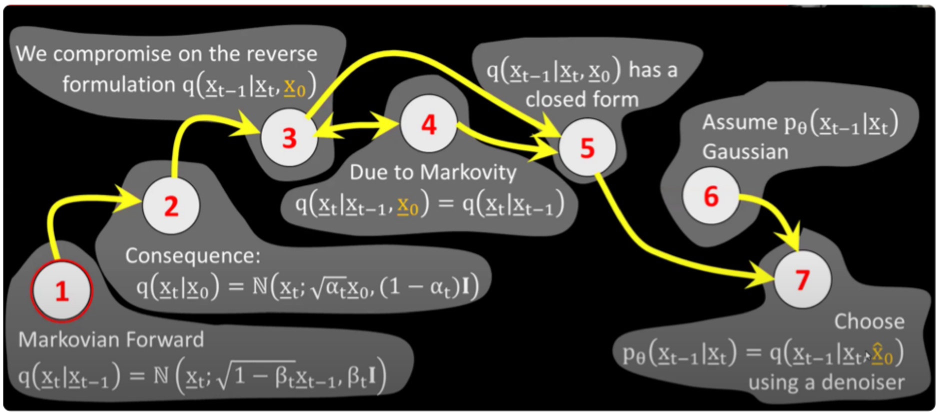
# DDIM

# Denoising diffusion implicit model.

This paper brought a true revolution.
→ a substantial speedup of DDPM
→ an alternative and elegant non-markovian diffusion process
→ a deterministic diffusion models (no noise injection)

## Summary : Creation of DDPM




# DDIM – Starting point

Step 5 $q(x_{t-1}|x_t, x_0)$ has a closed form.

They gave :

$$q(x_{t-1}|x_t, x_0) = N\left(x_{t-1}; \sqrt{\alpha_{t-1}}x_0 + \sqrt{1-\alpha_{t-1}-\sigma_t^2} \cdot \frac{x_t - \sqrt{\alpha_t}x_0}{\sqrt{1-\alpha_t}}, \sigma^2 I\right)$$

Why? How? What???

## Intuition

$$q(x_{t-1}|x_t, x_0) = N\left(x_{t-1}; \boxed{\sqrt{\alpha_{t-1}}x_0 + \sqrt{1-\alpha_{t-1}-\sigma^2 t} \cdot \boxed{\frac{x_t - \sqrt{\alpha_t}x_0}{\sqrt{1-\alpha_t}}}}, \sigma^2 I\right)$$

In terms of earlier

Therefore, taking this noise and computing $x_{t-1}$ with it by
$x_{t-1} = \sqrt{\alpha_{t-1}}\,x_0 + \sqrt{1-\alpha_{t-1}}\,\epsilon$
makes much more sense

equation, we had:
$x_t = \sqrt{\alpha_t}\,x_0 + \sqrt{1-\alpha_t}\,\epsilon$
Thus, this expression gives
$\epsilon \to$ the noise in $x_t$

☆ The introduction of $\sigma_t^2$ into this adds a softening and randomization effect.

# DDIM – The starting point

- In fact, DDIM starts by defining the joint distribution over all the steps via the following:

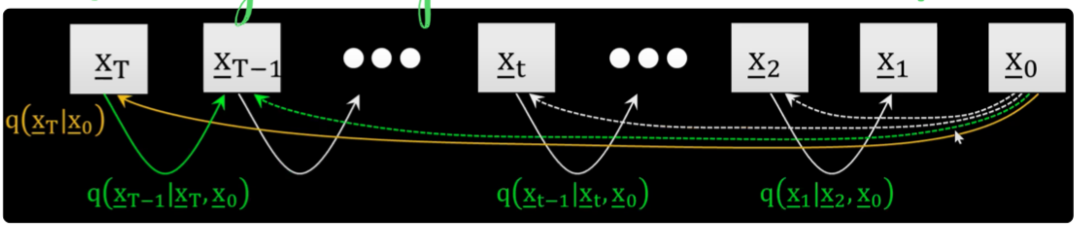
$$\boxed{q(\underline{x}_{1:T}|\underline{x}_0) = q(\underline{x}_T|\underline{x}_0)\prod_{t=2}^{T} q(\underline{x}_{t-1}|\underline{x}_t,\underline{x}_0)}$$

where $q(\underline{x}_T|\underline{x}_0) = \mathbb{N}(\underline{x}_T;\sqrt{\alpha_T}\underline{x}_0,(1-\alpha_T)\mathbf{I})$, and for all $1 < t < T$

$$q(\underline{x}_{t-1}|\underline{x}_t,\underline{x}_0) = \mathbb{N}\left(\underline{x}_{t-1};\sqrt{\alpha_{t-1}}\underline{x}_0 + \sqrt{1-\alpha_{t-1}-\sigma_t^2}\cdot\frac{\underline{x}_t-\sqrt{\alpha_t}\underline{x}_0}{\sqrt{1-\alpha_t}},\sigma_t^2\mathbf{I}\right)$$

- Strange! The order in which the process works is not a typical forward one …

What does this mean for the proposed forward path?
→ Start from $x_0$ and generate $x_T$.
→ Work your way backwards towards $x_1$.




# DDIM – Few amazing consequences

$$q(x_{t-1}|x_t,x_0) = N\left(x_{t-1};\sqrt{\alpha_{t-1}}\,x_0 + \sqrt{1-\alpha_{t-1}-\sigma_t^2}\cdot\frac{x_t-\sqrt{\alpha_t}x_0}{\sqrt{1-\alpha_t}},\sigma^2 I\right)$$

① A new parameter $\sigma_t^2$.

② When $\sigma_t^2 = \sqrt{\frac{1-\alpha_{t-1}}{1-\alpha_t}\left(1-\frac{\alpha_t}{\alpha_{t-1}}\right)}$ takes us back to DDPM

③ We can also choose $\sigma_t^2 = 0$ and this leads to a deterministic diffusion model in which we do not inject noise at all during the iterations.

④ The deterministic option has an ability to interpolate between images.

⑤ as step ④ very above is same, the denoiser for DDIM is same as DDPM, no change in training.

⑥ DDIM can run with much fewer steps (10-50) compared to DDPM that uses ~1000, while maintaining high quality. This is done by choosing a subset of temporal points and running over them while skipping the others.

What is it that enables DDIM to subsample the time without much harm?

→ ① the non-markovian nature of forward path.
(being non-markovian means $x_0$ keeps injecting itself at each iteration, carries information which helps us in reverse)

② the deterministic approach taken that has smoother trajectories
(moves from SDE to ODE, it samples more bigger steps rather that every tiny steps?

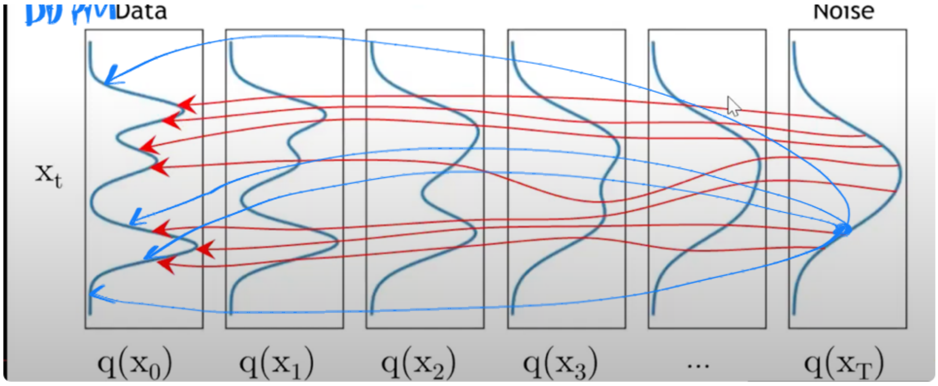
## DDIM - Interpolation results

→ as we have already said, for $\eta = 0$ DDIM becomes — deterministic, implying that each initialization $x_T$ is mapped into a unique and specific output image.

→ what does this mean? whereas the stochastic diffusion can create a variety from a given $x_T$, in deterministic DDIM there is a 1-to-1 mapping between $x_T$ and the $x_0$.

DDIM

Diffused Data Distributions




By interpolating in the $x_T$ - domain, we can move from one image to another while getting semantically meaningful images.

Example →