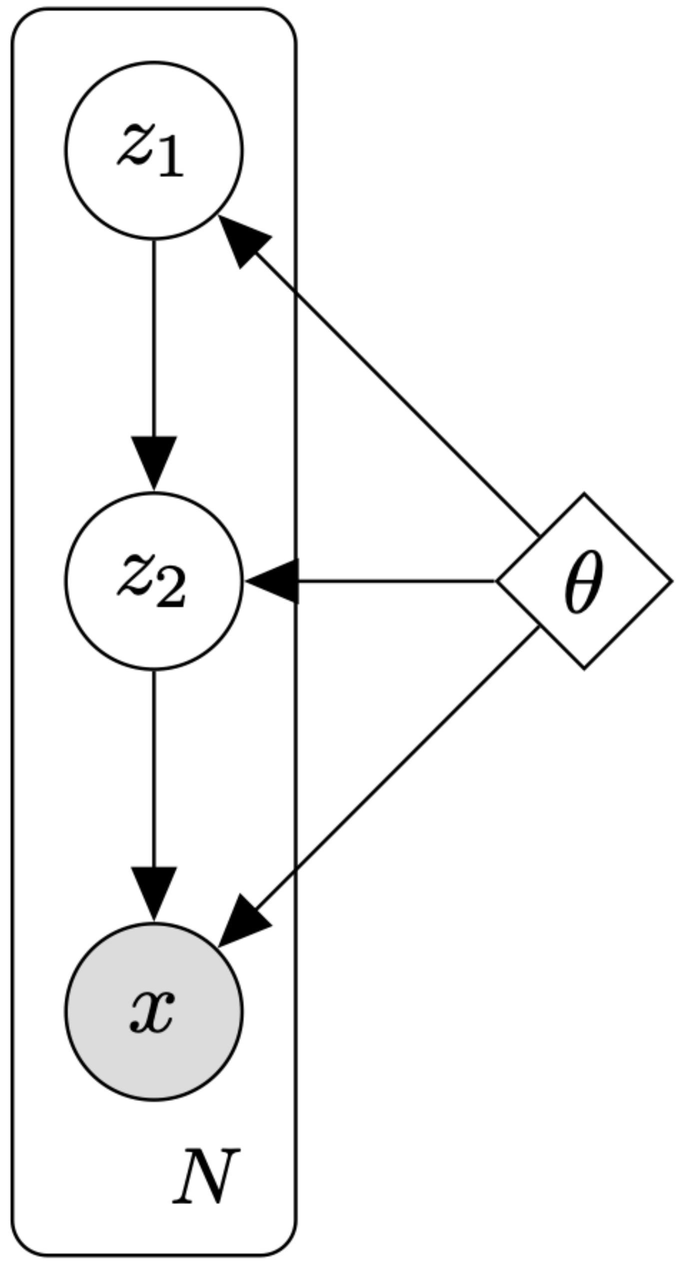
Neural Information Processing Systems 2024

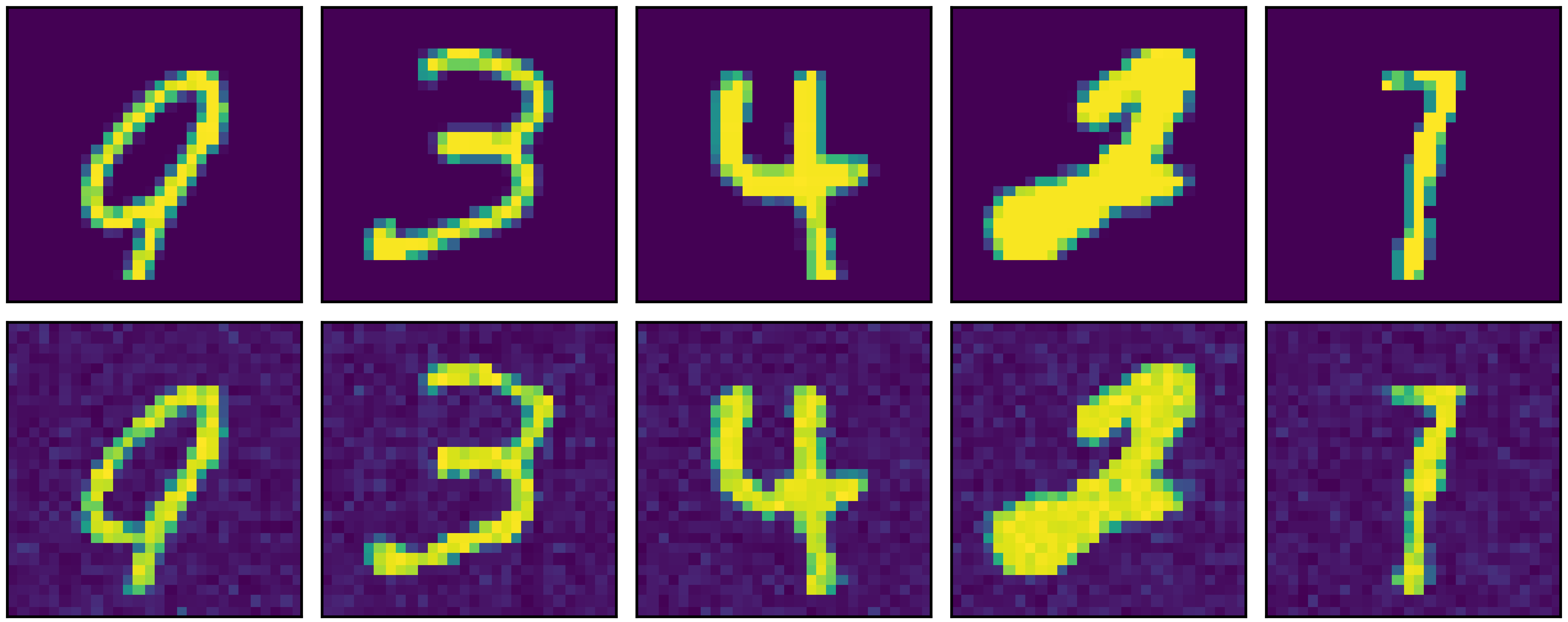
Divide-and-Conquer Predictive Coding

ELI SENNESH, HAO WU, TOMMASO SALVATORI; DECEMBER 2024

Predictive Coding: from Neuro to AI

- In NeuroAI, **predictive coding (PC)** provides
  - Biologically plausible credit-assignment in
  - Approximate *Bayesian inference*.
- But PC has trouble competing with ordinary deep learning on many tasks
- Our **Divide & Conquer PC (DCPC)** scales to compete with VAEs on structured problems.





DCPC trains hierarchical models

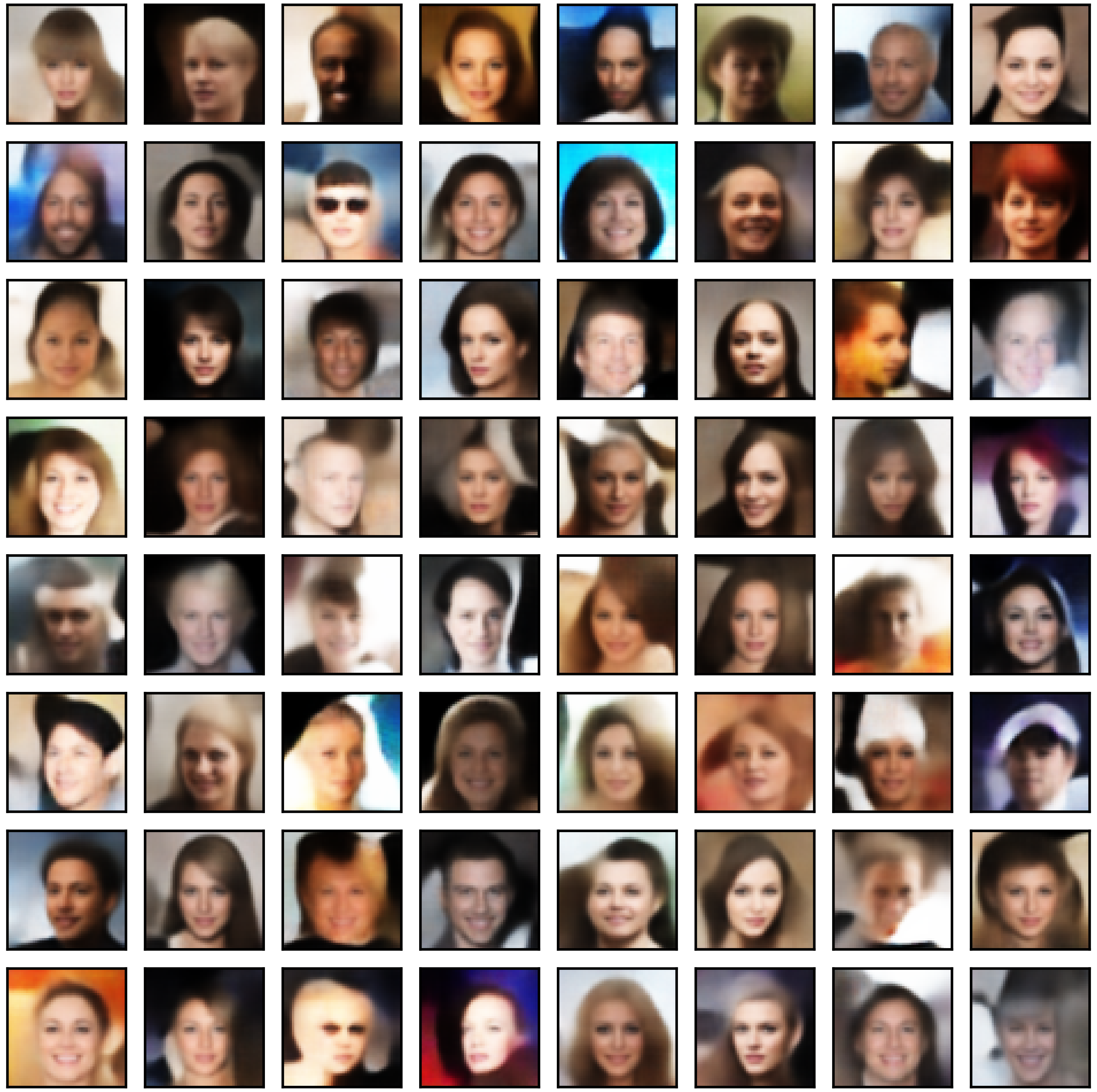
Inference algorithm	Dataset	NLL ↓	Mean Squared Error ↓
MCPC	MNIST	144.6 ± 0.7	$(8.29 \pm 0.05) \times 10^{-2}$
DCPC	MNIST	$\mathbf{102.5 \pm 0.01}$	$\mathbf{0.01 \pm 7.2 \times 10^{-6}}$
DCPC	EMNIST	160.8 ± 0.05	$3.3 \times 10^{-6} \pm 3.5 \times 10^{-9}$
DCPC	Fashion MNIST	284.1 ± 0.05	$0.03 \pm 2.7 \times 10^{-5}$

Table 2: Negative log-likelihood and mean squared error for MCPC against DCPC on held-out images from the MNISTs. Means and standard deviations are taken across five random seeds.

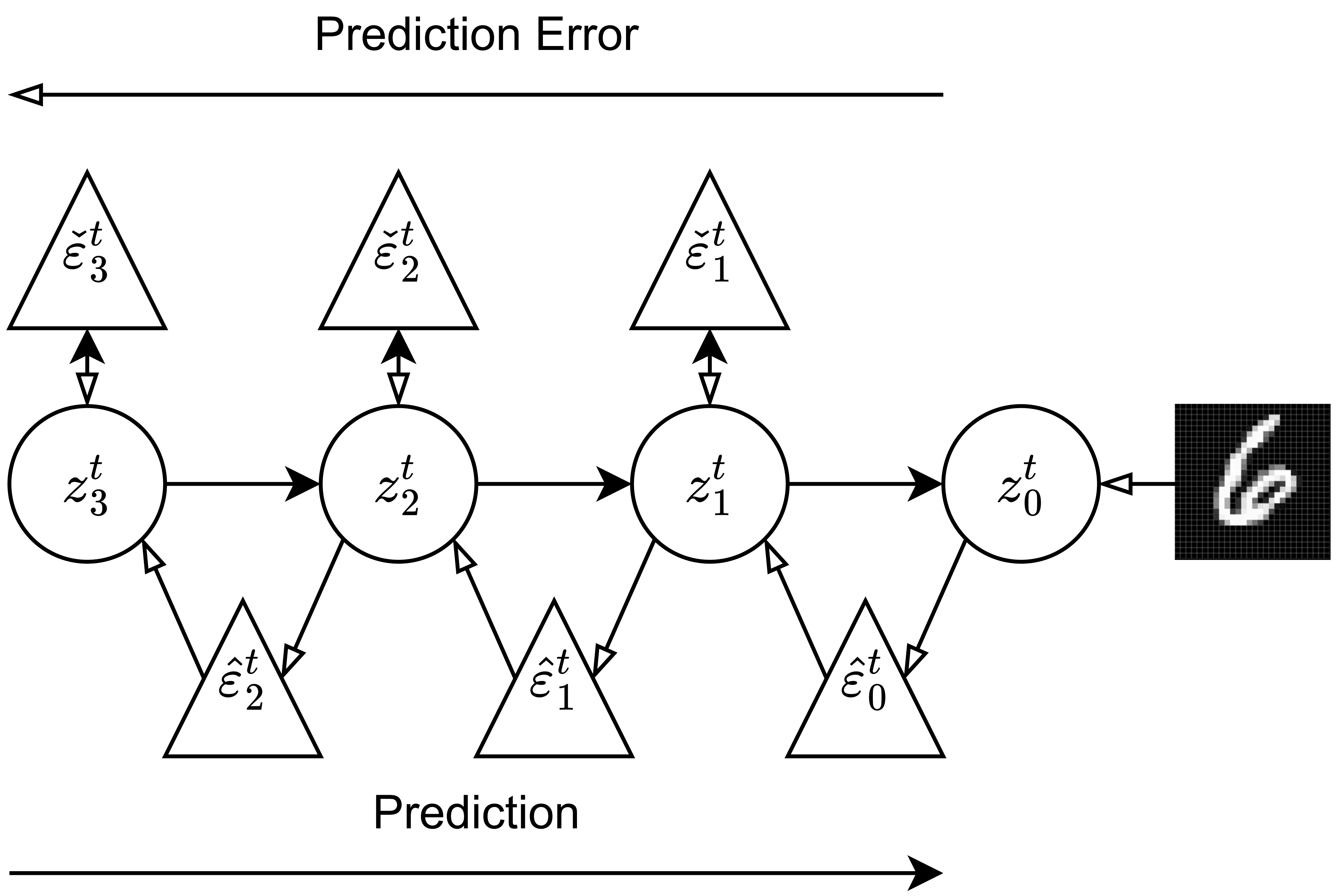
DCPC outperforms particle and PC methods on FID

Algorithm	Likelihood	Resolution ↑	$S\times$ Epochs ↓	FID ↓
PGD	$\mathcal{N}$	32×32	1×100	100 ± 2.7
DCPC (ours)	$\mathcal{N}$	32×32	1×100	$\mathbf{82.7 \pm 0.9}$
LPC	$\mathcal{DN}$	64×64	$300 \times 15 = 4500$	120 (approximate)
VAE	$\mathcal{DN}$	64×64	$1 \times 4500 = 4500$	86.3 ± 0.03
DCPC (ours)	$\mathcal{DN}$	64×64	$30 \times 150 = 4500$	$\mathbf{79.0} \pm 0.9$

Table 3: FID score comparisons on the CelebA dataset [Liu et al., 2015]. The score for LPC comes from Figure 2 in Zahid et al. [2024], where they ablated warm-starts and initialized from the prior.



How It Works



- Ancestor sampling initializes *predictions* for all latent variables $\mathbf{z}$.
- During inference, *prediction errors* ε_z measure the mismatch between predictions $z \in \mathbf{z}$ and the conditional $p(z \mid \mathbf{z}_{\setminus z})$ given the Markov blanket.
- **Langevin dynamics** and **Sequential Monte Carlo** let us sample updated predictions $z' \sim p(z \mid \mathbf{z}_{\setminus z})$.

Summary

- **Predictive coding** enjoys biological plausibility for Bayesian inference, a hard task.
- But it had trouble scaling to structured problems with correlated posteriors.
- **Divide & Conquer Predictive Coding (DCPC)** scales PC to substitute for amortized encoders
- Come see us at **Poster Session 6 (Fri 13 Dec 4:30 p.m. PST — 7:30 p.m. PST)** to talk about the methods and future extensions!

Many thanks to our supporters!

BASTOS LAB

VANDERBILT UNIVERSITY

VERSES

TU WIEN