A.4. Proof of Theorem 3.1

Proof. According to Eq. (2), when $\gamma = 1$ and $\sigma_t = 0$, we have the following formula:

$$\boldsymbol{x}_{t-1} = \sqrt{\frac{\alpha_{t-1}}{\alpha_t}}\boldsymbol{x}_t - \sqrt{\alpha_{t-1}}\psi(\alpha_t, \alpha_{t-1}, 0)\epsilon_\theta(\boldsymbol{x}_t, t, \mathcal{C}_{\boldsymbol{y}})$$

We can then represent $\boldsymbol{x}_t$ as:

$$\begin{aligned}
\boldsymbol{x}_t &= \sqrt{\frac{\alpha_t}{\alpha_{t-1}}}\boldsymbol{x}_{t-1} + \sqrt{\alpha_t}\left(\sqrt{\frac{1}{\alpha_t}-1} - \sqrt{\frac{1}{\alpha_{t-1}}-1}\right)\epsilon_\theta(\boldsymbol{x}_t, t, \mathcal{C}_{\boldsymbol{y}}) \\
&= \sqrt{\frac{\alpha_t}{\alpha_{t-1}}}\left(\sqrt{\frac{\alpha_{t-1}}{\alpha_{t-2}}}\boldsymbol{x}_{t-2} + \sqrt{\alpha_{t-1}}\left(\sqrt{\frac{1}{\alpha_{t-1}}-1} - \sqrt{\frac{1}{\alpha_{t-2}}-1}\right)\epsilon_\theta(\boldsymbol{x}_{t-1}, t-1, \mathcal{C}_{\boldsymbol{y}})\right) \\
&\quad + \sqrt{\alpha_t}\left(\sqrt{\frac{1}{\alpha_t}-1} - \sqrt{\frac{1}{\alpha_{t-1}}-1}\right)\epsilon_\theta(\boldsymbol{x}_t, t, \mathcal{C}_{\boldsymbol{y}}) \\
&= \sqrt{\frac{\alpha_t}{\alpha_{t-1}}}\left(\sqrt{\frac{\alpha_{t-1}}{\alpha_{t-2}}}\left[\sqrt{\frac{\alpha_{t-2}}{\alpha_{t-3}}}\boldsymbol{x}_{t-3} + \sqrt{\alpha_{t-2}}\left(\sqrt{\frac{1}{\alpha_{t-2}}-1} - \sqrt{\frac{1}{\alpha_{t-3}}-1}\right)\epsilon_\theta(\boldsymbol{x}_{t-2}, t-2, \mathcal{C}_{\boldsymbol{y}})\right]\right. \\
&\quad \left. + \sqrt{\alpha_{t-1}}\left(\sqrt{\frac{1}{\alpha_{t-1}}-1} - \sqrt{\frac{1}{\alpha_{t-2}}-1}\right)\right)\epsilon_\theta(\boldsymbol{x}_{t-1}, t-1, \mathcal{C}_{\boldsymbol{y}}) \\
&\quad + \sqrt{\alpha_t}\left(\sqrt{\frac{1}{\alpha_t}-1} - \sqrt{\frac{1}{\alpha_{t-1}}-1}\right)\epsilon_\theta(\boldsymbol{x}_t, t, \mathcal{C}_{\boldsymbol{y}}) \\
&= \sqrt{\frac{\alpha_t\alpha_{t-1}\ldots\alpha_1}{\alpha_{t-1}\alpha_{t-2}\ldots\alpha_0}}\boldsymbol{x}_0 + \sum_{i=1}^{t}\sqrt{\alpha_t}\left(\sqrt{\frac{1}{\alpha_i}-1} - \sqrt{\frac{1}{\alpha_{i-1}}-1}\right)\epsilon_\theta(\boldsymbol{x}_i, i, \mathcal{C}_{\boldsymbol{y}}).
\end{aligned} \tag{16}$$

Let $\epsilon_\theta(\boldsymbol{x}_i, i, \mathcal{C}_{\boldsymbol{y}}) - \epsilon_\theta(\boldsymbol{x}_{i-1}, i, \mathcal{C}_{\boldsymbol{y}}) = \delta_i$. Then, we have

$$\epsilon_\theta(\boldsymbol{x}_t, t, \mathcal{C}_{\boldsymbol{y}}) = -\sqrt{1-\bar{\alpha}_t}\nabla\log p_\theta(\boldsymbol{x}_{t-1} \mid y) + \delta_t,$$

which leads to:

$$\begin{aligned}
\boldsymbol{x}_t = &\sqrt{\frac{\alpha_t\alpha_{t-1}\ldots\alpha_1}{\alpha_{t-1}\alpha_{t-2}\ldots\alpha_0}}\boldsymbol{x}_0 - \sum_{i=0}^{t-1}\sqrt{\alpha_t(1-\bar{\alpha}_{i+1})}\left(\sqrt{\frac{1}{\alpha_{i+1}}-1} - \sqrt{\frac{1}{\alpha_i}-1}\right)\nabla_{\boldsymbol{x}_i}\log p_\theta(\boldsymbol{x}_i \mid y) \\
&+ \sum_{i=0}^{t-1}\sqrt{\alpha_t}\left(\sqrt{\frac{1}{\alpha_{i+1}}-1} - \sqrt{\frac{1}{\alpha_i}-1}\right)\delta_{i+1}.
\end{aligned} \tag{17}$$

Let $s_i = \sqrt{\alpha_t(1-\bar{\alpha}_{i+1})}\left(\sqrt{\frac{1}{\alpha_{i+1}}-1} - \sqrt{\frac{1}{\alpha_i}-1}\right)$, for $0 \le i \le t-1$. Then the above expression simplifies to:

$$\boldsymbol{x}_t = \sqrt{\alpha_t}\boldsymbol{x}_0 - \sum_{i=0}^{t-1}s_i\nabla_{\boldsymbol{x}_i}\log p_\theta(\boldsymbol{x}_i \mid y) + \sum_{i=0}^{t-1}\frac{s_i}{1-\sqrt{\bar{\alpha}_{i+1}}}\delta_{i+1}. \tag{18}$$

Applying Bayes' theorem, the conditional term can be rewritten as:

$$\boldsymbol{x}_t = \sqrt{\alpha_t}\boldsymbol{x}_0 - \sum_{i=0}^{t-1}\nabla_{\boldsymbol{x}_i}\log\left(\frac{p_\theta(\boldsymbol{x}_i)p_\theta(y \mid \boldsymbol{x}_i)}{p_\theta(y)}\right)^{s_i} + \sum_{i=0}^{t-1}\frac{s_i}{1-\sqrt{\bar{\alpha}_{i+1}}}\delta_{i+1}. \tag{19}$$

Since the gradient of $\log p_\theta(y)$ with respect to $\boldsymbol{x}_i$ is zero, we obtain:

$$\boldsymbol{x}_t = \sqrt{\alpha_t}\boldsymbol{x}_0 - \sum_{i=0}^{t-1}\left[\nabla_{\boldsymbol{x}_i}\log p_\theta(\boldsymbol{x}_i)^{s_i} + \nabla_{\boldsymbol{x}_i}\log p_\theta(y \mid \boldsymbol{x}_i)^{s_i}\right] + \sum_{i=0}^{t-1}\frac{s_i}{1-\sqrt{\bar{\alpha}_{i+1}}}\delta_{i+1}. \tag{20}$$

The proof is complete. □

A.5. Proof of the Forward Process in Negative Instance Generation

Proof. From Eq. (2), when $\gamma = 1$ and $\sigma_t = 0$, the following equation holds:

$$\boldsymbol{x}_{t-1} = \sqrt{\frac{\alpha_{t-1}}{\alpha_t}}\boldsymbol{x}_t - \sqrt{\alpha_{t-1}}\psi(\alpha_t, \alpha_{t-1}, 0)\epsilon_\theta(\boldsymbol{x}_t, t, \mathcal{C}_{\boldsymbol{y}}).$$

Thus, we can represent $\boldsymbol{x}_t$ as:

$$\boldsymbol{x}_t = \sqrt{\frac{\alpha_t}{\alpha_{t-1}}}\boldsymbol{x}_{t-1} + \sqrt{\alpha_t}\psi(\alpha_t, \alpha_{t-1}, 0)\epsilon_\theta(\boldsymbol{x}_t, t, \mathcal{C}_{\boldsymbol{y}}).$$

Since $\epsilon_\theta(\boldsymbol{x}_t, t, \mathcal{C}_{\boldsymbol{y}})$ is not directly accessible, we adopt a forward Euler approximation, replacing $\epsilon_\theta(\boldsymbol{x}_t, t, \mathcal{C}_{\boldsymbol{y}})$ with $\epsilon_\theta(\boldsymbol{x}_{t-1}, t, \mathcal{C}_{\boldsymbol{y}})$ following DDIM (Song et al., 2020a). As a result, we obtain:

$$\boldsymbol{x}_t \approx \sqrt{\frac{\alpha_t}{\alpha t - 1}}\boldsymbol{x}_{t-1} + \sqrt{\alpha_t}\psi(\alpha_t, \alpha_{t-1}, 0), \epsilon_\theta(\boldsymbol{x}_{t-1}, t, \mathcal{C}_{\boldsymbol{y}}).$$

The proof is complete. □

A.6. Proof of Theorem 3.2

Proof. We begin with the following equation:

$$\tilde{\boldsymbol{x}}_{t-1} = \sqrt{\frac{\alpha_{t-1}}{\alpha_t}}\tilde{\boldsymbol{x}}_t - \sqrt{\alpha_{t-1}}\left(\sqrt{\frac{1}{\alpha_t} - 1} - \sqrt{\frac{1}{\alpha_{t-1}} - 1}\right)\epsilon_\theta(_t, t) \tag{21}$$

Rearranging, we obtain:

$$\tilde{\boldsymbol{x}}_t = \sqrt{\frac{\alpha_t}{\alpha_{t-1}}}\tilde{\boldsymbol{x}}_{t-1} + \sqrt{\alpha_t}\left(\sqrt{\frac{1}{\alpha_t} - 1} - \sqrt{\frac{1}{\alpha_{t-1}} - 1}\right)\epsilon_\theta(\tilde{\boldsymbol{x}}_t, t)$$

Following a similar derivation process as in Eq. (16), we obtain Eq. (22):

$$\tilde{\boldsymbol{x}}_t = \sqrt{\frac{\alpha_t\alpha_{t-1}\dots\alpha_1}{\alpha_{t-1}\alpha_{t-2}\dots\alpha_0}}\tilde{\boldsymbol{x}}_0 + \sum_{i=1}^{t}\sqrt{\alpha_t}\left(\sqrt{\frac{1}{\alpha_i} - 1} - \sqrt{\frac{1}{\alpha_{i-1}} - 1}\right)\epsilon_\theta(\tilde{\boldsymbol{x}}_i, i). \tag{22}$$

Let $\epsilon_\theta(\tilde{\boldsymbol{x}}_i, i) - \epsilon_\theta(\tilde{\boldsymbol{x}}_{i-1}, t) = \tilde{\delta}_i$. Based on Eq. (3), we have:

$$\epsilon_\theta(\tilde{\boldsymbol{x}}_t, t) = -\sqrt{1 - \bar{\alpha}_t}\nabla_{\tilde{\boldsymbol{x}}_{t-1}}\log p_\theta(\tilde{\boldsymbol{x}}_{t-1}) + \tilde{\delta}_t.$$

Therefore, we obtain the following equation for $\tilde{\boldsymbol{x}}_t$:

$$\tilde{\boldsymbol{x}}_t = \sqrt{\alpha_t}\tilde{\boldsymbol{x}}_0 - \sum_{i=0}^{t-1}\nabla_{\tilde{\boldsymbol{x}}_i}\log p_\theta(\tilde{\boldsymbol{x}}_i)^{s_i} + \sum_{i=0}^{t-1}\frac{s_i}{\sqrt{1 - \bar{\alpha}_{i+1}}}\tilde{\delta}_{i+1}. \tag{23}$$

In the initial reversion step, where $\tilde{\boldsymbol{x}}_t = \boldsymbol{x}_t$, we have:

$$\nabla\log p_\theta(\tilde{\boldsymbol{x}}_t) = \nabla\log p_\theta(\boldsymbol{x}_t).$$

Since

$$\nabla \log p_\theta(\tilde{\boldsymbol{x}}_t) = \nabla \log p_\theta(\tilde{\boldsymbol{x}}_{t-1}) - \frac{1}{\sqrt{1-\bar{\alpha}_t}}\tilde{\delta}_t, \quad \text{and} \quad \nabla \log p_\theta(\boldsymbol{x}_t) = \nabla \log p_\theta(\boldsymbol{x}_{t-1}) - \frac{1}{\sqrt{1-\bar{\alpha}_t}}\delta_t,$$

we obtain:

$$\nabla \log p_\theta(\tilde{\boldsymbol{x}}_{t-1}) = \nabla \log p_\theta(\boldsymbol{x}_{t-1}) - \frac{1}{\sqrt{1-\bar{\alpha}_t}}\delta_t + \frac{1}{\sqrt{1-\bar{\alpha}_t}}\tilde{\delta}_t.$$

By induction, we get:

$$\nabla \log p_\theta(\tilde{\boldsymbol{x}}_{t-k}) = \nabla \log p_\theta(\boldsymbol{x}_{t-k}) + \sum_{j=t-k}^{t-1} \frac{1}{\sqrt{1-\bar{\alpha}_{j+1}}}\left[\tilde{\delta}_{j+1} - \delta_{j+1}\right],$$

or more generally:

$$\nabla \log p_\theta(\tilde{\boldsymbol{x}}_i) = \nabla \log p_\theta(\boldsymbol{x}_i) + \sum_{j=i}^{t-1} \frac{1}{\sqrt{1-\bar{\alpha}_{j+1}}}\left[\tilde{\delta}_{j+1} - \delta_{j+1}\right].$$

Substituting into Eq. (23), we get:

$$\tilde{\boldsymbol{x}}_t = \sqrt{\alpha_t}\tilde{\boldsymbol{x}}_0 - \sum_{i=0}^{t-1} \nabla_{\boldsymbol{x}_i} \log p_\theta(\boldsymbol{x}_i)^{s_i} - \sum_{i=0}^{t-1}\sum_{j=i}^{t-1} \frac{s_i}{\sqrt{1-\bar{\alpha}_{j+1}}}\left[\tilde{\delta}_{j+1} - \delta_{j+1}\right] + \sum_{i=0}^{t-1} \frac{s_i}{\sqrt{1-\bar{\alpha}_{i+1}}}\tilde{\delta}_{i+1}. \tag{24}$$

Meanwhile, the corresponding equation from Theorem 3.1 is:

$$\boldsymbol{x}_t = \sqrt{\alpha_t}\boldsymbol{x}_0 - \sum_{i=0}^{t-1} \left[\nabla_{\boldsymbol{x}_i} \log p_\theta(\boldsymbol{x}_i)^{s_i} + \nabla_{\boldsymbol{x}_i} \log p_\theta(y|\boldsymbol{x}_i)^{s_i}\right] + \sum_{i=0}^{t-1} \frac{s_i}{\sqrt{1-\bar{\alpha}_{i+1}}}\delta_{i+1}. \tag{25}$$

Equating both sides, we obtain:

$$\begin{aligned}
&\sqrt{\alpha_t}\boldsymbol{x}_0 - \sum_{i=0}^{t-1} \left[\nabla_{\boldsymbol{x}_i} \log p_\theta(\boldsymbol{x}_i)^{s_i} + \nabla_{\boldsymbol{x}_i} \log p_\theta(y|\boldsymbol{x}_i)^{s_i}\right] + \sum_{i=0}^{t-1} \frac{s_i}{\sqrt{1-\bar{\alpha}_{i+1}}}\delta_{i+1} \\
&\quad = \sqrt{\alpha_t}\tilde{\boldsymbol{x}}_0 - \sum_{i=0}^{t-1} \nabla_{\boldsymbol{x}_i} \log p_\theta(\boldsymbol{x}_i)^{s_i} - \sum_{i=0}^{t-1}\sum_{j=i}^{t-1} \frac{s_i}{\sqrt{1-\bar{\alpha}_{j+1}}}\left[\tilde{\delta}_{j+1} - \delta_{j+1}\right] + \sum_{i=0}^{t-1} \frac{s_i}{\sqrt{1-\bar{\alpha}_{i+1}}}\tilde{\delta}_{i+1}.
\end{aligned} \tag{26}$$

Solving for $\tilde{\boldsymbol{x}}_0$, we arrive at:

$$\tilde{\boldsymbol{x}}_0 = \boldsymbol{x}_0 - \frac{1}{\sqrt{\alpha_t}}\sum_{i=0}^{t-1} \nabla_{\boldsymbol{x}_i} \log p_\theta(y|\boldsymbol{x}_i)^{s_i} + \sum_{i=1}^{t-1}\sum_{j=i}^{t-1} \frac{s_i}{\sqrt{\alpha_t(1-\bar{\alpha}_{j+1})}}\left[\tilde{\delta}_{j+1} - \delta_{j+1}\right]. \tag{27}$$

The proof is complete. □