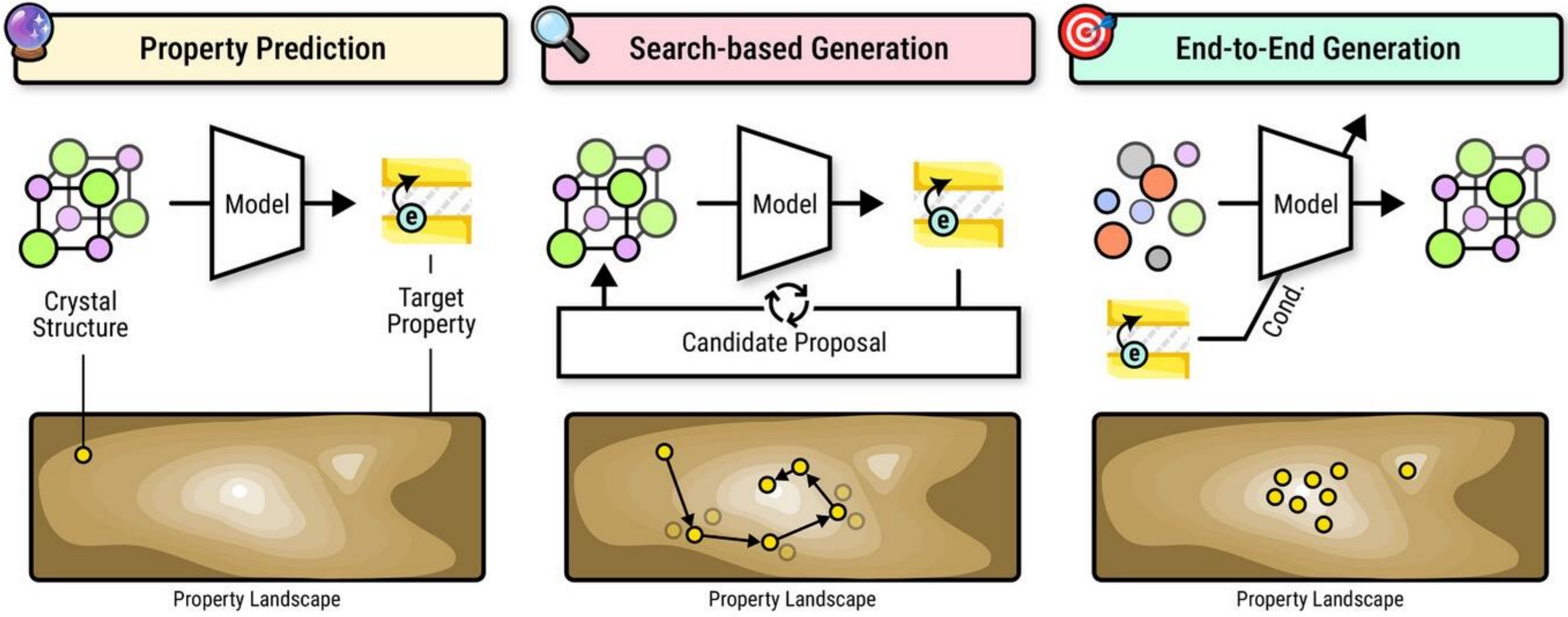
Cooking Up Materials with Generative Modelling

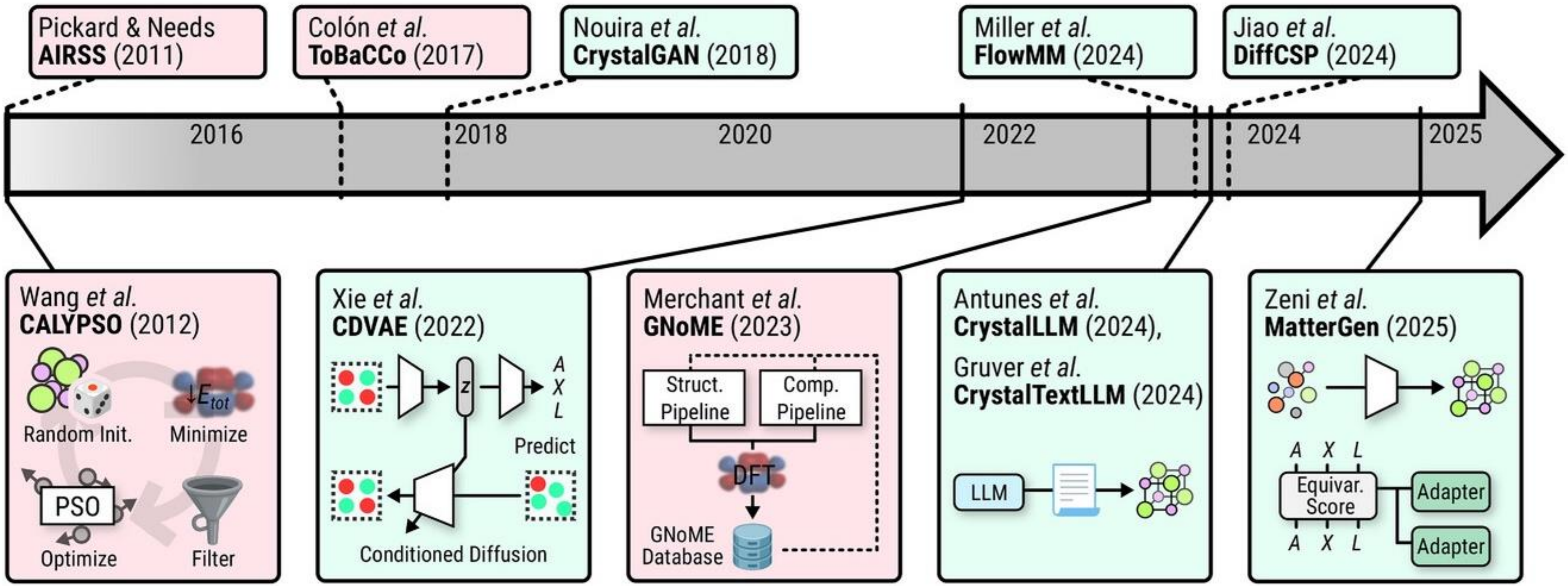
Presented by - Nikhil Singh
Research Scholar
Department of Chemistry
Indian Institute of Technology Delhi

ML models for Crytal Structures



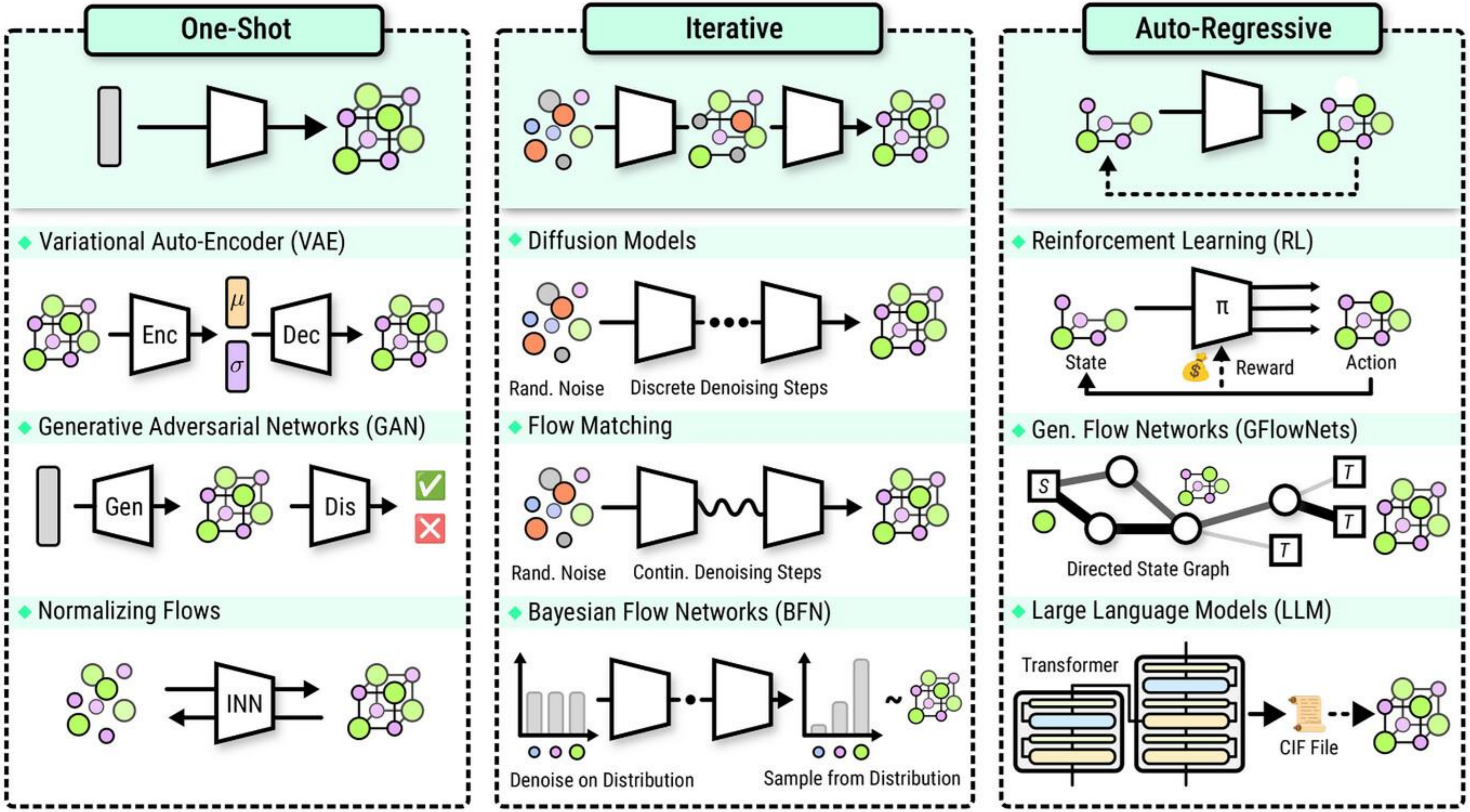
Metni et al., Advanced Materials (2026)

What has been done till now?



Metni et al., Advanced Materials (2026)

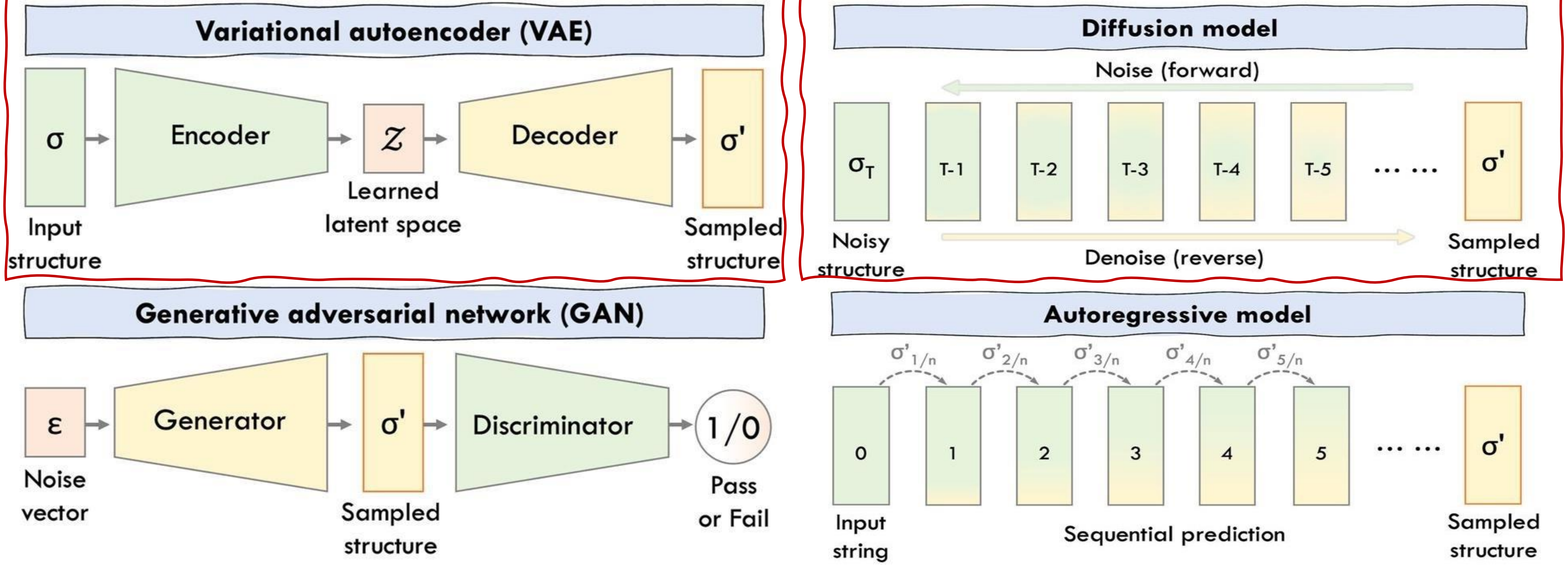
Different approaches



Metni et al., Advanced Materials (2026)

Advancements in generative models

Classic | **Progressive noise addition/removal**



Generator confuses discriminator with synthetic data

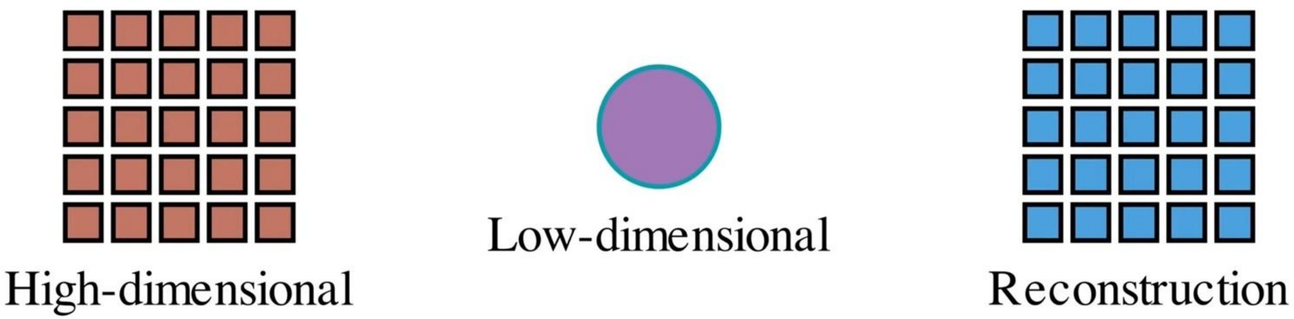
- **Beaten by diffusion models ☹**

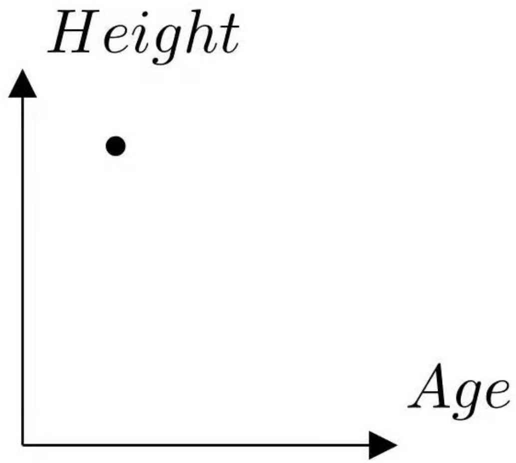
Sequential probability (language models)

Park et al., Matter 7, 2355-2367 (2024)

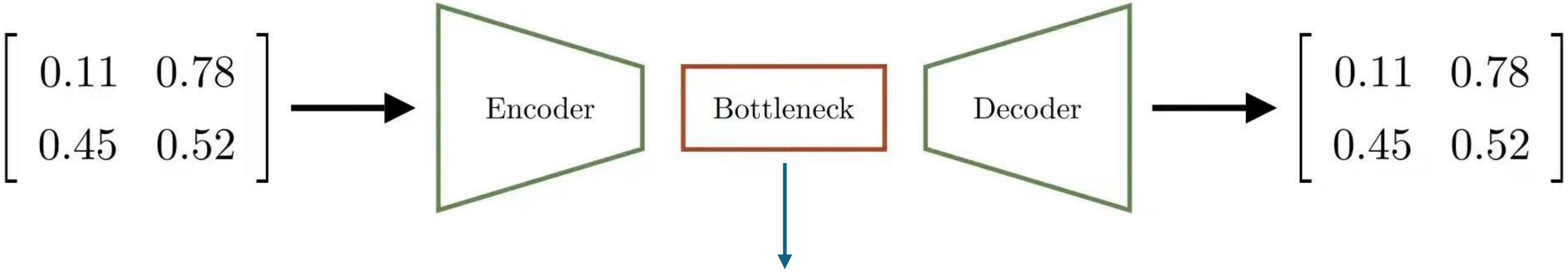
Autoencoders

Making the data representation compact

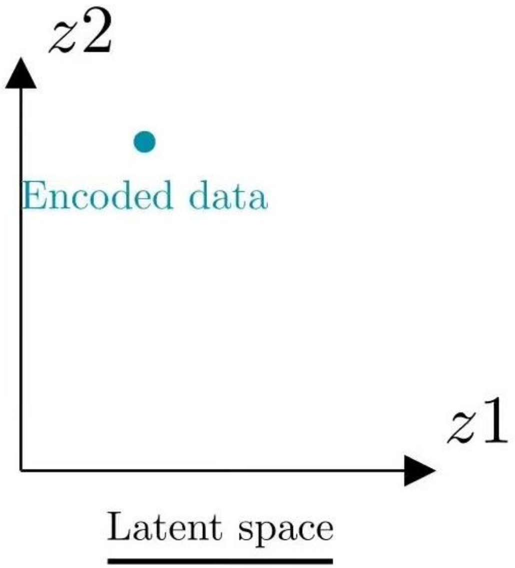




Basic Architecture



Also called 'Latent Space'



Youtube: https://youtu.be/qJeaCHQ1k2w?si=iDcfskIoGJMx4dvq (DeepIA)

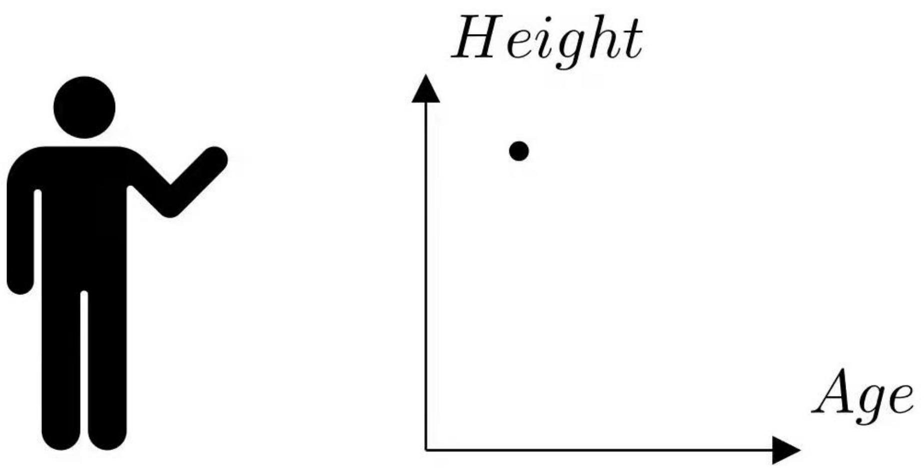
Autoencoders

Making the data representation compact

Data

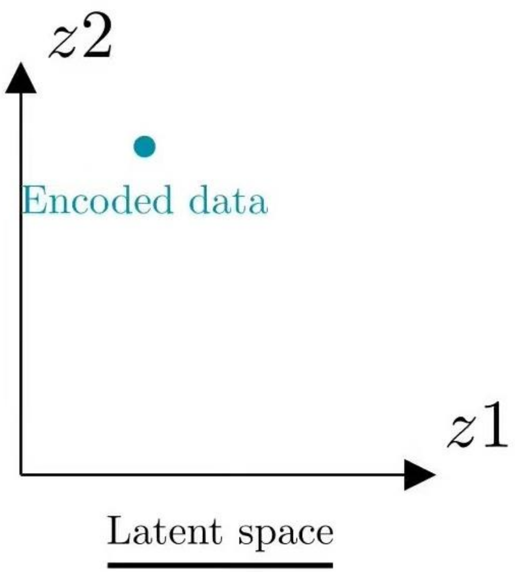
$$\begin{bmatrix} 66 & 68 & 97 & 67 & 79 \\ 66 & 27 & 89 & 29 & 24 \\ 37 & 92 & 44 & 60 & 54 \\ 29 & 53 & 87 & 74 & 17 \\ 34 & 2 & 6 & 99 & 63 \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix} \longrightarrow \begin{bmatrix} 10 \\ 22 \\ 34 \end{bmatrix}$$

Features



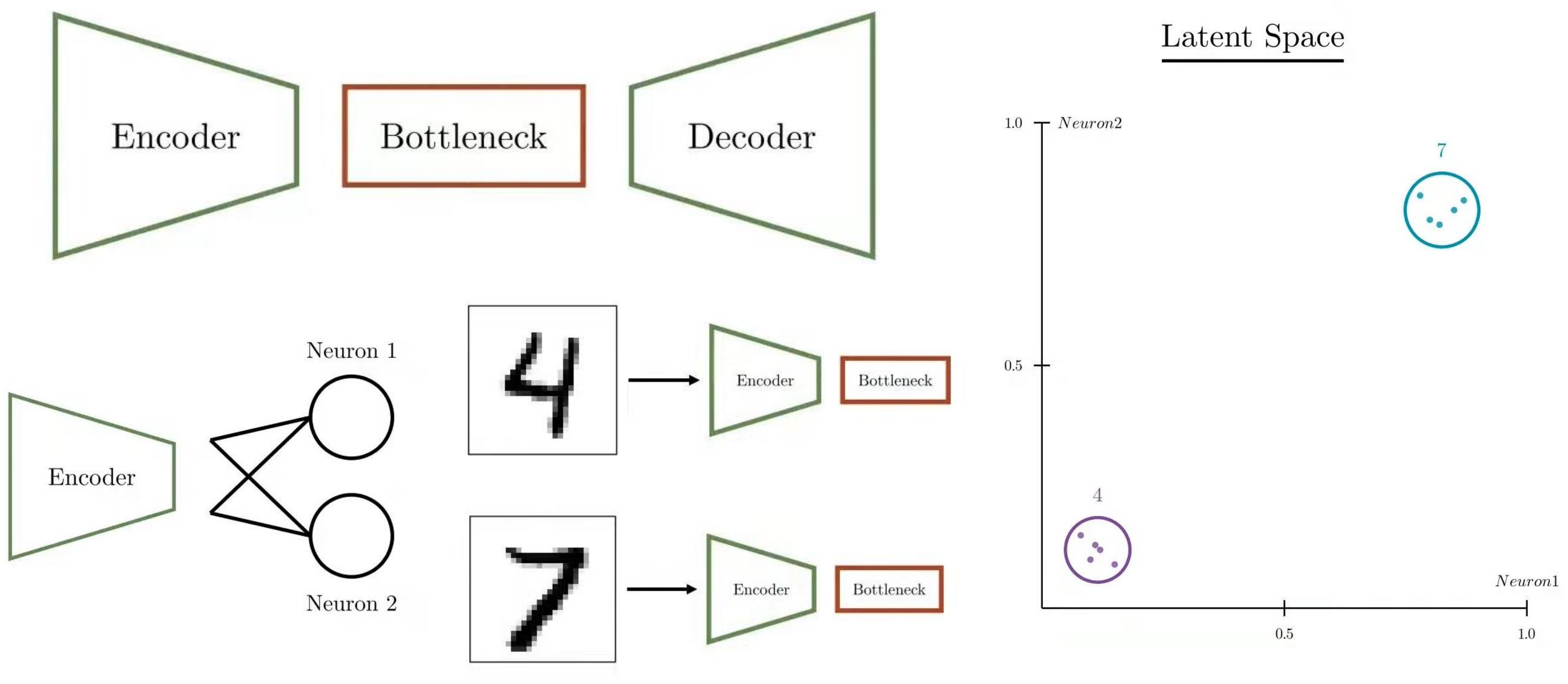
Basic Architecture

Also called 'Latent Space'

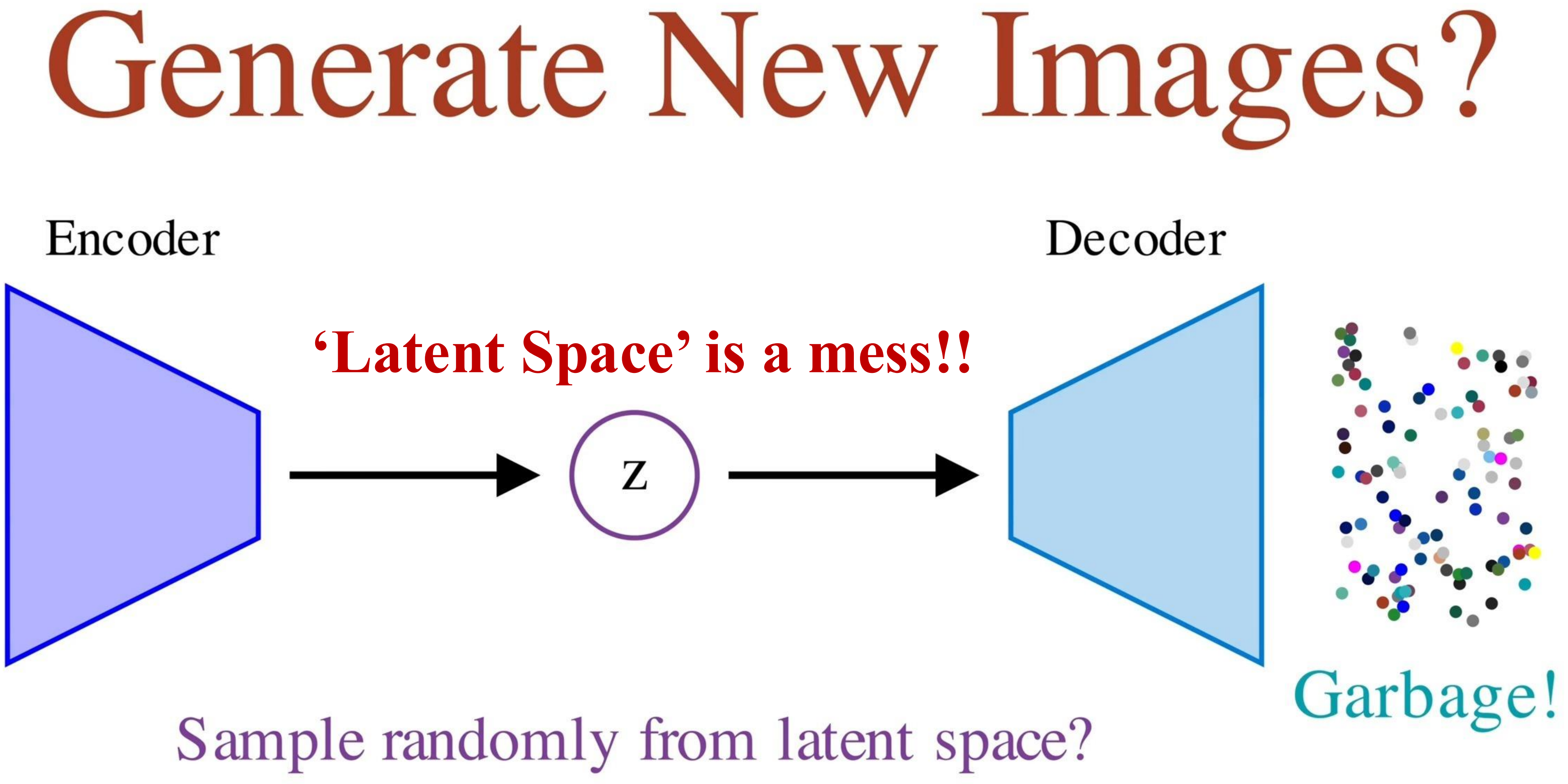


Why do we use Autoencoders?

Dimensionality reduction



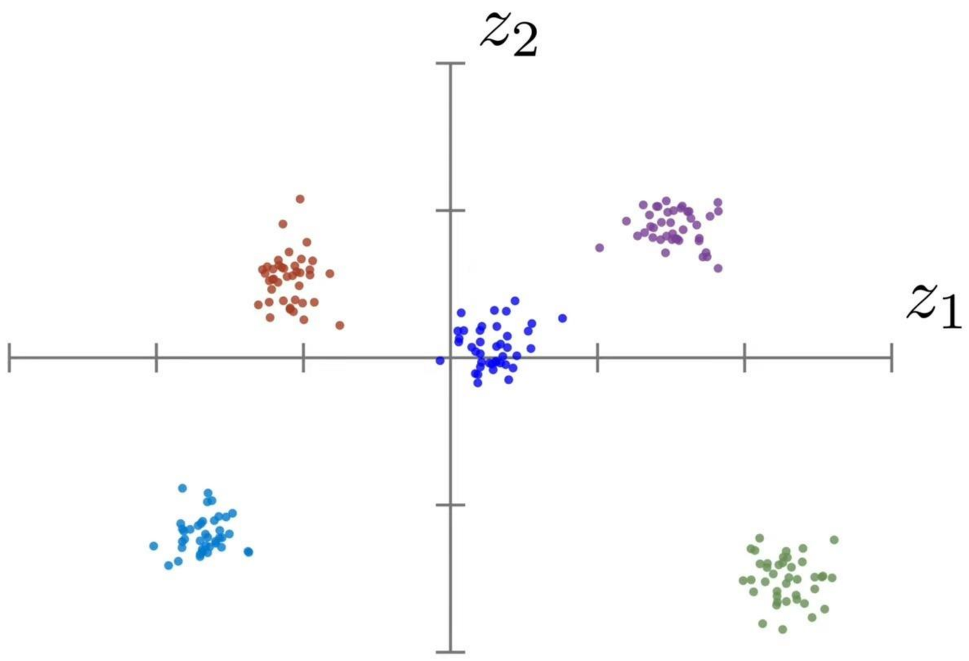
Issues with autoencoders?



Youtube: https://www.youtube.com/watch?v=geH5HnRapRs (DataMListic)

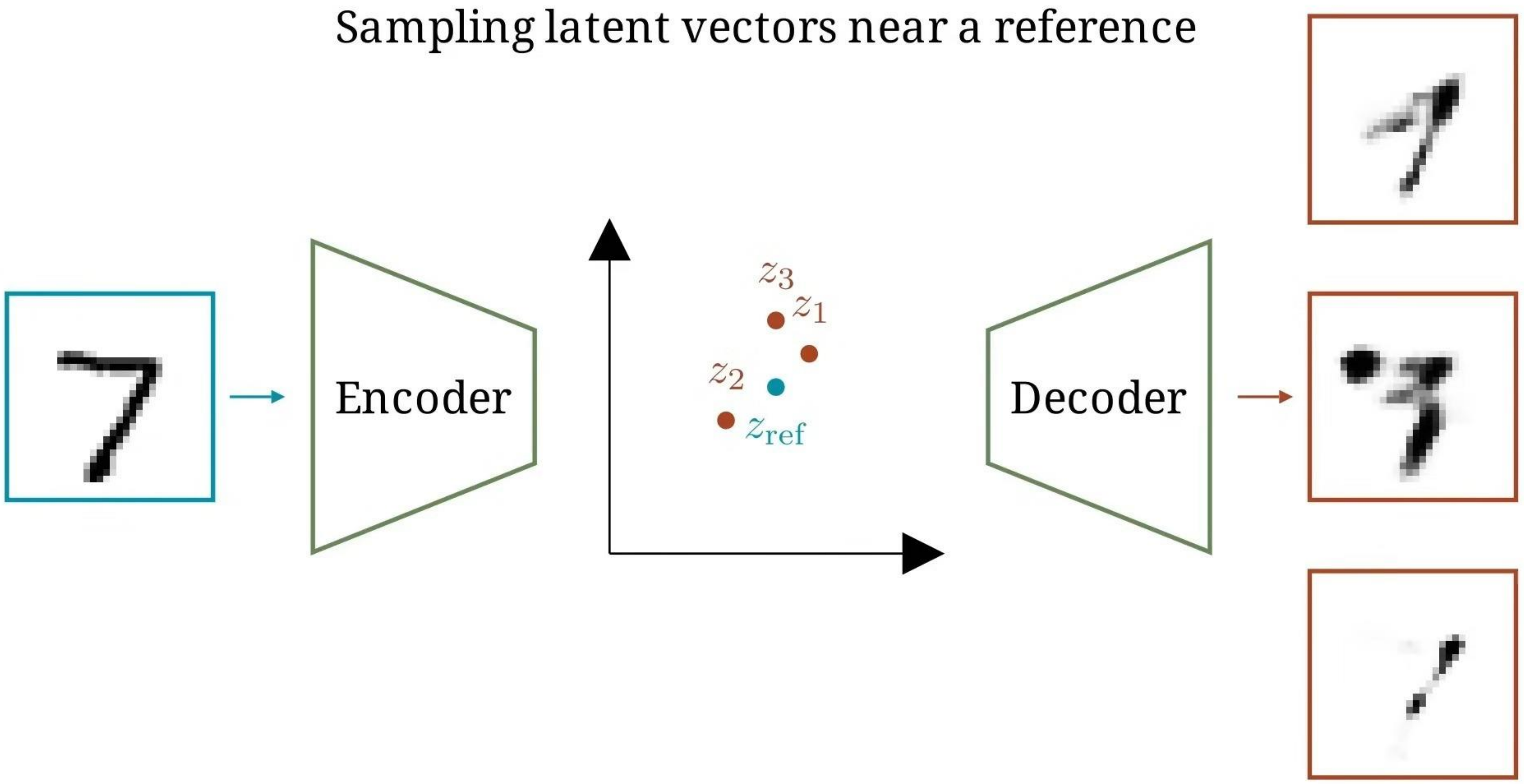
Issues with autoencoders?

The Latent Space Problem



Disconnected clusters!
No smooth transitions

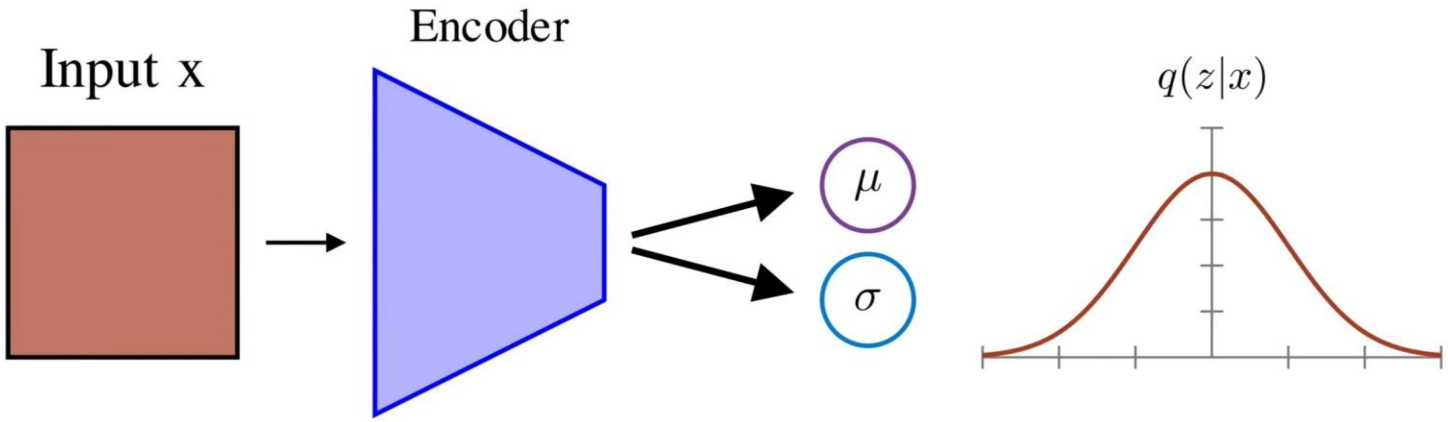
Issues with autoencoders?



Variational Autoencoders

Instead of mapping to points...
Map to probability distributions!

Probabilistic Encoding



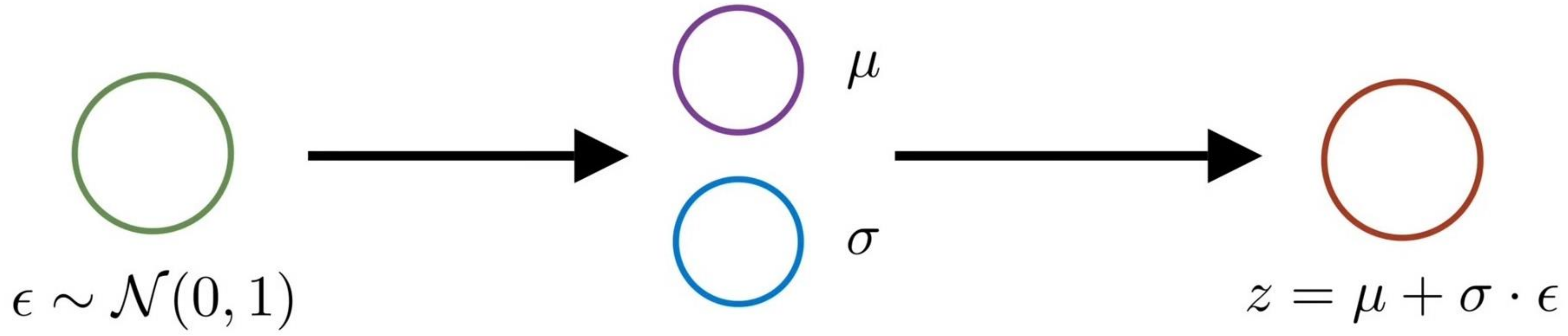
Distribution, not point!

Variational Autoencoders

The Reparameterization Trick

Problem: Can't backpropagate through random sampling!

Solution:



Now we can backpropagate!

Variational Autoencoders

The Training Objective

$$\mathcal{L} = \mathbb{E}_{q(z|x)}[\log p(x|z)] - D_{KL}(q(z|x)||p(z))$$

Evidence Lower Bound (ELBO)

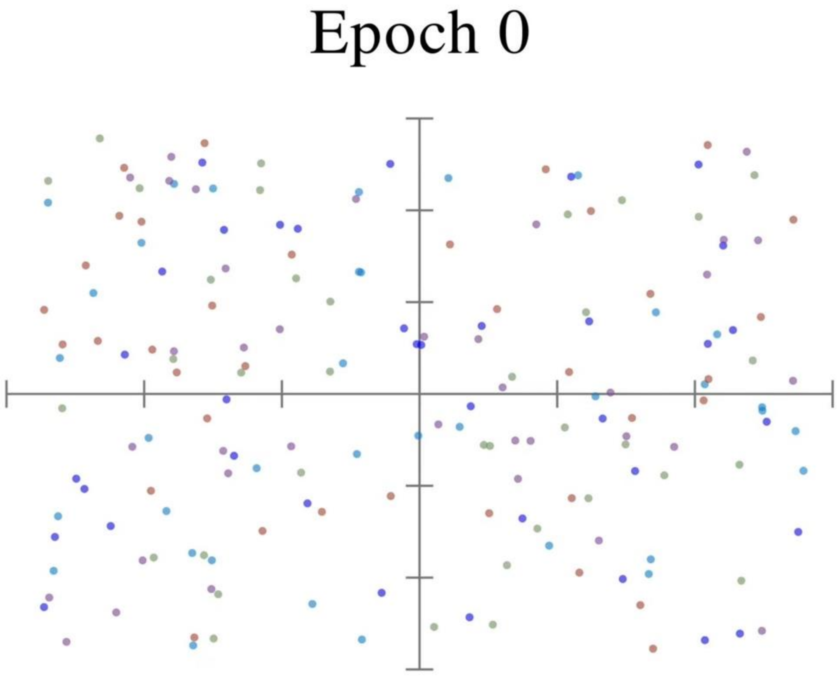
Reconstruction

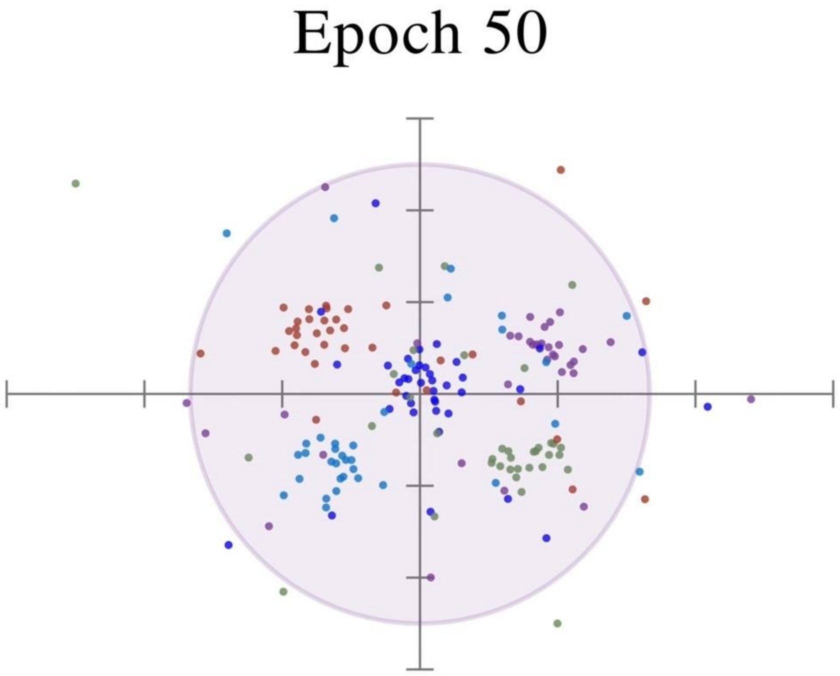
Makes sure the decoder can rebuild the original from the sample of latent distribution

Regularization

Acts as regulizer pulling all those individual gaussian distributions towards a standard normal distribution

Variational Autoencoders

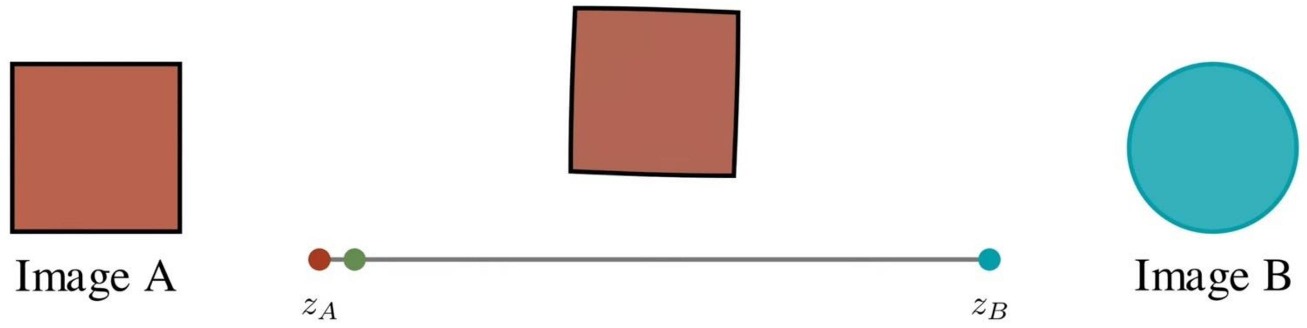




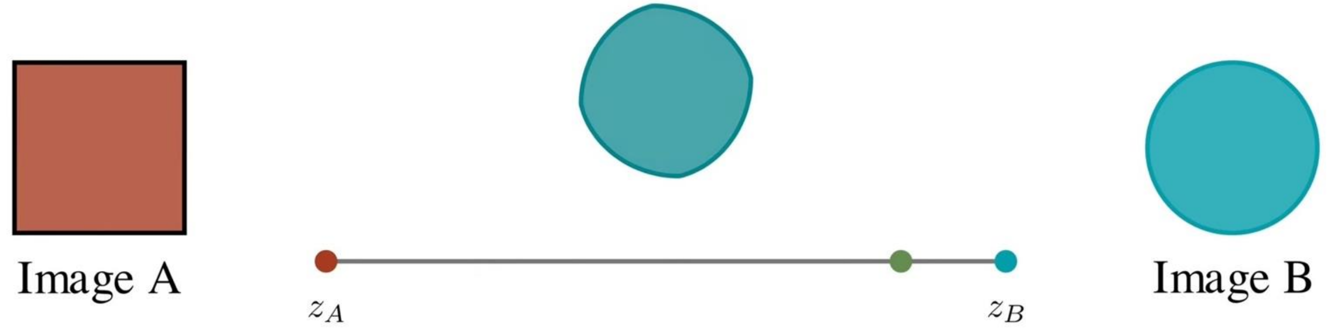
Gaussian distribution with clear structure

Variational Autoencoders

Smooth Interpolation

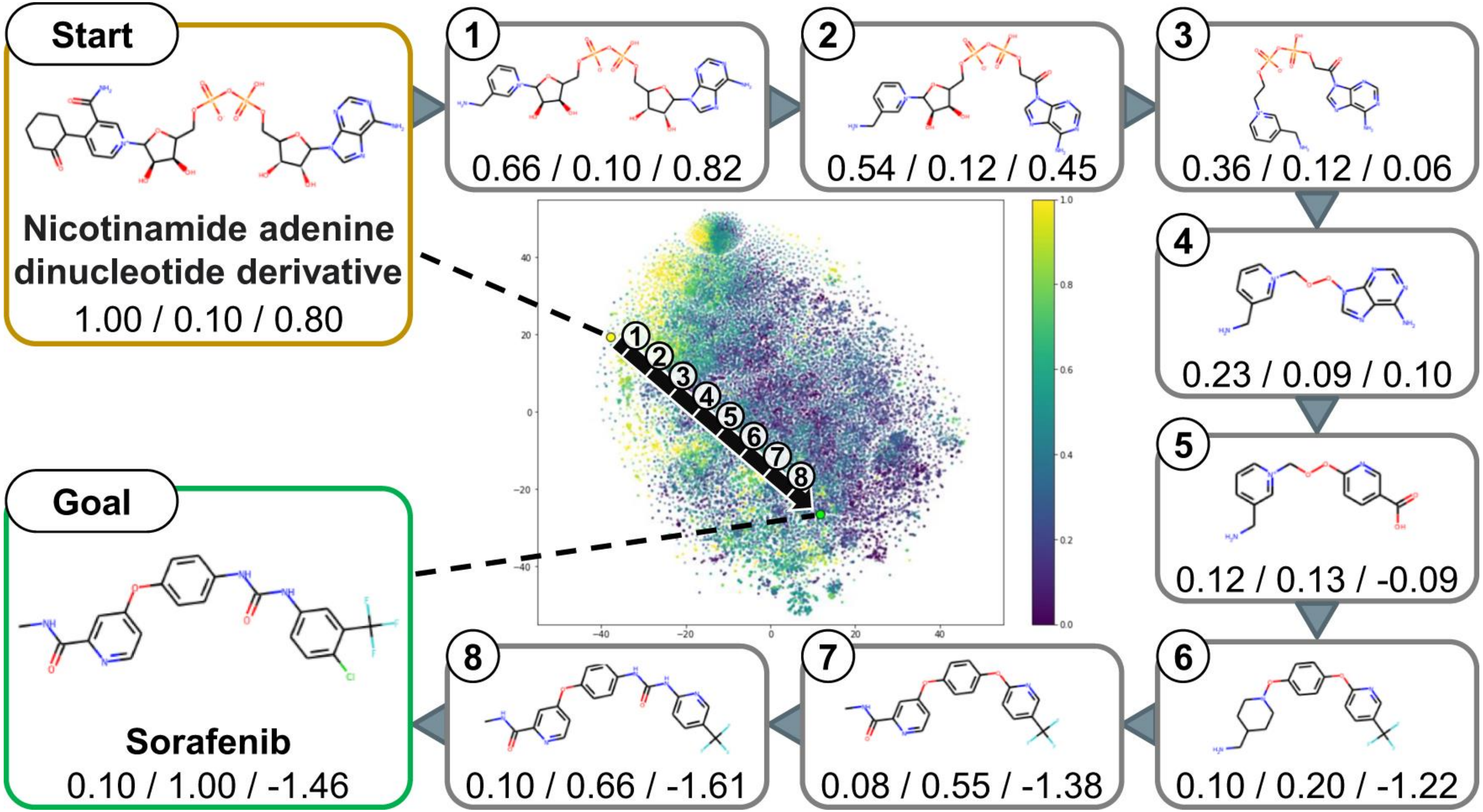


$$z_{interp} = (1 - \alpha) \cdot z_A + \alpha \cdot z_B$$



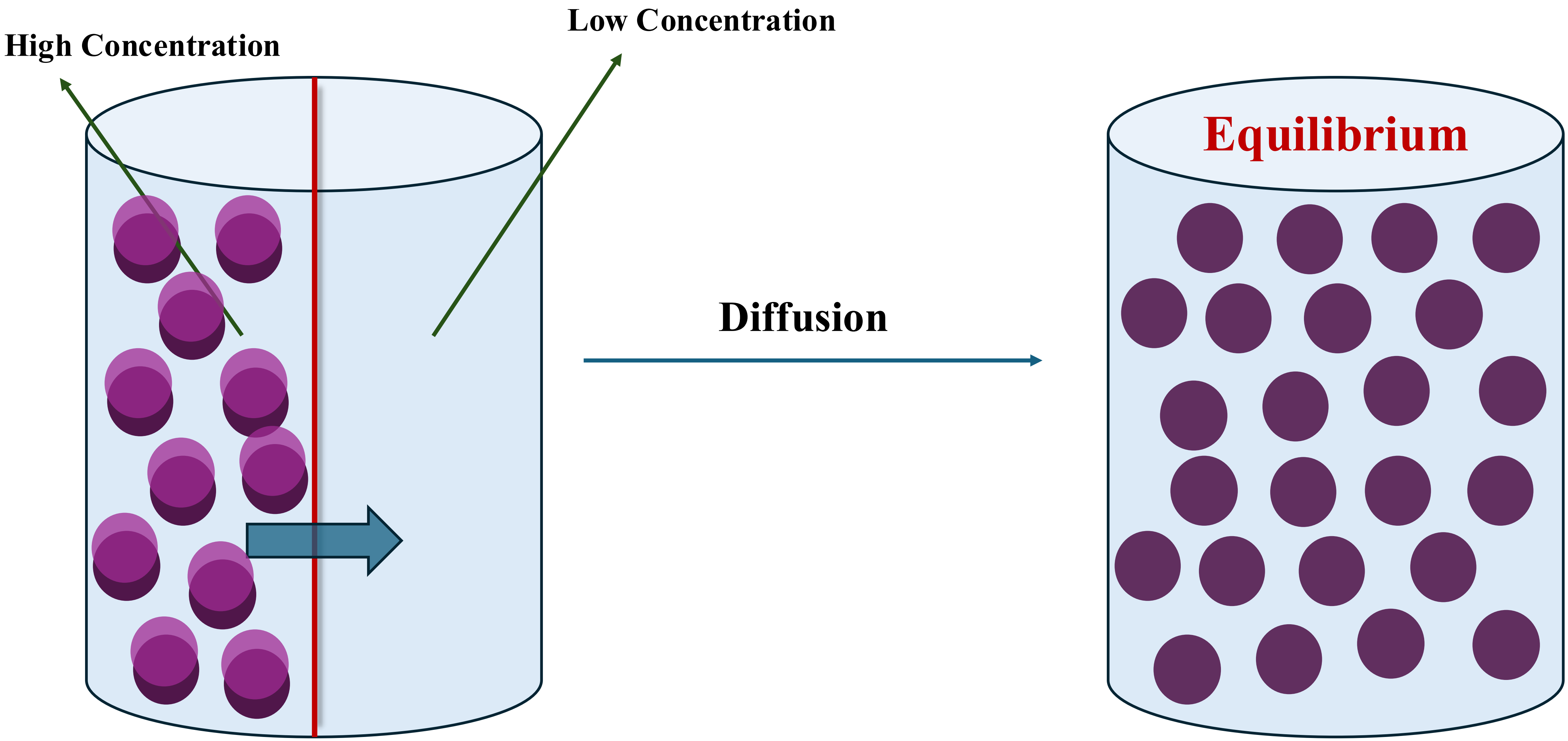
$$z_{interp} = (1 - \alpha) \cdot z_A + \alpha \cdot z_B$$

Variational Autoencoders



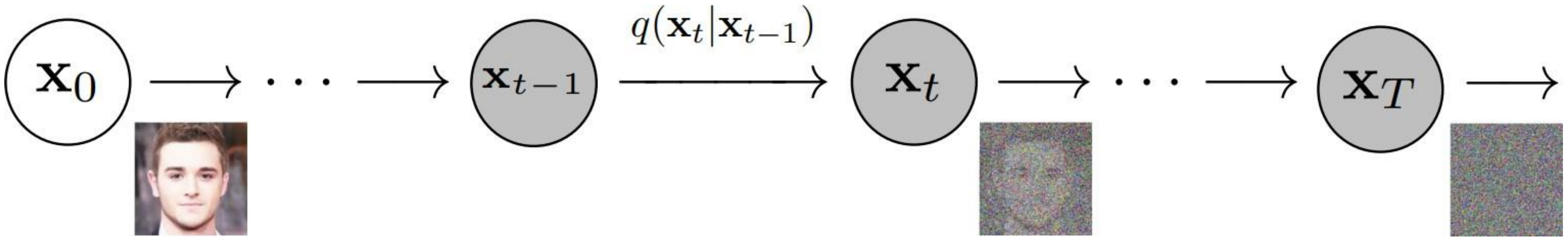
Commun Chem **6, 249 (2023)**

What is Diffusion?

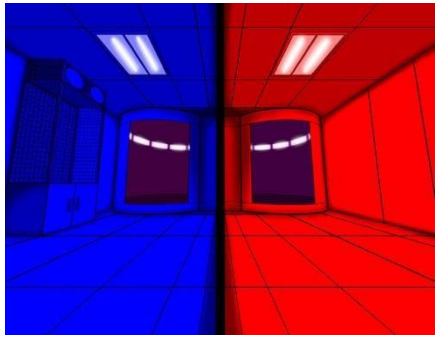


What is Diffusion?

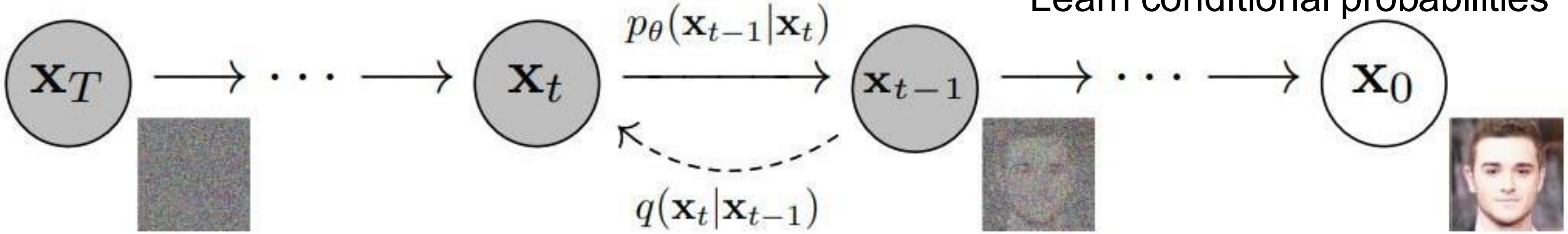
Forward diffusion



Gaussian noise added in a Markov chain

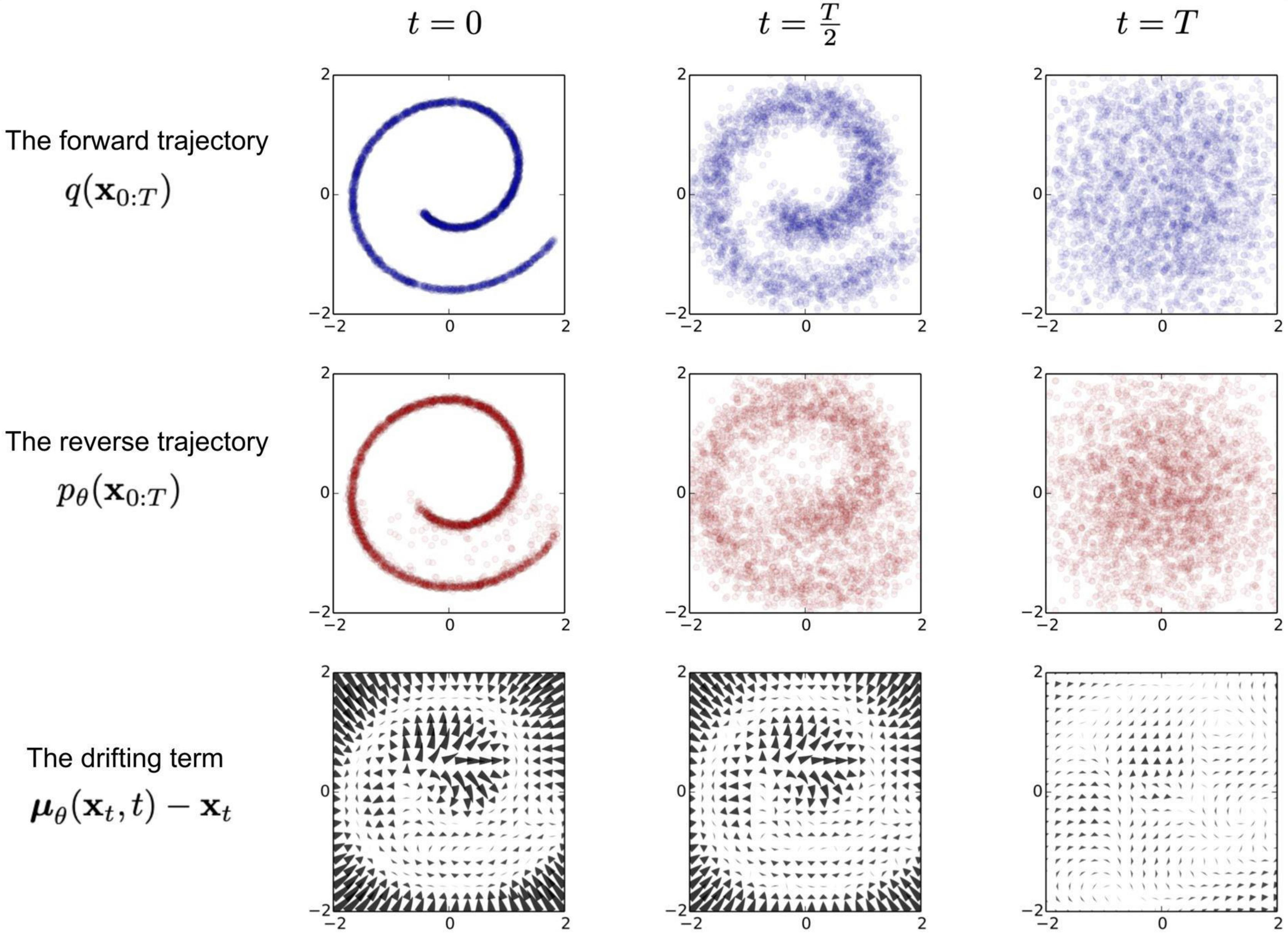


Learn conditional probabilities

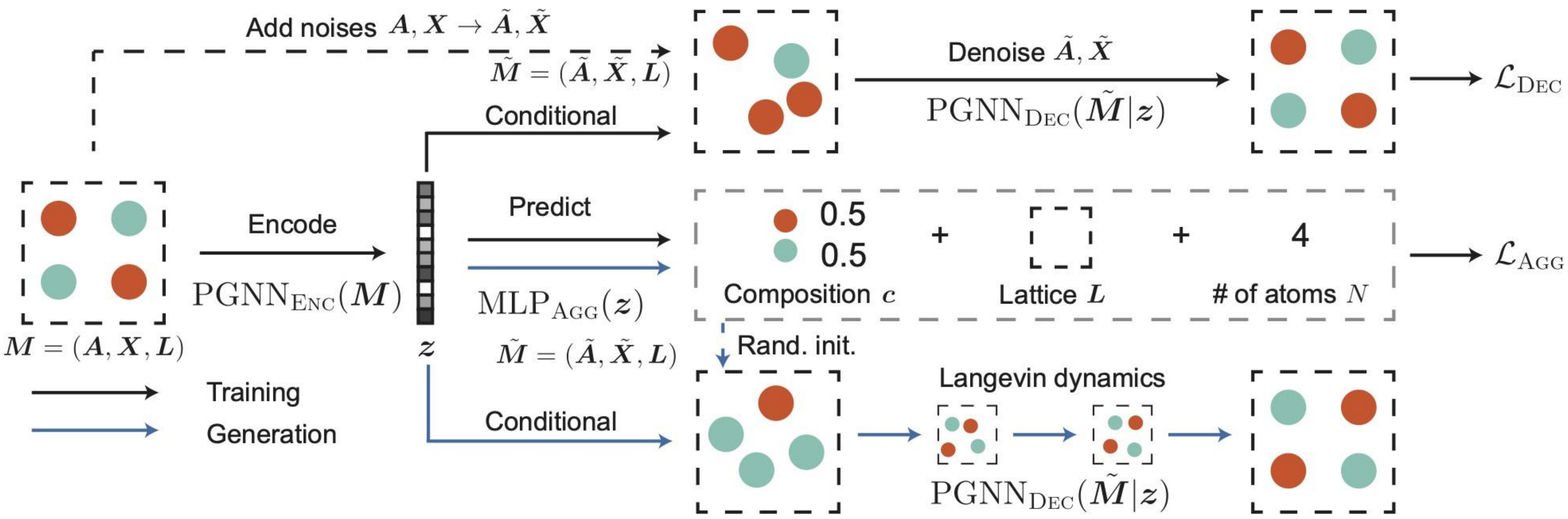


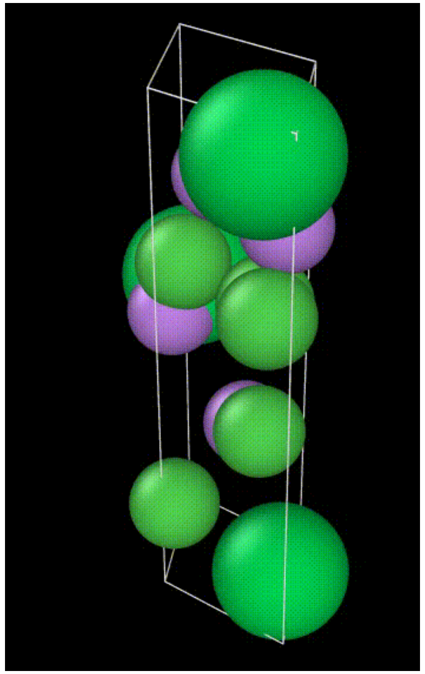
Reverse diffusion

What is Diffusion?



Diffusion models in materials



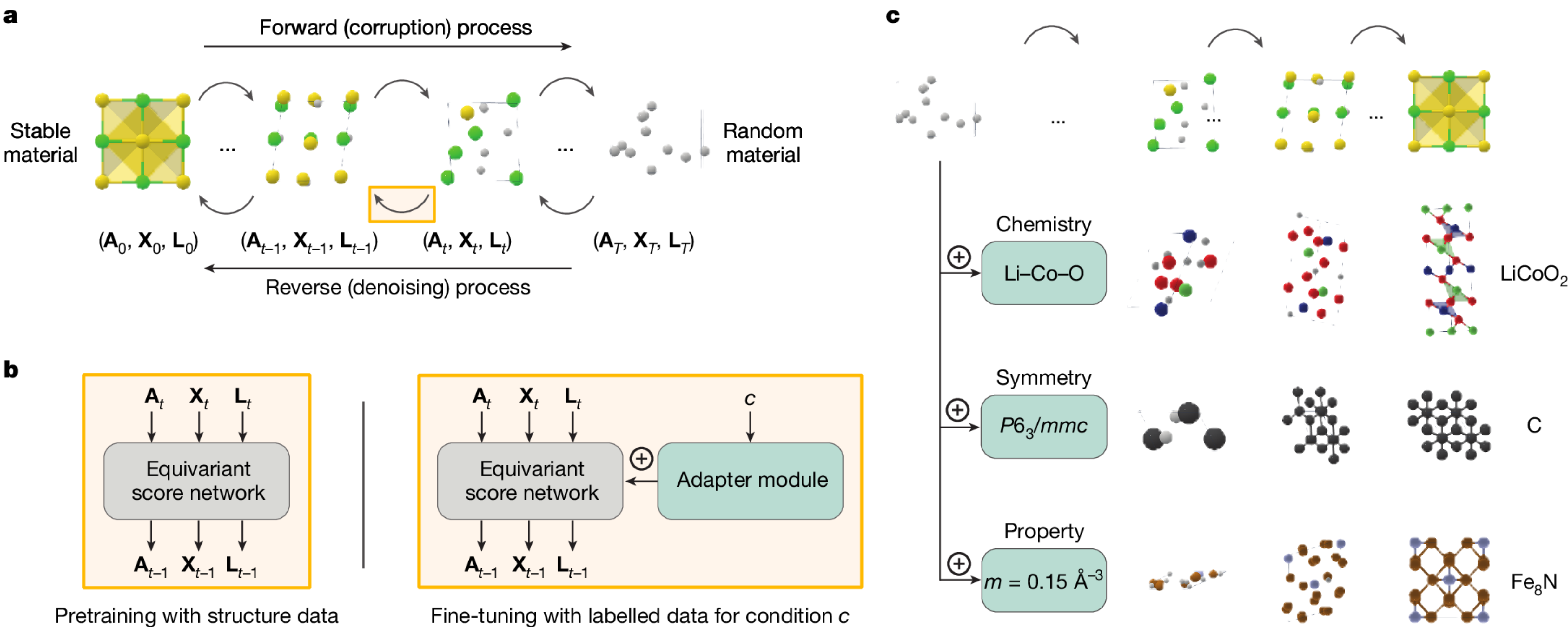


Crystal diffusion variational autoencoder (CDVAE)

- **One of the first diffusion models to be developed for structure prediction**
- **Periodic graph networks for encoding a latent space and denoising**
- **Property predictor: for composition, lattice, and number of atoms from latent space**
- **Langevin dynamics: final structure**

Diffusion models in materials

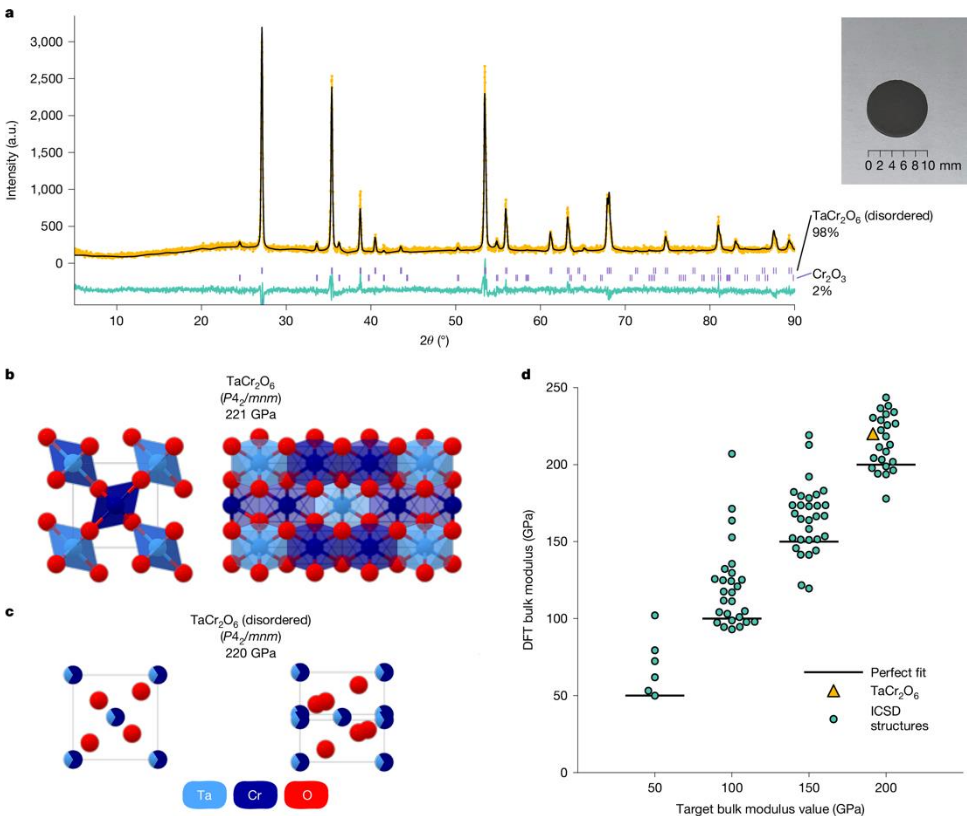
MatterGen by Microsoft

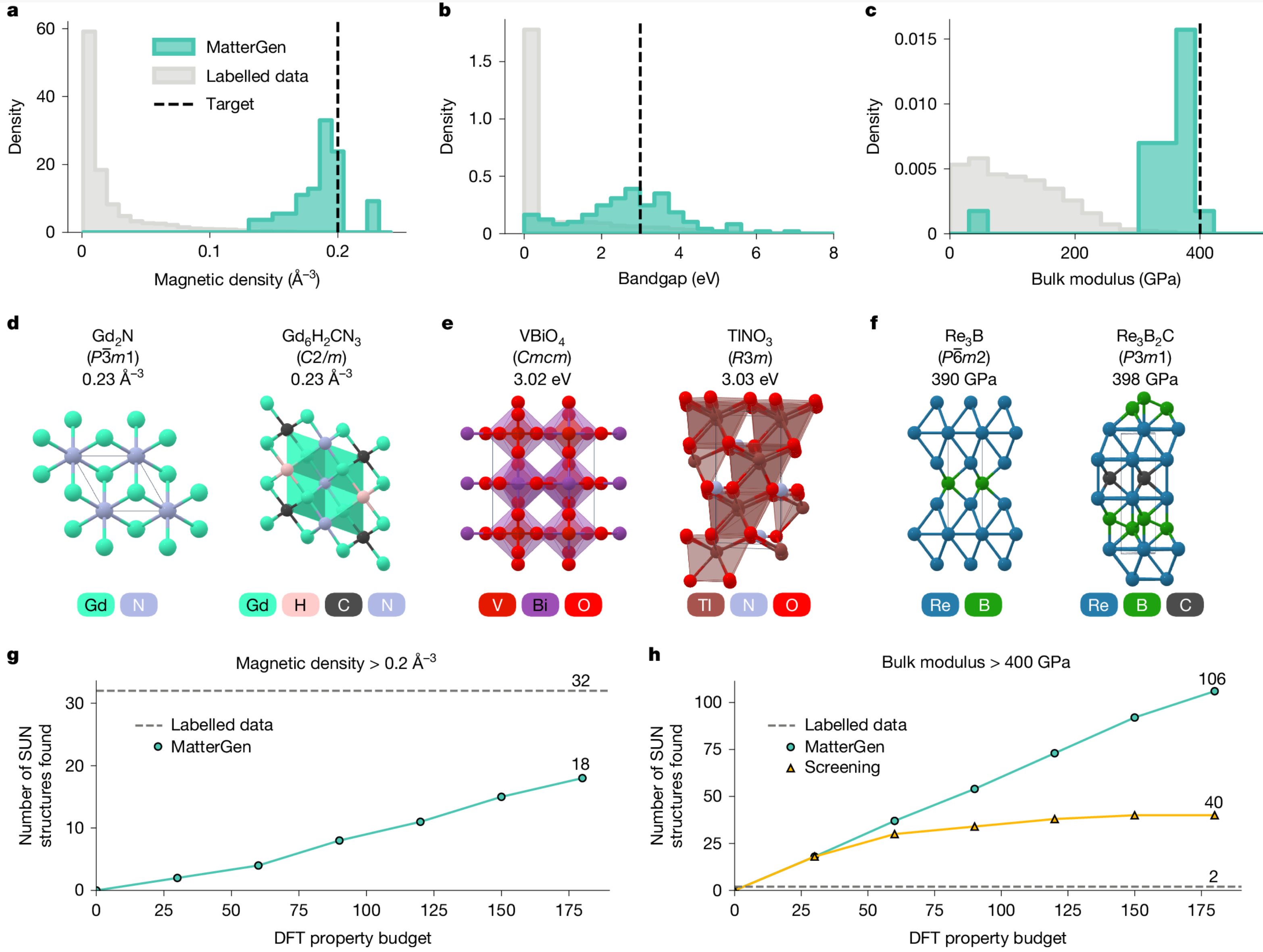


Nature 639, 624–632 (2025)

Diffusion models in materials

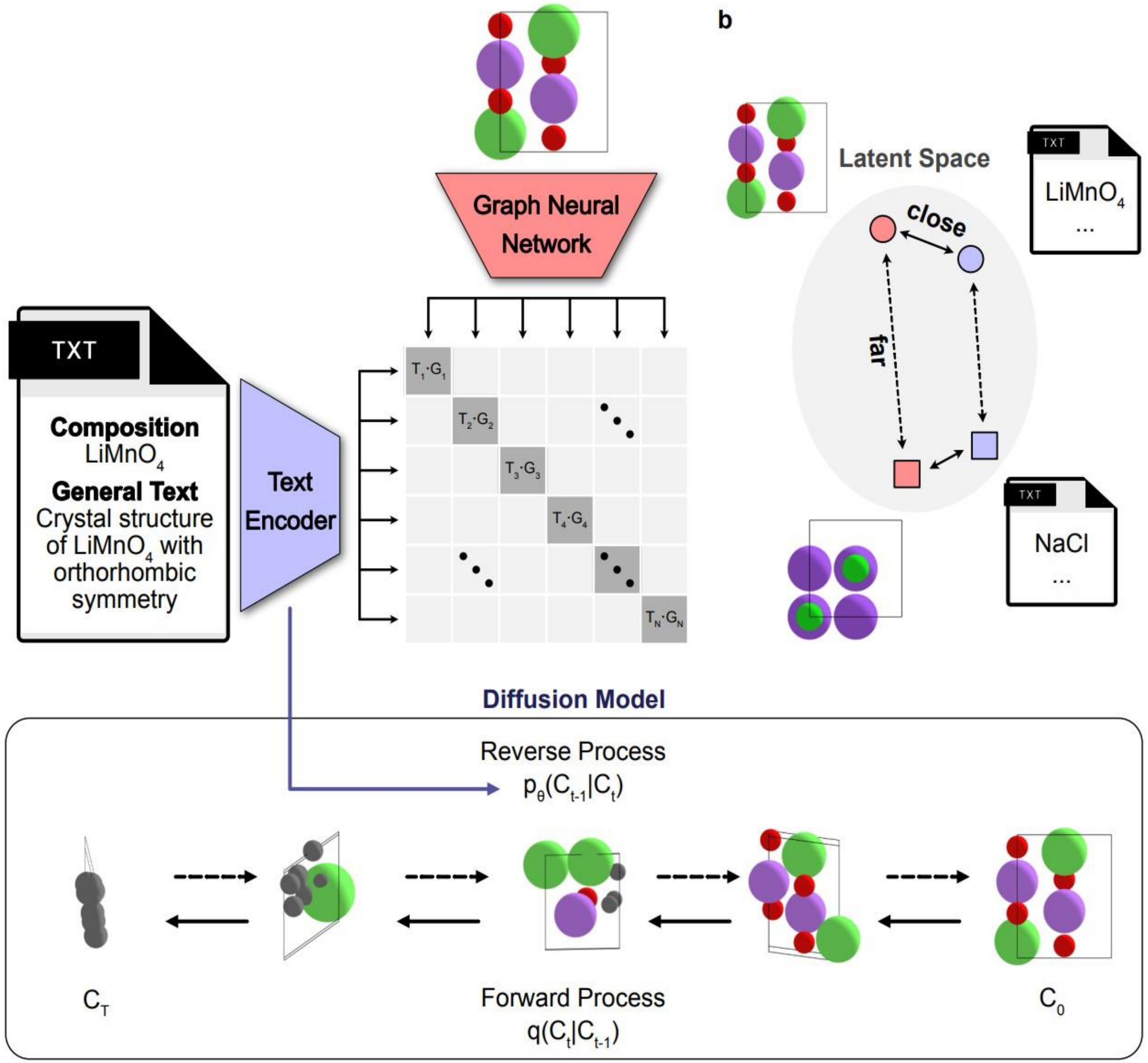
MatterGen by Microsoft





Nature 639, 624–632 (2025)

Chameleon: Text + Structure



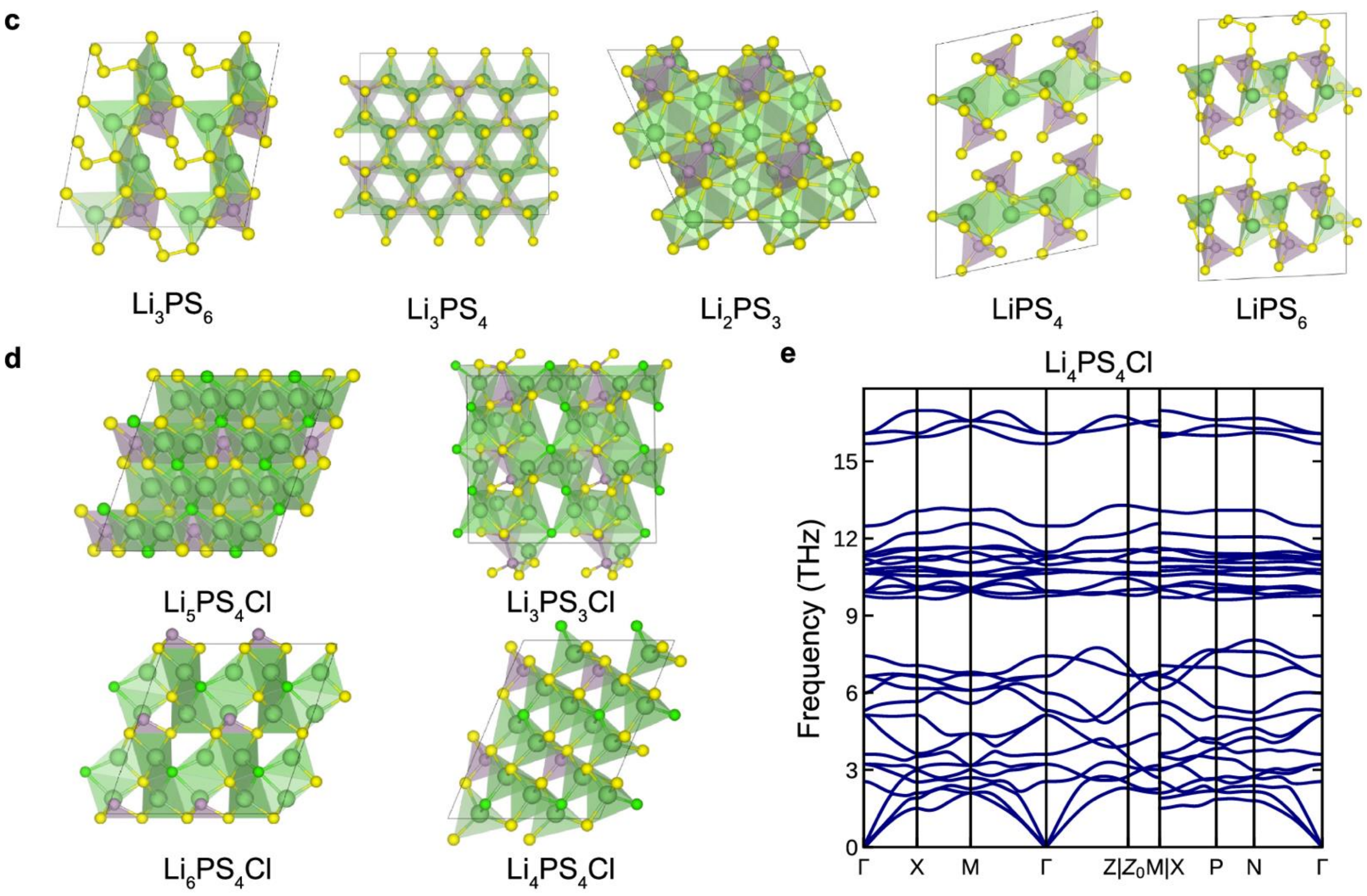
Text is aligned with graph embeddings for a given structure (contrastive learning): Crystal CLIP

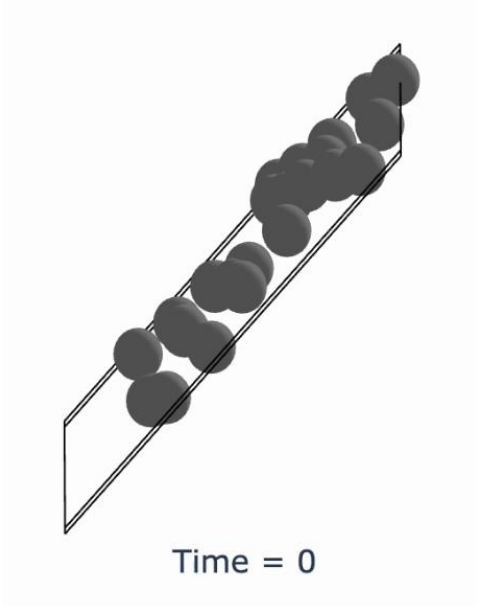
Text used to predict noise during denoising

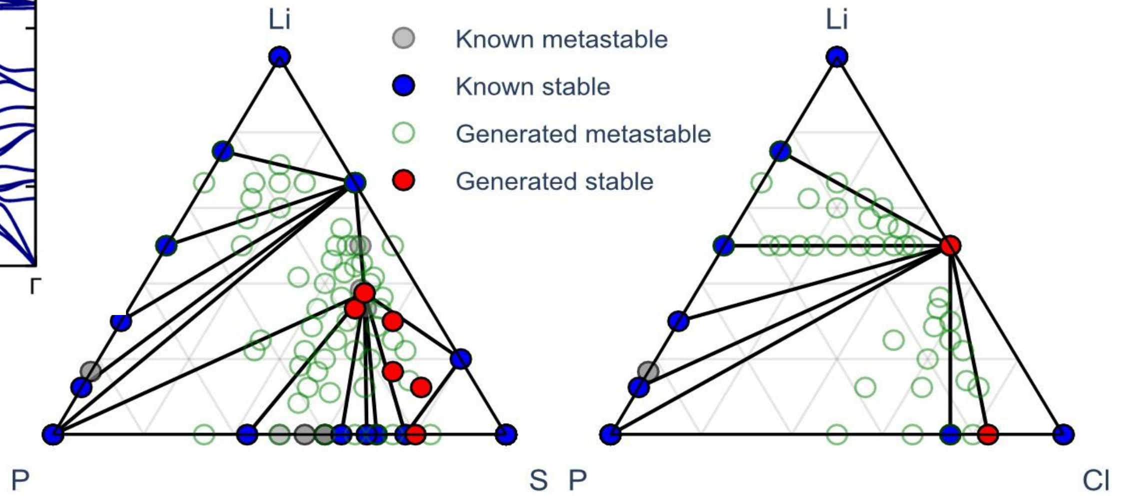
- **Can eliminate need for a secondary model for property prediction**

Nat Commun 16, 4379 (2025)

Chameleon in action!

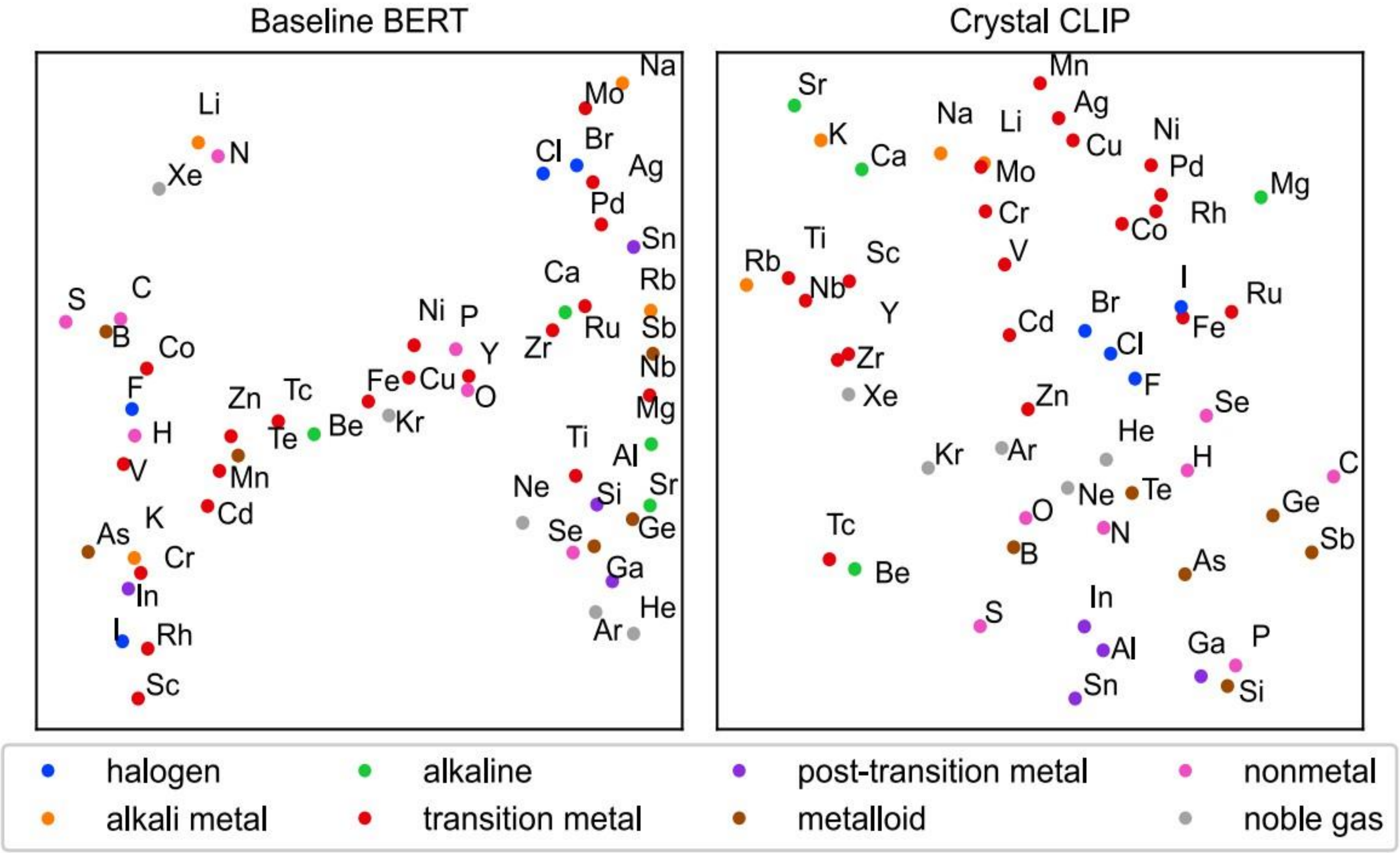






Nat Commun **16, 4379 (2025)**

Why Text + Graph



Too complex?

```python
from chemeleon import Chemeleon
from chemeleon.visualize import Visualizer
from ase.io import write
import os

# Load model
composition_model = Chemeleon.load_composition_model()

# Set parameters
n_samples = 10
n_atoms = 8
prompt = "Ti O S"

# Generate crystal structures
atoms_list = composition_model.sample(prompt, n_atoms, n_samples)

# Visualise
visualizer = Visualizer(atoms_list)
visualizer.view(index=0)

# Save cifs
output_folder = "chemeleon_structures"
os.makedirs(output_folder, exist_ok=True)

for i, atoms in enumerate(atoms_list):
    filename = os.path.join(output_folder, f"structure_{i+1}.cif")
    write(filename, atoms)
    print(f"Structure saved as {filename}")
```

Code front – No

Theory – Maybe?

Evaluating diffusion models

1. **Validity**

   Overlapping atoms? Large lattice parameters? Charge-neutral?

2. **Uniqueness**

   How diverse are the generated structures (in a randomly chosen sample size)?

3. **Structure matching**

   Do one of the 'best' generated structures include the known ground state?

4. **(Meta)stability**

   How (meta or un)stable are the generated structures?

   Evaluated using density functional theory or a foundational interatomic potential

Park et al., https://doi.org/10.26434/chemrxiv-2024-rw8p5

Summary

- **Generative models can facilitate structure generation/enumeration**
  - (Identify structures beyond simple human intuition)

- **Classical: autoencoders**
- **Modern: diffusion**

- **Diffusion:** learn probability distributions associated with noising/denoising

Nascent stage: can produce 'bad' structures or 'incorrect' compositions
- **More chemical constraints?**
- **Experimental validation?**

Thanks!

“Generative AI after generating a material’s structure from a text prompt”