

From [22, Section 3], for every θ , we have

$$\mathbb{E}[\|U(\mathbf{x}, k\Delta t; \theta) - U_n(\mathbf{x}, k\Delta t; \theta)\|^2] \leq K_{k\Delta t, U(\cdot, 0), \theta} \lambda_{N+1}^{-1}(\theta), \quad (\text{D.15})$$

where $K_{k\Delta t, U(\cdot, 0), \theta}$ is a constant depending on time $k\Delta t$ and the initial condition $U(\cdot, 0)$, and θ . Without loss of generality, we assume that $K_{k\Delta t, U(\cdot, 0), \theta}$ is non-decreasing in k (otherwise we can replace $K_{k\Delta t, U(\cdot, 0), \theta}$ with $\tilde{K}_{k\Delta t, U(\cdot, 0), \theta} := \max_{1 \leq i \leq k} K_{i\Delta t, U(\cdot, 0), \theta}$). Given the initial condition $U(\mathbf{x}, 0)$ and $P_n U(\mathbf{x}, 0)$, we denote the probability measures of $U(\mathbf{x}, k\Delta t; \theta)$ and $U_n(\mathbf{x}, T; \theta)$ by $\nu_{U(\cdot, 0)}(k\Delta t)$ and $\nu_{n, U_n(\cdot, 0)}(k\Delta t)$, respectively. Moreover, the joint probability measure of $(U(\mathbf{x}, k\Delta t; \theta), U_n(\mathbf{x}, k\Delta t; \theta))$ has marginal distributions $\nu_{U(\cdot, 0)}(k\Delta t)$ and $\nu_{n, U_n(\cdot, 0)}(k\Delta t)$, respectively. From Eq. (D.15), we can deduce that:

$$\begin{aligned} W_2^2(\nu_{U(\cdot, 0)}(k\Delta t), \nu_{n, U_n(\cdot, 0)}(k\Delta t)) &\leq \mathbb{E}[\|U(\mathbf{x}, k\Delta t; \theta) - U_n(\mathbf{x}, k\Delta t; \theta)\|^2] \\ &\leq \sup_{\theta, U(\mathbf{x}, 0)} K_{T, U(\cdot, 0), \theta} \lambda_{N+1}^{-1}(\theta). \end{aligned} \quad (\text{D.16})$$

Furthermore, using the definition of the local squared W_2 distance in Eq. (2.4), we have:

$$\begin{aligned} W_{2, \delta}^{2, e}(U(\cdot, k\Delta t; \theta), U_n(\cdot, k\Delta t; \theta)) &\leq \sup_{\theta, U(\mathbf{x}, 0)} \mathbb{E}[\|U(\mathbf{x}, k\Delta t; \theta) - U_n(\mathbf{x}, k\Delta t; \theta)\|^2] \\ &\leq \sup_{\theta, U(\mathbf{x}, 0)} K_{T, U(\cdot, 0), \theta} \lambda_{N+1}^{-1}(\theta). \end{aligned} \quad (\text{D.17})$$

Similarly, we can conclude that:

$$W_{2, \delta}^{2, e}(\hat{U}(\cdot, k\Delta t; \hat{\theta}), \hat{U}_n(\cdot, k\Delta t; \hat{\theta})) \leq \sup_{\hat{\theta}, U(\mathbf{x}, 0)} K_{T, U(\cdot, 0), \theta} \lambda_{N+1}^{-1}(\hat{\theta}). \quad (\text{D.18})$$

Given the same initial condition $U(\mathbf{x}, 0) = \hat{U}(\mathbf{x}, 0)$, using the triangle inequality of the Wasserstein distance [14], we have

$$\begin{aligned} W_{2, \delta}^{2, e}(U(\mathbf{x}, t; \theta), \hat{U}(\mathbf{x}, t; \hat{\theta})) &\leq 3W_{2, \delta}^{2, e}(U(\cdot, k\Delta t; \theta), U_n(\cdot, k\Delta t; \theta)) \\ &\quad + 3W_{2, \delta}^{2, e}(\hat{U}(\cdot, k\Delta t; \hat{\theta}), \hat{U}_n(\cdot, k\Delta t; \hat{\theta})) + 3W_{2, \delta}^{2, e}(U_n(\mathbf{x}, t; \theta), \hat{U}_n(\mathbf{x}, t; \hat{\theta})) \\ &\leq 3W_{2, \delta}^{2, e}(U_n(\mathbf{x}, t; \theta), \hat{U}_n(\mathbf{x}, t; \hat{\theta})) + 3 \sup_{\theta, U(\mathbf{x}, 0)} K_{T, U(\cdot, 0), \theta} \lambda_{N+1}^{-1}(\theta) \\ &\quad + 3 \sup_{\hat{\theta}, U(\mathbf{x}, 0)} K_{T, U(\cdot, 0), \theta} \lambda_{N+1}^{-1}(\hat{\theta}). \end{aligned} \quad (\text{D.19})$$