Algorithm 1 Training

1: **repeat**
2:    $\mathbf{x}_0 \sim q(\mathbf{x}_0)$
3:    $t \sim \mathrm{Uniform}(\{1, \ldots, T\})$
4:    $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
5:    Take gradient descent step on

$$\nabla_\theta \left\| \boldsymbol{\epsilon} - \boldsymbol{\epsilon}_\theta \left(\sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1 - \bar{\alpha}_t}\boldsymbol{\epsilon}, t \right) \right\|^2$$

6: **until** converged

Algorithm 2 Sampling

1: $\mathbf{x}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
2: **for** $t = T, ..., 1$ **do**
3:    $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ $if\ t > 1,\ else\ \mathbf{z} = \mathbf{0}$
4:    $\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(\mathbf{x}_t - \frac{1-\alpha_t}{\sqrt{1-\alpha_t}} \boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t) \right) + \sigma_t \mathbf{z}$
5: **end for**
6: **return** $\mathbf{x}_0$