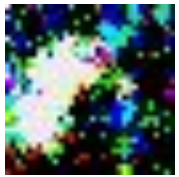


Image Distribution (One-hots)

Assigned Noise Distribution (Gaussian)

Then **the last diffusion step** provides **clear denoising results**!

Denoising Direction



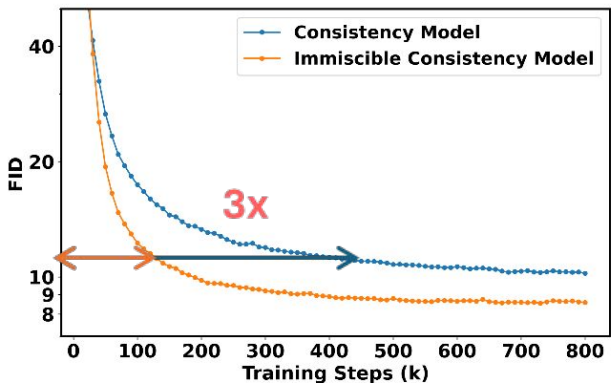
Predicted Image

True Image

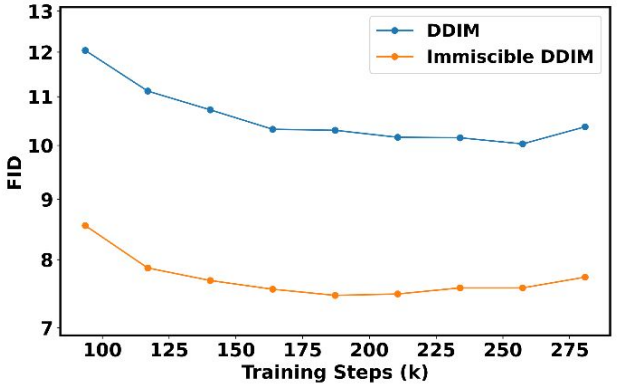
Training Efficiency Improvements

Unconditional Generation

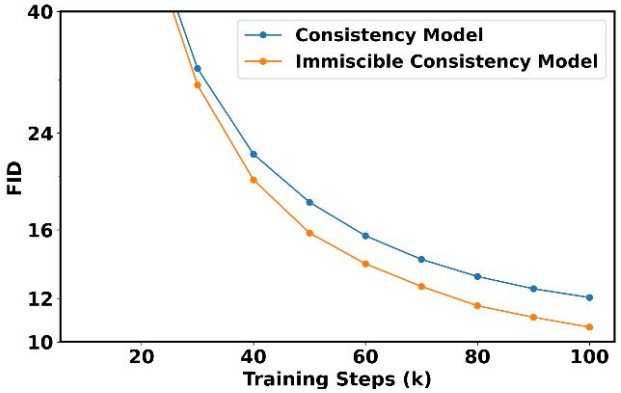
Consistency Model + CIFAR10



DDIM + CIFAR10



Consistency Model+ CelebA



Unconditional Generation

Stable Diffusion + ImageNet

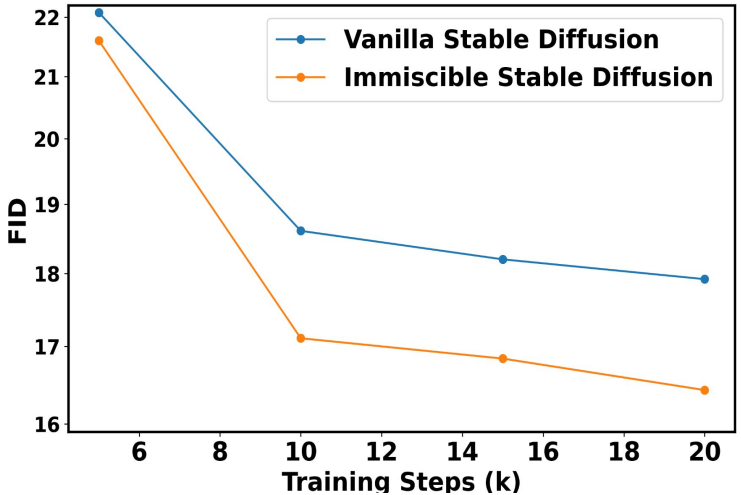
Vanilla (Baseline)

Immiscible

Training Efficiency Improvements

Class-Conditional Generation

Stable Diffusion + ImageNet



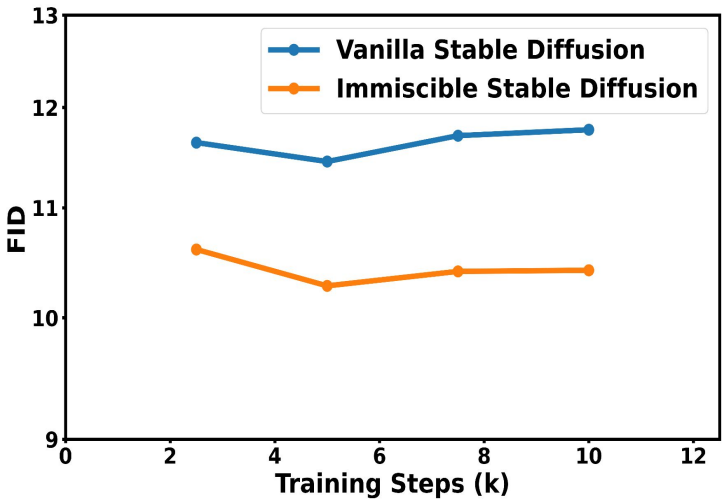
Vanilla

Immiscible

Training Efficiency Improvements

Class-Conditional **Fine-tuning**

Stable Diffusion v1.4 + ImageNet



Vanilla

Immiscible

Summary

1 line

of code:
image-noise assignment*

*Only one immiscible diffusion method;
Excluding Image Normalization for Some Baselines

One Line of Code
& Running Efficiently

```
noise_immiscible =
linear_sum_assignment
(image, noise)
```

6.7 ms*

Execution Time

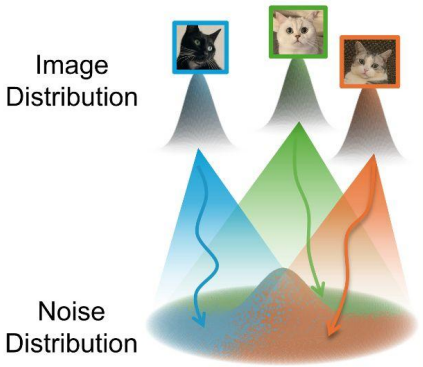
*For Batch Sizes = 256

2 %

image-noise data point
distance reduction

For Batch Sizes in [128, 1024]

Assignments of Corresponding
Images in Gaussian Noise Space

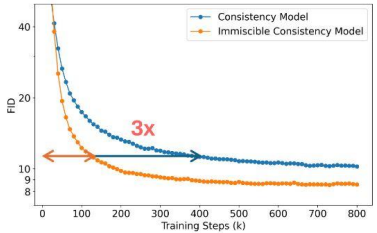


3 x

training efficiency
enhancement

On Consistency Model + CIFAR Dataset
On unconditional / conditional generation & fine-tuning

Effectiveness Observed
Both in FID and in Image Comparison



GitHub Site

Thank you!

NEURAL INFORMATION PROCESSING SYSTEMS

Berkeley
UNIVERSITY OF CALIFORNIA