Problem:

Given $q(x_{1:T}|x_0) = \prod_{t=1}^{T} q(x_t|x_{t-1})$

Show $q(x_{1:T}|x_0) = q(x_T|x_0) \prod_{t=T}^{2} q(x_{t-1}|x_t, x_0)$

proof:

$$q(x_{1:T}|x_0) = \prod_{t=1}^{T} q(x_t|x_{t-1}) = q(x_1|x_0) \prod_{t=2}^{T} q(x_t|x_{t-1})$$

$$= q(x_1|x_0) \prod_{t=2}^{T} q(x_{t-1}|x_t, x_0) \frac{q(x_t|x_0)}{q(x_{t-1}|x_0)}$$

$$= q(x_1|x_0) \prod_{t=2}^{T} q(x_{t-1}|x_t, x_0) \left[\frac{q(\cancel{x_2}|x_0)}{q(x_1|x_0)} \cdot \frac{q(\cancel{x_3}|x_0)}{q(\cancel{x_2}|x_0)} \cdot \ldots \cdot \frac{q(x_T|x_0)}{q(\cancel{x_{T-1}}|x_0)} \right]$$

$$= q(x_1|x_0) \frac{q(x_T|x_0)}{q(x_1|x_0)} \prod_{t=2}^{T} q(x_{t-1}|x_t, x_0)$$

$$= q(x_T|x_0) \prod_{t=T}^{2} q(x_{t-1}|x_t, x_0)$$

Prove Eq (4): $q(x_t|x_0) = \mathcal{N}(x_t; \sqrt{\bar{\alpha}_t}\, x_0, (1-\bar{\alpha}_t) I)$

proof:

Given $q(x_t|x_{t-1}) = \mathcal{N}(\sqrt{1-\beta_t}\, x_{t-1}, \beta_t I)$. $\epsilon_t \sim \mathcal{N}(0, I)$ $\forall t = 1 \sim T$

$$x_t = \sqrt{1-\beta_t}\, x_{t-1} + \sqrt{\beta_t}\, \epsilon_t$$

$$= \sqrt{\alpha_t}\, x_{t-1} + \sqrt{1-\alpha_t}\, \epsilon_t = \sqrt{\alpha_t} \left(\sqrt{\alpha_{t-1}}\, x_{t-2} + \sqrt{1-\alpha_{t-1}}\, \epsilon_{t-1} \right) + \sqrt{1-\alpha_t}\, \epsilon_t$$

$$= \sqrt{\alpha_t \alpha_{t-1}}\, x_{t-2} + \left(\sqrt{\alpha_t(1-\alpha_{t-1})}\, \epsilon_{t-1} + \sqrt{1-\alpha_t}\, \epsilon_t \right)$$

$\because \sqrt{\alpha_t(1-\alpha_{t-1})}\, \epsilon_{t-1} \sim \mathcal{N}(0, \alpha_t(1-\alpha_{t-1}) I)$ and $\sqrt{1-\alpha_t}\, \epsilon_t \sim \mathcal{N}(0, (1-\alpha_t) I)$

$\therefore \left(\sqrt{\alpha_t(1-\alpha_{t-1})}\, \epsilon_{t-1} + \sqrt{1-\alpha_t}\, \epsilon_t \right) \sim \mathcal{N}(0, [\alpha_t(1-\alpha_{t-1}) + (1-\alpha_t)] I) = \mathcal{N}(0, (1-\alpha_t \alpha_{t-1}) I)$

$\Rightarrow x_t = \sqrt{\alpha_t \alpha_{t-1}}\, x_{t-2} + \sqrt{1-\alpha_t \alpha_{t-1}}\, \epsilon'$, $\epsilon' \sim \mathcal{N}(0, (1-\alpha_t \alpha_{t-1}) I)$

$$= \cdots$$

$$= \sqrt{\bar{\alpha}_t}\, x_0 + \sqrt{1-\bar{\alpha}_t}\, \epsilon$$

$\Rightarrow q(x_t|x_0) = \mathcal{N}(x_t; \sqrt{\bar{\alpha}_t}\, x_0, (1-\bar{\alpha}_t) I)$

Prove Eq (6): $q(x_{t-1}|x_t,x_0)=\mathcal{N}(x_{t-1};\tilde{\mu}_t(x_t,x_0),\tilde{\beta}_t I)$

where $\tilde{\mu}_t(x_t,x_0)=\frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t}x_0+\frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}x_t$ and $\tilde{\beta}_t=\frac{1-\bar{\alpha}_{t-1}}{1-\bar{\alpha}_t}\beta_t$

proof:

$$q(x_{t-1}|x_t,x_0)=q(x_t|x_{t-1},x_0)\frac{q(x_{t-1}|x_0)}{q(x_t|x_0)}$$

$$\propto\exp\left(-\frac{1}{2}\left(\frac{(x_t-\sqrt{\alpha_t}x_{t-1})^2}{\beta_t}+\frac{(x_{t-1}-\sqrt{\bar{\alpha}_{t-1}}x_0)^2}{1-\bar{\alpha}_{t-1}}-\frac{(x_t-\sqrt{\bar{\alpha}_t}x_0)^2}{1-\bar{\alpha}_t}\right)\right)$$

$$=\exp\left(-\frac{1}{2}\left(\left(\frac{\alpha_t}{\beta_t}+\frac{1}{1-\bar{\alpha}_{t-1}}\right)x_{t-1}^2-\left(\frac{2\sqrt{\alpha_t}}{\beta_t}x_t+\frac{2\sqrt{\bar{\alpha}_{t-1}}}{1-\bar{\alpha}_{t-1}}x_0\right)x_{t-1}+C(x_t,x_0)\right)\right)$$

Given $\mathcal{N}(\mu,\sigma^2)\propto\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)=\exp\left(-\frac{1}{2}\left(\frac{1}{\sigma^2}x^2-\frac{2\mu}{\sigma^2}x+\frac{\mu^2}{\sigma^2}\right)\right)$

$$\Rightarrow\frac{1}{\sigma^2}=\frac{1}{\tilde{\beta}_t}=\left(\frac{\alpha_t}{\beta_t}+\frac{1}{1-\bar{\alpha}_{t-1}}\right)=\frac{\alpha_t-\alpha_t\bar{\alpha}_{t-1}+\beta_t}{\beta_t(1-\bar{\alpha}_{t-1})}=\frac{1-\bar{\alpha}_t}{\beta_t(1-\bar{\alpha}_{t-1})}\Rightarrow\tilde{\beta}_t=\frac{1-\bar{\alpha}_{t-1}}{1-\bar{\alpha}_t}\beta_t$$

$$\Rightarrow\frac{2\mu}{\sigma^2}=\frac{2\tilde{\mu}(x_t,x_0)}{\tilde{\beta}_t}=\left(\frac{2\sqrt{\alpha_t}}{\beta_t}x_t+\frac{2\sqrt{\bar{\alpha}_{t-1}}}{1-\bar{\alpha}_{t-1}}x_0\right)\Rightarrow\tilde{\mu}(x_t,x_0)=\frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}x_t+\frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t}x_0$$

$$\Rightarrow q(x_{t-1}|x_t,x_0)=\mathcal{N}(x_{t-1};\tilde{\mu}(x_t,x_0),\tilde{\beta}_t I)$$

Prove Eq (8): $L_{t-1}=\mathbb{E}_q\left[\frac{1}{2\sigma_t^2}\|\tilde{\mu}_t(x_t,x_0)-\mu_\theta(x_t,t)\|^2\right]+C$

proof:

Given $q(x_{t-1}|x_t,x_0)=\mathcal{N}(x_{t-1};\tilde{\mu}_t(x_t,x_0),\tilde{\beta}_t I)$, $p_\theta(x_{t-1}|x_t)=\mathcal{N}(x_{t-1};\mu_\theta(x_t,t),\sigma_t^2 I)$

let $\tilde{\beta}_t=\sigma_t^2$

$$L_{t-1}=\mathbb{E}_q\left[D_{KL}\left(q(x_{t-1}|x_t,x_0)\|p_\theta(x_{t-1}|x_t)\right)\right]$$

by theorem $D_{KL}(N_0\|N_1)=\frac{1}{2}\left(tr(\Sigma_1^{-1}\Sigma_0)-k+(\mu_1-\mu_0)^T\Sigma_1^{-1}(\mu_1-\mu_0)+\ln\left(\frac{\det\Sigma_1}{\det\Sigma_0}\right)\right)$, k is dimension

$$\Rightarrow L_{t-1}=\mathbb{E}_q\left[\frac{1}{2}\left(tr((\sigma_t^2 I)^{-1}(\sigma_t^2 I))-k+(\mu_\theta(x_t,t)-\tilde{\mu}_t(x_t,x_0))^T(\sigma_t^2 I)^{-1}(\mu_\theta(x_t,t)-\tilde{\mu}_t(x_t,x_0))+\ln\left(\frac{\det\sigma_t^2 I}{\det\sigma_t^2 I}\right)\right)\right]$$

$$=\mathbb{E}_q\left[\frac{1}{2\sigma_t^2}(\mu_\theta(x_t,t)-\tilde{\mu}_t(x_t,x_0))^T(\mu_\theta(x_t,t)-\tilde{\mu}_t(x_t,x_0))\right]+C$$

$$=\mathbb{E}_q\left[\frac{1}{2\sigma_t^2}\|\tilde{\mu}_t(x_t,x_0)-\mu_\theta(x_t,t)\|^2\right]+C$$