Title

# Reparametrization

## Reparameterization

Note that

Substitute this into our objective.

Note: is the noise added to data sample at time step to get (when we are using the diffusion kernel, which is going straight from to

Note that so

Observe that if we multiply the numerator and denominator by in the bolded term, the form of the first half of the sum matches the second half.

Now is our model, so we can do whatever we want with it. It’s just a formula we use with some parameters and So, let

Which we are totally allowed to do, since it’s just a function of

Then we get

This is our objective.

We ignore the constants out front and simply it further:

So we are just predicting the noise added.

Is another way to write it.

**Thus, to train a diffusion model:**

For all data:

For all time steps:

1. Generating a sample according to the diffusion kernel
2. Try to predict epsilon, using MSE loss.

OR

for a batch of data:

Generate random time step

Generate noise

Optimize

For inference:

Sample from N(0,I) to get

Compute f:

Sample from

Eventually, we compute ), which is our data sample.

Sigmas here are predetermined.