

Step1

Denoising network:
operate on edges and nodes



Sampling

Step 3

Parameters for
updating model

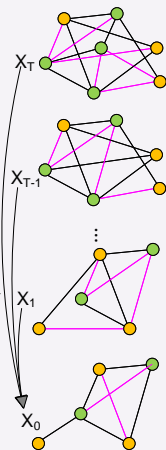
$$\sum_N \nabla_{\theta} \log p_{\theta}(\mathbf{X}_0|\mathbf{X}_t)$$

Accumulate the
obtained gradient

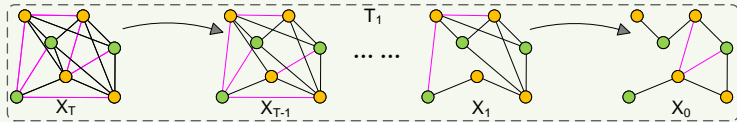
$$\nabla_{\theta} \log p_{\theta}(\mathbf{X}_0|\mathbf{X}_t)$$

Gradient calculation
for all trajectories

Estimate reward mean
and variance



Step 2



N sampled trajectories

Reward function

Reward

