Supplementary Material

1 THE RUNGE–KUTTA METHOD FOR DIFFUSION MODELS

We first revisit the Runge–Kutta method (RK4) which solves initial value problems (Runge, 1895; Kutta, 1901; Wikipedia, 2023). Given an initial value problem as $f(\mathbf{x}_t, t) = d\mathbf{x}_t/dt$, where $\mathbf{x}_t$ is associated with time t, the estimation of $\mathbf{x}_t$ at time $(t + h)$ with step size h is computed by

$$\mathbf{x}_{t+h} = \mathbf{x}_t + \frac{h}{6}\left(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4\right) \ , \tag{S1}$$

where

$$\mathbf{k}_1 = f(\mathbf{x}_t, t) \ , \tag{S2}$$

$$\mathbf{k}_2 = f(\mathbf{x}_t + \frac{h}{2}\mathbf{k}_1, t + \frac{h}{2}) \ , \tag{S3}$$

$$\mathbf{k}_3 = f(\mathbf{x}_t + \frac{h}{2}\mathbf{k}_2, t + \frac{h}{2}) \ , \tag{S4}$$

$$\mathbf{k}_4 = f(\mathbf{x}_t + h\mathbf{k}_3 \, , t + h) \ . \tag{S5}$$

For an initial value $\mathbf{x}_0$ at time $(t = 0)$, one can estimate $\mathbf{x}_T$ iteratively by using Eq. (S1) from $(t = 0)$ to the terminate time $(t = T)$.

For the backward process of a diffusion model, the reverse step can be written as

$$\mathbf{x}_{t-h} = g\left(\mathbf{x}_t, \boldsymbol{\epsilon}_{t-h}, t - h\right) \ , \tag{S6}$$

where $g(\cdot)$ refers to the backward step in DDPM (Ho et al., 2020) with Gaussian noises or DDIM (Song et al., 2021) without Gaussian noises and $\boldsymbol{\epsilon}_{t-h}$ is the reversing gradient from $\mathbf{x}_t$ to $\mathbf{x}_{t-h}$, denoted as "model prediction". To apply RK4 to the reverse step, we follow the same rule as Eq. (S1) but change the moving step as

$$\boldsymbol{\epsilon}_{t-h} = \frac{1}{6}\left(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4\right) \ , \tag{S7}$$

where

$$\mathbf{k}_1 = f(\mathbf{x}_t, t) \ , \tag{S8}$$

$$\mathbf{k}_2 = f(g(\mathbf{x}_t, \mathbf{k}_1, t - \lfloor\frac{h}{2}\rfloor), t - \lfloor\frac{h}{2}\rfloor) \ , \tag{S9}$$

$$\mathbf{k}_3 = f(g(\mathbf{x}_t, \mathbf{k}_2, t - \lfloor\frac{h}{2}\rfloor), t - \lfloor\frac{h}{2}\rfloor) \ , \tag{S10}$$

$$\mathbf{k}_4 = f(g(\mathbf{x}_t, \mathbf{k}_3, t - h), t - h) \ , \tag{S11}$$

where $\lfloor\cdot\rfloor$ rounds to the smaller integer due to the discretization of the sampling time space. The difference of moving steps between Eq. (S7) and Eq. (S1) is the multiplication of h or $(h/2)$ to $\mathbf{k}_i$ for $i = \{1, 2, 3, 4\}$. Empirically, applying these multiplications in diffusion sampling leads to strongly unrealistic image generation even with small h. We thus hypothesize that the moving steps are already considered in $g(\cdot)$ 17 because of the joint effect of the sampling coefficients in $g(\cdot)$, $\boldsymbol{\epsilon}_t$, $\boldsymbol{\epsilon}_{t-\lfloor h/2 \rfloor}$, and $\boldsymbol{\epsilon}_{t-h}$.

REFERENCES

Ho, J., Jain, A., and Abbeel, P. (2020). Denoising diffusion probabilistic models. *Conference on Neural Information Processing Systems (NeurIPS)*

Kutta, W. (1901). Beitrag zur naherungsweisen integration totaler differentialgleichungen. *Zeitschrift für Mathematik und Physik*

Runge, C. D. T. (1895). Über die numerische auflösung von differentialgleichungen. *Mathematische Annalen, Springer*

Song, J., Meng, C., and Ermon, S. (2021). Denoising diffusion implicit models. *International Conference on Learning Representations (ICLR)*

Wikipedia (2023). Runge-Kutta methods. `https://en.wikipedia.org/wiki/Runge-Kutta_methods`