算法1：基于概率掩码的高效去噪扩散（EDDPM）

输入：随机初始化的扩散模型F_θ，训练轮数N，最终掩码率γ_f。
初始化：各时间步采样概率$s = 1 \in \mathbb{R}^T$（初始全激活）

1. 对于轮次$e = 1, 2, \ldots, N$：

    2. 根据式(10)计算当前掩码率γ_e。

    ### 【掩码率的渐进增加】

    为控制模型复杂度，定义最终掩码率γ_f（即$K = \gamma_f T$）。为稳定训练，掩码率从1（全步骤）渐进过渡到γ_f，采用（Zhu & Gupta, 2017）的增长函数：

    $$\gamma_e = \begin{cases} 1, & \text{若 } e < e_1, \\ \gamma_f + (1-\gamma_f)\left(1 - \frac{e-e_1}{N-e_1}\right)^3, & \text{否则}, \end{cases} \tag{10}$$

    其中e_1为初始全步骤训练轮数，γ_e为当前轮次e的剩余步骤比例。

    3. 对于每个训练迭代：

        4. 采样数据 mini-batch X_B。
        5. 基于分数s进行伯努利采样，生成扩散步骤掩码。
        6. 根据式(6)，基于采样掩码更新方差调度。

            $x_t \sim \mathcal{N}\left(\sqrt{\alpha_t(m)}x_0,\ (1-\alpha_t(m))\mathbf{I}\right)$
            其中$\alpha_t(m) = \prod_{i=1}^{t} \hat{\alpha}_i(m)$，且$\hat{\alpha}_t(m) = 1 - \beta_t m_t$。(6)

        7. 随机采样未掩码的扩散步骤用于训练。
        8. 计算扩散模型损失$\mathcal{L}_\theta^t$。

            其中损失函数$\mathcal{L}_\theta^t(x_0, \epsilon, m)$的形式为：
            $\mathcal{L}_\theta^t(x_0, \epsilon, m) = C_t \left\| \epsilon - \epsilon_\theta\left(\sqrt{\alpha_t(m)}x_0 + \sqrt{1-\alpha_t(m)}\epsilon,\ t\right) \right\|^2$

        9. 对F_θ反向传播，估计$\nabla_\theta \Phi(\theta, s)$。

            $\nabla_\theta \Phi(\theta, s) = \mathbb{E}_{m\sim p(m|s)} \mathbb{E}_{x_0,\epsilon,t|m} \nabla_\theta \mathcal{L}_\theta^t(x_0, \epsilon, m) \quad (8)$

        10. 根据式(9)估计$\nabla_s \Phi(\theta, s)$。

            采用策略梯度法估计梯度：

            $$\begin{aligned} \nabla_s \Phi(\theta, s) &= \nabla_s \sum_m \left[\mathbb{E}_{x_0,\epsilon,t|m} \mathcal{L}_\theta^t(x_0, \epsilon, m)\right] p(m|s) \\ &= \sum_m \left[\mathbb{E}_{x_0,\epsilon,t|m} \mathcal{L}_\theta^t(x_0, \epsilon, m)\right] \nabla_s p(m|s) \\ &= \sum_m \mathbb{E}_{x_0,\epsilon,t|m} \mathcal{L}_\theta^t(x_0, \epsilon, m) \nabla_s \ln p(m|s) \cdot p(m|s) \\ &= \mathbb{E}_{m\sim p(m|s)} \mathbb{E}_{x_0,\epsilon,t|m} \mathcal{L}_\theta^t(x_0, \epsilon, m) \nabla_s \ln p(m|s). \quad (9) \end{aligned}$$

            因此，$\mathcal{L}_\theta^t(x_0, \epsilon, m) \nabla_s \ln p(m|s)$是$\Phi(\theta, s)$的随机梯度。

        11. 根据式(11)更新θ和s。

            通过投影梯度下降更新θ和s：
            $\theta = \theta - \eta \nabla_\theta \Phi(\theta, s), \quad s = \text{proj}_S\left(s - \eta \nabla_s \Phi(\theta, s)\right), \quad (11)$
            其中$S = \{s \in \mathbb{R}^T : \|s\|_1 \le K_e,\ s \in [0,1]^T\}$ $(K_e = \gamma_e T)$

            投影计算细节

给定向量 z，其在约束区域 $\{s \in \mathbb{R}^T\text{：}\|s\|_1 \leq Ke\text{，}s \in [0,1]^T\}$ 上的投影 s 可按如下方式计算：

$s = \min(1, \max(0, z - v_2^* 1))$。

其中 $v_2^* = \max(0, v_1^*)$，且 v_1^* 是以下方程的解：

$1^\top[\min(1, \max(0, z - v_1^* 1))] - Ke = 0$。(12)

方程（12）可使用二分法高效求解。

12.结束迭代

13.结束训练