NON-ASYMPTOTIC ESTIMATES FOR ACCELERATED HIGH ORDER LANGEVIN MONTE CARLO ALGORITHMS

ARIEL NEUFELD AND YING ZHANG

ABSTRACT. In this paper, we propose two new algorithms, namely aHOLA and aHOLLA, to sample from high-dimensional target distributions with possibly super-linearly growing potentials. We establish non-asymptotic convergence bounds for aHOLA in Wasserstein-1 and Wasserstein-2 distances with rates of convergence equal to $1 + q/2$ and $1/2 + q/4$, respectively, under a local Hölder condition with exponent $q \in (0, 1]$ and a convexity at infinity condition on the potential of the target distribution. Similar results are obtained for aHOLLA under certain global continuity conditions and a dissipativity condition. Crucially, we achieve state-of-the-art rates of convergence of the proposed algorithms in the non-convex setting which are higher than those of the existing algorithms. Numerical experiments are conducted to sample from several distributions and the results support our main findings.

1. INTRODUCTION

We consider the problem of sampling from a high-dimensional target distribution

$$\pi_\beta(\mathrm{d}\theta) \propto \exp(-\beta U(\theta))\,\mathrm{d}\theta, \tag{1}$$

where $\theta \in \mathbb{R}^d$, $\beta > 0$ is the so-called inverse temperature parameter, and $U : \mathbb{R}^d \to \mathbb{R}$ is some (smooth enough) function. The distribution π_β in (1) can be viewed as the invariant measure of the Langevin stochastic differential equation (SDE) given by

$$Z_0 := \theta_0, \quad \mathrm{d}Z_t = -h(Z_t)\,\mathrm{d}t + \sqrt{2\beta^{-1}}\mathrm{d}B_t, \quad t \geq 0, \tag{2}$$

where θ_0 is an $\mathbb{R}^d$-valued random variable, $h := \nabla U$, and $(B_t)_{t\geq 0}$ is a d-dimensional Brownian motion. Thus, to sample from π_β, one may consider using algorithms that track (2). Widely used algorithms of this type include the unadjusted Langevin algorithm (ULA) (or the Langevin Monte Carlo (LMC) algorithm) and the stochastic gradient Langevin dynamics (SGLD) algorithm, which are the Euler discretization of (2). Theoretical guarantees for ULA and SGLD to sample approximately from π_β have been well established in the literature under the conditions that the (stochastic) gradient of the potential of π_β is globally Lipschitz continuous and is strongly convex, see [1, 2, 6, 7, 8, 9]. Recent research focuses on the analysis of ULA and SGLD under weaker conditions so as to accommodate a variety of distributions. To relax the strong convexity condition of h, [4, 5, 23, 24, 27, 28, 29] considered replacing it with certain (local) dissipativity or convexity at infinity condition, and obtain convergence results using techniques developed in [10, 11]. To relax the global Lipschitz condition (in θ) and replace it with a local Lipschitz (or Hölder) condition, certain techniques need to be applied to modify the algorithms (see, e.g., [13, 19, 21, 25] and references therein). This is due to the fact that the absolute moments of the aforementioned algorithms could diverge to infinity at finite time point [12]. In [3], a taming technique proposed in [13, 25] is applied to obtain the tamed ULA algorithm (TULA) and convergence results are obtained under a local Lipschitz condition. Several variants of ULA and SGLD have been developed by applying the taming techniques while their convergence results are established under relaxed conditions to accommodate π_β with super-linearly growing potentials [16, 17, 18, 20, 22].

The algorithms mentioned above are first-order methods which make use of the (stochastic) gradient of the potential of π_β. The state-of-the-art rates of convergence of these algorithms in Wasserstein-2 distances are shown to be 1 in the convex case while they are 1/2 in the non-convex case, which are obtained under certain conditions imposed on the Hessian of the potential. In [7], an LMC algorithm with Ozaki discretization (LMCO) and its variant LMCO' are proposed, which are second-order methods making use of the Hessian of the potential. It is shown in [7] that second-order methods improve on

Key words and phrases. Sampling problem, non-asymptotic estimates, taming technique, super-linearly growing coefficients, high order algorithm.

Financial support by the MOE AcRF Tier 2 Grant MOE-T2EP20222-0013 and the Guangzhou-HKUST(GZ) Joint Funding Program (No. 2024A03J0630) is gratefully acknowledged.

the first-order methods in ill-conditioned cases. However, no improvements are made on the rates of convergences of these methods in Wasserstein distances. In [26], a high order LMC algorithm (HOLA) is developed by applying a taming technique to an order 1.5 numerical scheme introduced in [14]. HOLA makes use of the third derivative of the potential of π_β, and [26] shows that the rate of convergence of HOLA in Wasserstein-2 distance in the convex setting is 3/2, which is higher than that of the first and second-order methods.

In this paper, we mainly consider an algorithm that can be used to sample from π_β with a super-linearly growing potential. To this end, we propose an accelerated HOLA algorithm (aHOLA), which is a variant of HOLA in [26] obtained using a new taming factor. For completeness, we also propose an accelerated high order linear LMC algorithm (aHOLLA) which is the counterpart of aHOLA in the linear case (i.e., in the case where the derivatives of the potential of π_β are growing at most linearly). Crucially, we obtain non-asymptotic error bounds for aHOLA in Wasserstein distances under a local Hölder condition with exponent $q \in (0, 1]$ and a convexity at infinity condition. Our results are applicable to various distributions including, e.g., the double-well potential distribution, which cannot be covered by the corresponding results in Wasserstein distances for HOLA in [26]. In addition, in the non-convex setting, we obtain state-of-the-art rates of convergence of aHOLA in Wasserstein-1 and Wasserstein-2 distances equal to $1 + q/2$ and $1/2 + q/4$, respectively, which improve the rates of convergence of the first and second-order algorithms in the existing literature. We achieve similar convergence results for aHOLLA under certain global Hölder and Lipschitz conditions and a dissipativity condition. To illustrate the applicability of our results, we conduct experiments using aHOLA and aHOLLA to sample from target distributions including a multivariate standard Gaussian distribution, a multivariate Gaussian mixture distribution, and a double-well potential distribution. Numerical results show that the proposed algorithms can sample approximately from the aforementioned distributions, which support our main results.

The rest of the paper is organised as follows. Section 2 presents the setting and assumptions together with the main results for aHOLA where the target distributions have super-linearly growing potentials. Section 3 presents the setting and assumptions together with the main results for aHOLLA where the target distributions have at most linearly growing potentials. Section 4 discusses the related results in the literature in comparison with our work to highlight our contributions. Section 5 illustrates numerical results which support our main findings. Section 6 contains the proofs of main results in Section 2. Finally, Appendix A contains the proofs for the auxiliary results in Sections 2, 5, and 6, while Appendices B and C present tables summarising full expressions of all constants appearing in the statements for aHOLA and aHOLLA, respectively.

We conclude this section by introducing some notation. Let $(\Omega, \mathcal{F}, P)$ be a probability space. We denote by $\mathbb{E}[Z]$ the expectation of a random variable Z. For $1 \leq p < \infty$, L^p is used to denote the usual space of p-integrable real-valued random variables. Fix integers $d, m \geq 1$. For an $\mathbb{R}^d$-valued random variable Z, its law on $\mathcal{B}(\mathbb{R}^d)$, i.e. the Borel sigma-algebra of $\mathbb{R}^d$, is denoted by $\mathcal{L}(Z)$. For a positive real number a, we denote by $\lfloor a \rfloor$ its integer part, and $\lceil a \rceil := \min\{b \in \mathbb{Z} | b \geq a\}$. The Euclidean scalar product is denoted by $\langle \cdot, \cdot \rangle$, with $| \cdot |$ standing for the corresponding norm (where the dimension of the space may vary depending on the context). Denote by $|M|_{\mathsf{F}}$ and M^{T} the Frobenius norm and the transpose of any given matrix $M \in \mathbb{R}^{d\times m}$, respectively. Let $f : \mathbb{R}^d \to \mathbb{R}$ and $g : \mathbb{R}^d \to \mathbb{R}^d$ be twice continuously differentiable functions. Denote by ∇f, $\nabla^2 f$ and Δf the gradient of f, the Hessian of f, and the Laplacian of f, respectively. Denote by $\vec{\Delta} g : \mathbb{R}^d \to \mathbb{R}^d$ the vector Laplacian of g, i.e., for all $\theta \in \mathbb{R}^d$, $\vec{\Delta} g(\theta)$ is a vector in $\mathbb{R}^d$ whose i-th entry is $\sum_{j=1}^d \partial^2_{\theta^{(j)}} g^{(i)}(\theta)$. For any integer $q \geq 1$, let $\mathcal{P}(\mathbb{R}^q)$ denote the set of probability measures on $\mathcal{B}(\mathbb{R}^q)$. For $\mu \in \mathcal{P}(\mathbb{R}^d)$ and for a μ-integrable function $f : \mathbb{R}^d \to \mathbb{R}$, the notation $\mu(f) := \int_{\mathbb{R}^d} f(\theta)\mu(\mathrm{d}\theta)$ is used. For $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$, let $\mathcal{C}(\mu, \nu)$ denote the set of probability measures ζ on $\mathcal{B}(\mathbb{R}^{2d})$ such that its respective marginals are μ, ν. For two Borel probability measures μ and ν defined on $\mathbb{R}^d$ with finite p-th moments, the Wasserstein distance of order $p \geq 1$ is defined as

$$W_p(\mu, \nu) := \left(\inf_{\zeta \in \mathcal{C}(\mu,\nu)} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |\theta - \overline{\theta}|^p \zeta(\mathrm{d}\theta \mathrm{d}\overline{\theta}) \right)^{1/p}.$$

2. Assumptions and main results for aHOLA

Let $U:\mathbb{R}^d\to\mathbb{R}$ be a three times continuously differentiable function satisfying $\int_{\mathbb{R}^d} e^{-\beta U(\theta)}\,\mathrm{d}\theta<\infty$ for any $\beta>0$ and denote by

$$h:=\nabla U,\quad H:=\nabla^2 U,\quad \Upsilon:=\vec{\Delta}(\nabla U)$$

its gradient, Hessian, and vector Laplacian, respectively. Furthermore, define, for any $\beta>0$,

$$\pi_\beta(A):=\frac{\int_A e^{-\beta U(\theta)}\,\mathrm{d}\theta}{\int_{\mathbb{R}^d} e^{-\beta U(\theta)}\,\mathrm{d}\theta},\quad A\in\mathcal{B}(\mathbb{R}^d). \tag{3}$$

Let $\rho\in[2,\infty)\cap\mathbb{N}$ and $q\in(0,1]$. Then, for any $i,j\in\mathbb{N}$, $\lambda>0$, and for any function $f:\mathbb{R}^d\to\mathbb{R}^{i\times j}$, denote by

$$f_\lambda(\theta):=\frac{f(\theta)}{(1+\lambda^{3/2}|\theta|^{3(\rho+q-1)})^{1/3}},\quad \theta\in\mathbb{R}^d. \tag{4}$$

The accelerated high order Langevin Monte Carlo algorithm (aHOLA) for SDE (2) is given by

$$\theta_0^{\mathsf{aHOLA}}:=\theta_0,\quad \theta_{n+1}^{\mathsf{aHOLA}}=\theta_n^{\mathsf{aHOLA}}+\lambda\phi^\lambda(\theta_n^{\mathsf{aHOLA}})+\sqrt{2\lambda\beta^{-1}}\psi^\lambda(\theta_n^{\mathsf{aHOLA}})\xi_{n+1},\quad n\in\mathbb{N}_0, \tag{5}$$

where $\lambda>0$ is the step size, $\beta>0$, $(\xi_n)_{n\in\mathbb{N}_0}$ are i.i.d. standard d-dimensional Gaussian random variables, and where for all $\theta\in\mathbb{R}^d$,

$$\phi^\lambda(\theta):=-h_\lambda(\theta)+(\lambda/2)\left(H_\lambda(\theta)h_\lambda(\theta)-\beta^{-1}\Upsilon_\lambda(\theta)\right), \tag{6}$$

and

$$\psi^\lambda(\theta):=\sqrt{\mathsf{I}_d-\lambda H_\lambda(\theta)+(\lambda^2/3)(H_\lambda(\theta))^2} \tag{7}$$

with I_d being the $d\times d$ identity matrix.

Remark 2.1. *We note that, as mentioned in [7, 14], instead of taking the matrix square root as in (7), $\psi^\lambda(\theta_n^{\mathsf{aHOLA}})\xi_{n+1}$ in (5) can be computed by considering the transformation*

$$\left(\mathsf{I}_d-(1/2)\lambda H_\lambda(\theta_n^{\mathsf{aHOLA}})\right)\widehat{\xi}_{n+1}+(\sqrt{3}/6)\lambda H_\lambda(\theta_n^{\mathsf{aHOLA}})\xi'_{n+1} \tag{8}$$

with $(\widehat{\xi}_n)_{n\in\mathbb{N}_0}$ and $(\xi'_n)_{n\in\mathbb{N}_0}$ being two independent standard Gaussian vectors independent of $(\xi_n)_{n\in\mathbb{N}_0}$ and θ_0. This is due to the fact that (8) has the same distribution as that of (7).

Furthermore, we note that aHOLA (5)-(7) is developed based on the (tamed) order 1.5 scheme of SDE (2), see [14, Chapter 10], which is given by

$$\begin{aligned}\theta_{n+1}^\lambda&=\theta_n^\lambda-\lambda h_\lambda(\theta_n^\lambda)+(\lambda^2/2)\left(H_\lambda(\theta_n^\lambda)h_\lambda(\theta_n^\lambda)-\beta^{-1}\Upsilon_\lambda(\theta_n^\lambda)\right)\\&\quad+\sqrt{2\lambda\beta^{-1}}\widetilde{\xi}_{n+1}-\lambda\sqrt{2\lambda\beta^{-1}}H_\lambda(\theta_n^\lambda)\overline{\xi}_{n+1},\quad n\in\mathbb{N}_0,\end{aligned} \tag{9}$$

where $\theta_0^\lambda:=\theta_0$, $(\widetilde{\xi}_n)_{n\in\mathbb{N}_0}$ is a sequence of i.i.d. standard d-dimensional Gaussian random variables, and $(\overline{\xi}_n)_{n\in\mathbb{N}_0}$ is a sequence of i.i.d. d-dimensional Gaussian random variables with mean 0 and covariance $(1/3)\mathsf{I}_d$. More precisely, for any $n\in\mathbb{N}_0$, $\overline{\xi}_{n+1}=\int_n^{n+1}\int_n^s \mathrm{d}B_r^\lambda\,\mathrm{d}s$, where $B_t^\lambda:=B_{\lambda t}/\sqrt{\lambda}$, $t\geq 0$. One observes that the law of aHOLA (5)-(7) coincides with that of the algorithm (9) at grid points, i.e., for each $n\in\mathbb{N}_0$, $\mathcal{L}(\theta_n^{\mathsf{aHOLA}})=\mathcal{L}(\theta_n^\lambda)$. It is assumed throughout the paper that the $\mathbb{R}^d$-valued random variable θ_0 (the initial condition) is independent of $(\xi_n)_{n\in\mathbb{N}_0}$, $(\widetilde{\xi}_n)_{n\in\mathbb{N}_0}$, and $(\overline{\xi}_n)_{n\in\mathbb{N}_0}$.

2.1. Assumptions.

Let $U:\mathbb{R}^d\to\mathbb{R}$ be three times continuously differentiable, and let $\rho\in[2,\infty)\cap\mathbb{N}$ and $q\in(0,1]$ be fixed. The following assumptions are stated.

We first impose a condition on the initial value θ_0.

Assumption 1. *The initial condition θ_0 is independent of $(\xi_n)_{n\in\mathbb{N}_0}$ and has a finite $16(\rho+1)$-th moment, i.e., $\mathbb{E}[|\theta_0|^{16(\rho+1)}]<\infty$.*

Then, we impose a local Hölder condition on the third derivative of U.

Assumption 2. *There exists $L>0$ such that, for all $i=1,\dots,d$, $\theta,\overline{\theta}\in\mathbb{R}^d$,*

$$|\nabla^2 h^{(i)}(\theta)-\nabla^2 h^{(i)}(\overline{\theta})|\leq L(1+|\theta|+|\overline{\theta}|)^{\rho-2}|\theta-\overline{\theta}|^q.$$

In addition, there exist $K_h,K_H>0$ such that, for all $\theta\in\mathbb{R}^d$,

$$|h(\theta)|\leq K_h(1+|\theta|^{\rho+q}),\quad |H(\theta)|\leq K_H(1+|\theta|^{\rho+q-1}).$$

By Assumption 2, we obtain local Lipschitz (or Hölder) conditions and growth conditions on the first, second, and third derivatives of U as presented in the following remark. The proof is postponed to Appendix A.1.

Remark 2.2. *By Assumption 2, for any $\theta, \overline{\theta} \in \mathbb{R}^d$, $i = 1, \dots, d$, we obtain the following estimates:*

$$|\nabla^2 h^{(i)}(\theta)| \leq \mathsf{K}_0(1 + |\theta|)^{\rho+q-2}, \tag{10}$$

$$|\nabla h^{(i)}(\theta) - \nabla h^{(i)}(\overline{\theta})| \leq \mathsf{K}_0(1 + |\theta| + |\overline{\theta}|)^{\rho+q-2}|\theta - \overline{\theta}|, \tag{11}$$

$$|\nabla h^{(i)}(\theta)| \leq \mathsf{K}_1(1 + |\theta|)^{\rho+q-1},$$

$$|h^{(i)}(\theta) - h^{(i)}(\overline{\theta})| \leq \mathsf{K}_1(1 + |\theta| + |\overline{\theta}|)^{\rho+q-1}|\theta - \overline{\theta}|,$$

$$|h^{(i)}(\theta)| \leq \mathsf{K}_2(1 + |\theta|)^{\rho+q},$$

$$|\Upsilon(\theta) - \Upsilon(\overline{\theta})| \leq d^{3/2}L(1 + |\theta| + |\overline{\theta}|)^{\rho-2}|\theta - \overline{\theta}|^q, \tag{12}$$

$$|\Upsilon(\theta)| \leq \mathsf{K}_{3,d}(1 + |\theta|)^{\rho+q-2},$$

where $\mathsf{K}_0 := 2^{1-q}\max\{L, |\nabla^2 h^{(1)}(0)|, \dots, |\nabla^2 h^{(d)}(0)|\}$, $\mathsf{K}_1 := \max\{\mathsf{K}_0, |\nabla h^{(1)}(0)|, \dots, |\nabla h^{(d)}(0)|\}$, $\mathsf{K}_2 := \max\{\mathsf{K}_1, |h^{(1)}(0)|, \dots, |h^{(d)}(0)|\}$, and $\mathsf{K}_{3,d} := \max\{d^{3/2}L, \Upsilon(0)\}$.

Remark 2.3. *One may notice that in Assumption 2, we assume separately growth conditions of h and H, which could have been deduced directly by using the polynomial Hölder condition of $\nabla^2 h^{(i)}$, $i = 1, \dots, d$, as shown in Remark 2.2. The reason is that the stepsize restriction $\lambda_{\max}$ given in (14) is reciprocally related to the growth constants of h and H (i.e., K_h and K_H, respectively), thus, separately imposing growth conditions allows us to optimize $\lambda_{\max}$. For example, consider the double well potential $U(\theta) = (1/4)|\theta|^4 - (1/2)|\theta|^2$, $\theta \in \mathbb{R}^d$. In this case, it can be shown that, for any $i = 1, \dots, d$, $\nabla^2 h^{(i)}(\theta) = 2\theta \mathsf{e}_i^{\mathsf{T}} + 2\theta^{(i)}\mathsf{I}_d + 2\mathsf{e}_i\theta^{\mathsf{T}}$, where e_i denotes the standard basis vector in $\mathbb{R}^d$ with its i-th element being 1. We have that, for any $\theta, \overline{\theta} \in \mathbb{R}^d$,*

$$|\nabla^2 h^{(i)}(\theta) - \nabla^2 h^{(i)}(\overline{\theta})| \leq 6|\theta - \overline{\theta}|,$$

which, then, by following the same arguments as in the proof of Remark 2.2, the above condition implies that

$$|h(\theta)| \leq 24\sqrt{d}(1 + |\theta|^3), \quad |H(\theta)| \leq 12\sqrt{d}(1 + |\theta|^2).$$

However, since $h(\theta) = \theta(|\theta|^2 - 1)$ and $H(\theta) = \mathsf{I}_d(|\theta|^2 - 1) + 2\theta\theta^{\mathsf{T}}$, we have that

$$|h(\theta)| \leq 2(1 + |\theta|^3), \quad |H(\theta)| \leq 3(1 + |\theta|^2).$$

From the above calculations, we observe that the growth constants of h and H obtained by using the expressions is much smaller than those deduced using the polynomial Hölder condition.

Next, we impose a convexity at infinity condition on U.

Assumption 3. *There exist constants $a, b > 0$, and $\overline{r} \in [0, r)$ with $r := \rho + q - 1$ such that, for all $\theta, \overline{\theta} \in \mathbb{R}^d$,*

$$\langle \theta - \overline{\theta}, h(\theta) - h(\overline{\theta}) \rangle \geq a|\theta - \overline{\theta}|^2(|\theta|^r + |\overline{\theta}|^r) - b|\theta - \overline{\theta}|^2(|\theta|^{\overline{r}} + |\overline{\theta}|^{\overline{r}}).$$

Under Assumptions 2 and 3, we obtain a dissipativity condition on h. The explicit statement is provided below and its proof is postponed to Appendix A.1.

Remark 2.4. *By Assumption 3, for any $\theta \in \mathbb{R}^d$, we obtain*

$$\langle \theta, h(\theta) \rangle \geq \mathsf{a}_{\mathsf{D}}|\theta|^{r+2} - \mathsf{b}_{\mathsf{D}},$$

where $\mathsf{a}_{\mathsf{D}} := a/2$ and $\mathsf{b}_{\mathsf{D}} := (a/2 + b)\mathsf{R}_{\mathsf{D}}^{\overline{r}+2} + |h(0)|^2/(2a)$ with $\mathsf{R}_{\mathsf{D}} := \max\{(4b/a)^{1/(r-\overline{r})}, 2^{1/r}\}$. Moreover, for any $\theta \in \mathbb{R}^d$, we have that

$$\langle \theta, h(\theta) \rangle \geq \mathsf{a}_{\mathsf{D}}|\theta|^2 - \overline{\mathsf{b}}_{\mathsf{D}}, \tag{13}$$

where $\overline{\mathsf{b}}_{\mathsf{D}} := \mathsf{a}_{\mathsf{D}} + \mathsf{b}_{\mathsf{D}}$.

Furthermore, we obtain a one-sided Lipschitz condition on h as presented below. The proof is postponed to Appendix A.1.

Remark 2.5. *By Assumptions 2 and 3, for any* $\theta, \overline{\theta} \in \mathbb{R}^d$*, we obtain*

$$\langle \theta - \overline{\theta}, h(\theta) - h(\overline{\theta}) \rangle \geq -\mathsf{L}_{\mathsf{OS}} |\theta - \overline{\theta}|^2,$$

where $\mathsf{L}_{\mathsf{OS}} := \sqrt{d}\mathsf{K}_1(1 + 2\mathsf{R}_{\mathsf{OS}})^{\rho+q-1} > 0$ *with* $\mathsf{R}_{\mathsf{OS}} := (b/a)^{1/(r-\overline{r})}$.

2.2. Main results.

Denote by

$$\lambda_{\max} := \min\{1, \mathsf{a}_{\mathsf{D}}^{-1}, (19\mathsf{a}_{\mathsf{D}}/240 K_h \max\{K_H, K_h\})^2, (\mathsf{a}_{\mathsf{D}}/120 K_h^2 K_H^2)^{2/3}, \mathsf{a}_{\mathsf{D}}/(480 K_h^2 K_H)\}. \quad (14)$$

Under Assumptions 1, 2, and 3, we obtain the following non-asymptotic error bound in Wasserstein-1 distance between the law of aHOLA (5)-(7) and π_β.

Theorem 2.6. *Let Assumptions 1, 2, and 3 hold. Then, for any* $\beta > 0$*, there exist positive constants* C_0, C_1, C_2 *such that, for any* $n \in \mathbb{N}_0$*,* $0 < \lambda \leq \lambda_{\max}$*,*

$$W_1(\mathcal{L}(\theta_n^{\mathsf{aHOLA}}), \pi_\beta) \leq C_1 e^{-C_0 \lambda n} (\mathbb{E}[|\theta_0|^{16(\rho+1)}] + 1) + C_2 \lambda^{1+q/2},$$

where C_0, C_1, C_2 *are given explicitly in* (35) *(see also Appendix B). Furthermore, for any* $\beta > 0$*,* $\delta > 0$*, if we choose*

$$\lambda \leq \min\left\{(\delta/(2C_2))^{2/(2+q)}, \lambda_{\max}\right\},$$
$$n \geq \max\left\{\delta^{-2/(2+q)}(2C_2)^{2/(2+q)}/C_0, 1/(\lambda_{\max} C_0)\right\} \ln(2C_1(\mathbb{E}[|\theta_0|^{16(\rho+1)}] + 1)/\delta),$$

then, we have $W_1(\mathcal{L}(\theta_n^{\mathsf{aHOLA}}), \pi_\beta) \leq \delta$.

Moreover, we can also obtain a non-asymptotic result in Wasserstein-2 distance between the law of aHOLA (5)-(7) and π_β as presented below.

Corollary 2.7. *Let Assumptions 1, 2, and 3 hold. Then, for any* $\beta > 0$*, there exist positive constants* C_3, C_4, C_5 *such that, for any* $n \in \mathbb{N}_0$*,* $0 < \lambda \leq \lambda_{\max}$*,*

$$W_2(\mathcal{L}(\theta_n^{\mathsf{aHOLA}}), \pi_\beta) \leq C_4 e^{-C_3 \lambda n} (\mathbb{E}[|\theta_0|^{16(\rho+1)}] + 1)^{1/2} + C_5 \lambda^{1/2+q/4}$$

where C_3, C_4, C_5 *are given explicitly in* (40) *(see also Appendix B). Furthermore, for any* $\beta > 0$*,* $\delta > 0$*, if we choose*

$$\lambda \leq \min\left\{(\delta/(2C_5))^{4/(2+q)}, \lambda_{\max}\right\},$$
$$n \geq \max\left\{\delta^{-4/(2+q)}(2C_5)^{4/(2+q)}/C_3, 1/(\lambda_{\max} C_3)\right\} \ln(2C_4(\mathbb{E}[|\theta_0|^{16(\rho+1)}] + 1)^{1/2}/\delta),$$

then, we have $W_2(\mathcal{L}(\theta_n^{\mathsf{aHOLA}}), \pi_\beta) \leq \delta$.

The proofs of Theorem 2.6 and Corollary 2.7 are deferred in Section 6.

3. Assumptions and main results for aHOLLA

In this section, we consider the case where the derivatives of U is growing at most linearly. The setting in this case is similar to that described in Section 2 except that there is no need to use tamed coefficients as in aHOLA (5)-(7). More precisely, with the assumptions and notation defined up to (3), the aHOLA algorithm in the linear setting, which we name the accelerated high order linear Langevin Monte Carlo algorithm (aHOLLA), is given by

$$\Theta_0^{\mathsf{aHOLLA}} := \theta_0, \quad \Theta_{n+1}^{\mathsf{aHOLLA}} = \Theta_n^{\mathsf{aHOLLA}} + \lambda \phi_{\mathsf{Lin}}^{\lambda}(\Theta_n^{\mathsf{aHOLLA}}) + \sqrt{2\lambda\beta^{-1}} \psi_{\mathsf{Lin}}^{\lambda}(\Theta_n^{\mathsf{aHOLLA}}) \xi_{n+1}, \quad n \in \mathbb{N}_0, \quad (15)$$

where for all $\theta \in \mathbb{R}^d$,

$$\phi_{\mathsf{Lin}}^{\lambda}(\theta) := -h(\theta) + (\lambda/2)\left(H(\theta)h(\theta) - \beta^{-1}\Upsilon(\theta)\right), \quad (16)$$

and

$$\psi_{\mathsf{Lin}}^{\lambda}(\theta) := \sqrt{\mathsf{I}_d - \lambda H(\theta) + (\lambda^2/3)(H(\theta))^2}. \quad (17)$$

Remark 3.1. *The corresponding order 1.5 scheme of aHOLLA* (15)-(17) *in the linear case is given by*

$$\begin{aligned} \Theta_0^{\lambda} := \theta_0, \quad \Theta_{n+1}^{\lambda} = \Theta_n^{\lambda} - \lambda h(\Theta_n^{\lambda}) + (\lambda^2/2)\left(H(\Theta_n^{\lambda})h(\Theta_n^{\lambda}) - \beta^{-1}\Upsilon(\Theta_n^{\lambda})\right) \\ + \sqrt{2\lambda\beta^{-1}}\widetilde{\xi}_{n+1} - \lambda\sqrt{2\lambda\beta^{-1}} H(\Theta_n^{\lambda})\overline{\xi}_{n+1}, \quad n \in \mathbb{N}_0. \end{aligned} \quad (18)$$

Then, it holds that $\mathcal{L}(\Theta_n^{\mathsf{aHOLLA}}) = \mathcal{L}(\Theta_n^{\lambda})$*, for each* $n \in \mathbb{N}_0$.

We provide below the assumptions and main results for the aHOLLA algorithm (15)-(17).

3.1. **Assumptions.** Let $U : \mathbb{R}^d \to \mathbb{R}$ be three times continuously differentiable and let $q \in (0, 1]$ be fixed. The following assumptions are stated, which can be viewed as counterparts to those stated in Section 2.

We first impose assumptions on the initial condition θ_0.

Assumption 4. *The initial condition θ_0 is independent of $(\xi_n)_{n\in\mathbb{N}_0}$ and has a finite fourth moment, i.e., $\mathbb{E}[|\theta_0|^4] < \infty$.*

Then, we impose conditions on the first, second, and third derivatives of U.

Assumption 5. *There exists $\overline{L}_1 > 0$ such that, for all $i = 1, \dots, d$, $\theta, \overline{\theta} \in \mathbb{R}^d$,*

$$|\nabla^2 h^{(i)}(\theta) - \nabla^2 h^{(i)}(\overline{\theta})| \le \overline{L}_1 |\theta - \overline{\theta}|^q.$$

In addition, there exist $\overline{L}_2, \overline{L}_3 > 0$ such that, for all $\theta, \overline{\theta} \in \mathbb{R}^d$,

$$|H(\theta) - H(\overline{\theta})| \le \overline{L}_2 |\theta - \overline{\theta}|,$$
$$|h(\theta) - h(\overline{\theta})| \le \overline{L}_3 |\theta - \overline{\theta}|.$$

Remark 3.2. *By Assumption 5, for any $\theta, \overline{\theta} \in \mathbb{R}^d$, $i = 1, \dots, d$, we obtain the following estimates:*

$$|\nabla^2 h^{(i)}(\theta)| \le \overline{\mathsf{K}}_0(1 + |\theta|^q), \qquad |H(\theta)\overline{\theta}| \le \overline{L}_3|\overline{\theta}|, \qquad |h(\theta)| \le \overline{\mathsf{K}}_1(1 + |\theta|),$$
$$|\Upsilon(\theta) - \Upsilon(\overline{\theta})| \le d^{3/2}\overline{L}_1|\theta - \overline{\theta}|^q, \qquad |\Upsilon(\theta)| \le d\overline{L}_2,$$

where $\overline{\mathsf{K}}_0 := \max\{\overline{L}_1, |\nabla^2 h^{(1)}(0)|, \dots, |\nabla^2 h^{(d)}(0)|\}$, $\overline{\mathsf{K}}_1 := \max\{\overline{L}_3, |h(0)|\}$.

Finally, we impose a dissipativity condition on h.

Assumption 6. *There exist constants $\overline{a}, \overline{b} > 0$ such that, for all $\theta \in \mathbb{R}^d$,*

$$\langle \theta, h(\theta) \rangle \ge \overline{a}|\theta|^2 - \overline{b}.$$

3.2. **Main results.** Denote by

$$\overline{\lambda}_{\max} := \min\left\{1, 1/\overline{a}, \overline{a}/(16\overline{L}_3\overline{\mathsf{K}}_1), \overline{a}/(16\overline{\mathsf{K}}_1^2), \overline{a}^{1/3}/(4\overline{L}_3^2\overline{\mathsf{K}}_1^2)^{1/3}, \overline{a}^{1/2}/(16\overline{L}_3\overline{\mathsf{K}}_1^2)^{1/3}\right\}. \tag{19}$$

Then, under Assumptions 4, 5, and 6, we deduce the following non-asymptotic convergence results in Wasserstein distances for aHOLLA (15)-(17).

Theorem 3.3. *Let Assumptions 4, 5, and 6 hold. Then, for any $\beta > 0$, there exist positive constants $C_{\mathsf{Lin},0}, C_{\mathsf{Lin},1}, C_{\mathsf{Lin},2}$ such that, for any $n \in \mathbb{N}_0$, $0 < \lambda \le \overline{\lambda}_{\max}$,*

$$W_1(\mathcal{L}(\Theta_n^{\mathsf{aHOLLA}}), \pi_\beta) \le C_{\mathsf{Lin},1} e^{-C_{\mathsf{Lin},0}\lambda n}(\mathbb{E}[|\theta_0|^4] + 1) + C_{\mathsf{Lin},2}\lambda^{1+q/2},$$

where $C_{\mathsf{Lin},0}, C_{\mathsf{Lin},1}, C_{\mathsf{Lin},2}$ are given explicitly in Appendix C. Furthermore, for any $\beta > 0$, $\delta > 0$, if we choose

$$\lambda \le \min\left\{(\delta/2C_{\mathsf{Lin},2})^{2/(2+q)}, \overline{\lambda}_{\max}\right\},$$
$$n \ge \max\left\{(2C_{\mathsf{Lin},2}/\delta)^{2/(2+q)}/C_{\mathsf{Lin},0}, 1/(\overline{\lambda}_{\max}C_{\mathsf{Lin},0})\right\} \ln(2C_{\mathsf{Lin},1}(\mathbb{E}[|\theta_0|^4] + 1)/\delta),$$

then, we have $W_1(\mathcal{L}(\Theta_n^{\mathsf{aHOLLA}}), \pi_\beta) \le \delta$.

Corollary 3.4. *Let Assumptions 4, 5, and 6 hold. Then, for any $\beta > 0$, there exist positive constants $C_{\mathsf{Lin},3}, C_{\mathsf{Lin},4}, C_{\mathsf{Lin},5}$ such that, for any $n \in \mathbb{N}_0$, $0 < \lambda \le \overline{\lambda}_{\max}$,*

$$W_2(\mathcal{L}(\Theta_n^{\mathsf{aHOLLA}}), \pi_\beta) \le C_{\mathsf{Lin},4} e^{-C_{\mathsf{Lin},3}\lambda n}(\mathbb{E}[|\theta_0|^4] + 1)^{1/2} + C_{\mathsf{Lin},5}\lambda^{1/2+q/4}$$

where $C_{\mathsf{Lin},3}, C_{\mathsf{Lin},4}, C_{\mathsf{Lin},5}$ are given explicitly in Appendix C. Furthermore, for any $\beta > 0$, $\delta > 0$, if we choose

$$\lambda \le \min\left\{(\delta/2C_{\mathsf{Lin},5})^{4/(2+q)}, \overline{\lambda}_{\max}\right\},$$
$$n \ge \max\left\{(2C_{\mathsf{Lin},5}/\delta)^{4/(2+q)}/C_{\mathsf{Lin},3}, 1/(\overline{\lambda}_{\max}C_{\mathsf{Lin},3})\right\} \ln(2C_{\mathsf{Lin},4}(\mathbb{E}[|\theta_0|^4] + 1)^{1/2}/\delta),$$

then, we have $W_2(\mathcal{L}(\Theta_n^{\mathsf{aHOLLA}}), \pi_\beta) \le \delta$.

The proofs of Theorem 3.3 and Corollary 3.4 follow the same ideas as in the proofs of Theorem 2.6 and Corollary 2.7 explained in Section 6, and hence are deferred in Appendix D.

4. Related work and comparison

In this section, we compare our results in Theorem 2.6, Corollary 2.7, Theorem 3.3, and Corollary 3.4 with the most relevant work [26, 22] to highlight our contributions.

In [26], the authors propose the HOLA algorithm [26, (2)] and its counterpart in the linear case [26, (8)], and obtain non-asymptotic convergence bounds for the two algorithms in Wasserstein-2 distance under the assumptions [26, H1-H3] and [26, H3-H6], respectively. [26, H1] assumes that the norm of $\nabla U(\theta)$ tends to infinity as $|\theta|$ tends to infinity, and that $\langle \theta, \nabla U(\theta)\rangle$ is bounded from below by a positive number when $|\theta|$ gets large. This condition is similar to (13) in Remark 2.4 deduced using Assumption 3, which is key in establishing the moment estimates of aHOLA. [26, H2] imposes a local Hölder condition on the third derivative of U, which is the same condition as in our Assumption 2. We note that the growth conditions imposed on h and H in Assumption 2 are merely to relax the stepsize restriction in (14), see Remark 2.3. [26, H3] assumes a strong convexity condition of ∇U, while in our work, we discard such an assumption and only assume a convexity at infinity condition in Assumption 3 for aHOLA. We refer to [22, Remark 2.3] for a detailed comparison between [26, H3] and our Assumption 3. In addition, in the linear case, we impose a dissipativity condition in Assumption 6 for aHOLLA which can be deduced from [26, H3] and is thus a weaker condition compared to [26, H3]. It is worth highlighting that both conditions, i.e., Assumptions 3 and 6, can accommodate distributions with non-convex potentials that cannot be covered by [26, H3]. [26, H4-H6] impose global Lipschitz conditions on the derivatives of U, which are the same conditions as those in Assumption 5 except that we consider a global Hölder condition for the third derivative of U which covers more general cases. To obtain our main results, we further impose conditions on θ_0 in Assumptions 1 and 4 which are easily satisfied by various choices in practical implementations. Consequently, under Assumptions 1-3 and Assumptions 4-6, we obtain non-asymptotic convergence results in Wasserstein-2 distance which are applicable to a wide range of distributions that cannot be covered by those established in [26] including, e.g., the double-well distribution.

In [22], the authors propose an mTULA algorithm and obtain non-asymptotic error bounds in Wasserstein-1 and Wasserstein-2 distances under [22, Assumptions 1-4] with rates of convergence equal to 1 and 1/2, respectively. [22, Assumption 1] imposes an initial condition on θ_0, which is similar to our Assumptions 1 and 4. [22, Assumptions 2 and 4] imposes local Lipschitz conditions on the first and second derivative of U which become corresponding global Lipschitz conditions in the linear setting. In our work, we impose a local Hölder condition on the third derivative of U in Assumption 2 in the super-linear case for aHOLA while we replace it with a global Hölder condition in Assumption 5 in the linear case for aHOLLA. This is due to the fact that we utilise high order derivatives of U in the design of aHOLA (5)-(7) and aHOLLA (15)-(17). [22, Assumption 3] assumes a convexity at infinity condition and a dissipativity condition which are the same as those in Assumptions 3 and 6, respectively. We highlight that, in both linear and super-linear cases, we obtain non-asymptotic convergence results in Wasserstein-1 and Wasserstein-2 distances with rates of convergence equal to $1 + q/2$ and $1/2 + q/4$, respectively, which are higher than those obtained in [22] due to the use of high order derivatives of U in our proposed algorithms.

5. Numerical experiments

In this section, we illustrate the applicability of our results in Theorem 2.6, Corollary 2.7, Theorem 3.3, and Corollary 3.4.

5.1. Sampling from target distributions.

We use aHOLA (5)-(7) and aHOLLA (15)-(17) to draw samples from various distributions[1]. More precisely, we consider the following high-dimensional target distributions:

(i) a multivariate standard Gaussian distribution with its potential given by

$$U(\theta) := \frac{1}{2}|\theta|^2, \quad \theta \in \mathbb{R}^d, \tag{20}$$

(ii) a multivariate Gaussian mixture distribution with its potential given by

$$U(\theta) := \frac{1}{2}|\theta - \hat{a}|^2 - \log(1 + \exp(-2\langle x, \hat{a}\rangle)), \quad \theta \in \mathbb{R}^d, \tag{21}$$

[1] The python code is available at `https://github.com/tracyyingzhang/aHOLA`.

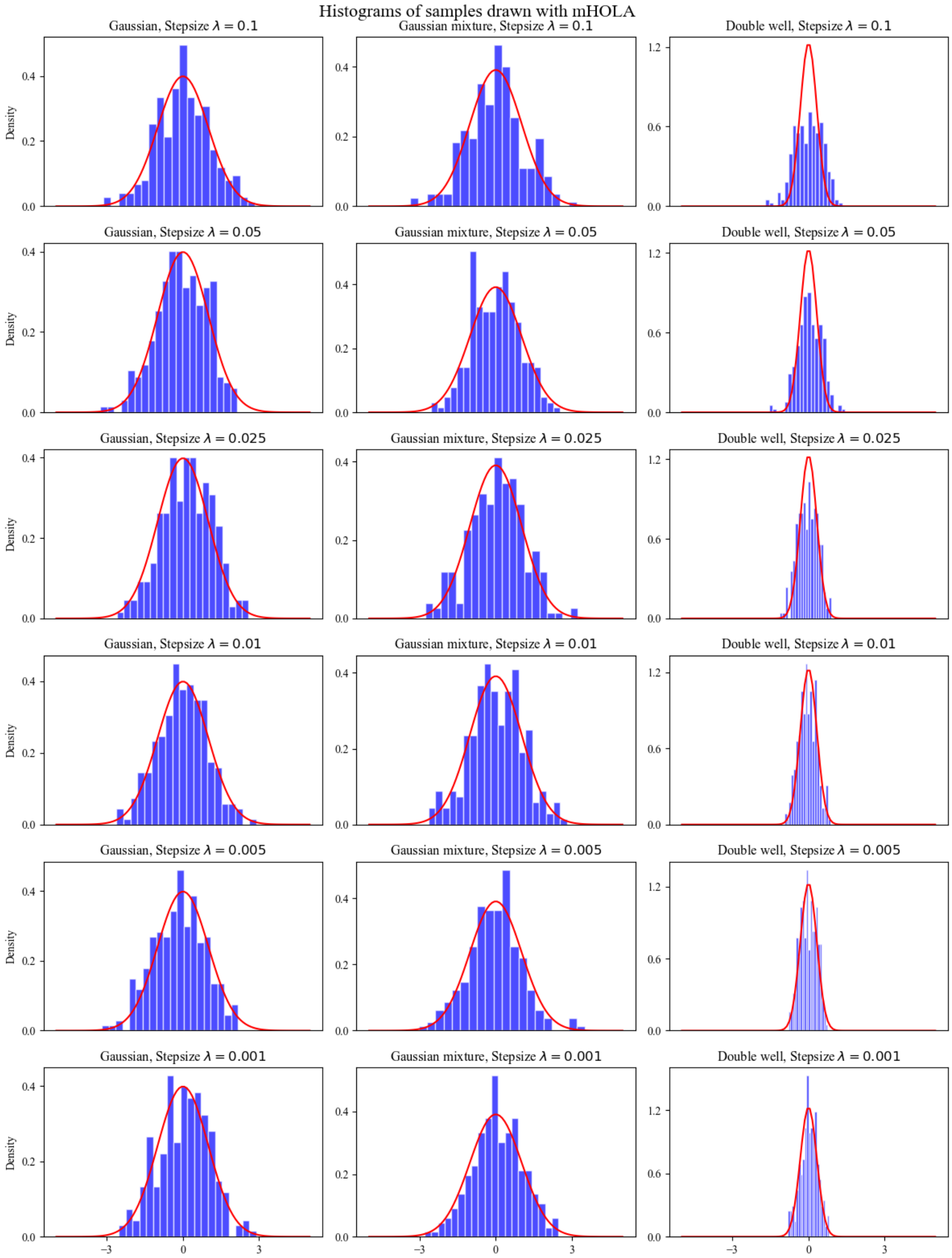

FIGURE 1. Normalised histograms of the first components of the samples drawn using aHOLA and aHOLLA.

where $\hat{a} \in \mathbb{R}^d$ is a given vector with $|\hat{a}| > 1$, and

(iii) a double-well potential distribution with its potential given by

$$U(\theta) := \frac{1}{4}|\theta|^4 - \frac{1}{2}|\theta|^2, \quad \theta \in \mathbb{R}^d. \tag{22}$$

We note that the first two distributions in (i) and (ii) have potentials (20) and (21) whose gradients are growing at most linearly, while the last distribution in (iii) has potential (22) whose gradient is growing super-linearly. We show in the following proposition that the examples (i) and (ii) satisfy Assumptions 5 and 6, and (iii) satisfies Assumptions 2 and 3.

Proposition 5.1. *The functions $U : \mathbb{R}^d \to \mathbb{R}$ given in* (20) *and* (21) *satisfy Assumptions 5 and 6 while that given in* (22) *satisfies Assumptions 2 and 3.*

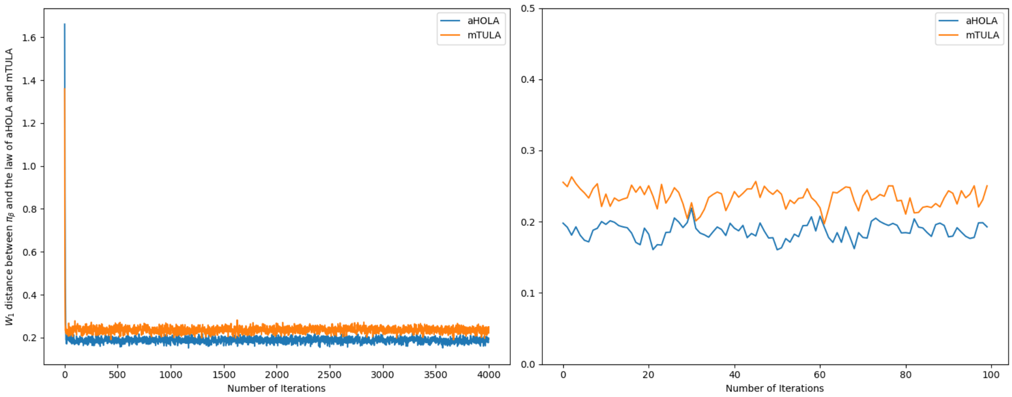

FIGURE 2. (Left) Wasserstein-1 distance between π_β^{dw} and the law of aHOLA and mTULA for the entire time period $\lambda n = 400$ with $n = 4000$. (Right) Wasserstein-1 distance between π_β^{dw} and the law of aHOLA and mTULA for the last 100 iterations.

Proof. See Appendix A.2. □

For numerical experiments, we set $d = 100$. The initial value θ_0 is set to be the null vector in $\mathbb{R}^d$ so as to satisfy Assumptions 1 and 4. This, together with Proposition 5.1, ensures that the proposed algorithms aHOLA (5)-(7) and aHOLLA (15)-(17) can sample approximately from the target distributions due to our main results. In addition, we set $\beta = 1$ and consider different stepsizes $\lambda = \{0.001, 0.005, 0.01, 0.025, 0.05, 0.1\}$. The number of iterations n is chosen such that $\lambda n = 400$ is fixed. For each distribution and stepsize mentioned above, we run 250 independent aHOLA or aHOLLA Markov chains and collect outputs from the last iterations of these Markov chains.

To illustrate the performance of the proposed algorithms, in Figure 1, we plot the normalised histograms of the first components of the 250 samples generated using the method described above, and compare the histograms obtained numerically with their corresponding marginal density functions which are red lines superimposed on the plots. We note that the marginal density functions for the first component of a multivariate standard Gaussian distribution, a multivariate Gaussian mixture distribution, and a doubel-well potential distribution are given by

$$
\begin{aligned}
\theta^{(1)} &\ni \frac{1}{\sqrt{2\pi}} e^{-(\theta^{(1)})^2/2}, \\
\theta^{(1)} &\ni \frac{1}{2\sqrt{2\pi}} \left(e^{-(\theta^{(1)}-\hat{a}^{(1)})^2/2} + e^{-(\theta^{(1)}+\hat{a}^{(1)})^2/2} \right), \\
\theta^{(1)} &\ni \frac{\Gamma(d/2)}{\sqrt{\pi}\Gamma((d-1)/2)} \frac{\int_0^\infty x^{(d-3)/2} \exp\left(-(x+(\theta^{(1)})^2)^2/4 + (x+(\theta^{(1)})^2)/2\right) \,\mathrm{d}x}{\int_0^\infty x^{d/2-1} \exp\left(-x^2/4 + x/2\right) \,\mathrm{d}x},
\end{aligned} \tag{23}
$$

respectively, where we denote by $\theta^{(1)}, \hat{a}^{(1)} \in \mathbb{R}$ the first component of $\theta, \hat{a} \in \mathbb{R}^d$, respectively, and where Γ denotes the gamma function. In addition, for the multivariate Gaussian mixture distribution, we set $|\hat{a}| = 2$ in the experiments with all its components being equal.

In Figure 1, we observe that the histograms obtained using samples generated by aHOLA (5)-(7) and aHOLLA (15)-(17) are close to their corresponding theoretical marginal density functions. This illustrates that the proposed algorithms can generate samples approximately from given target distributions with appropriately chosen stepsizes. The numerical results support our main findings, i.e., Theorem 2.6, Corollary 2.7, Theorem 3.3, and Corollary 3.4.

5.2. **Comparsion of aHOLA and mTULA.** Denote by π_β^{dw} the marginal distribution of the double-well potential distribution with its density given in (23). We show the convergence of aHOLA (5)-(7) in Wasserstein-1 distance and compare the performance of aHOLA with that of mTULA proposed in [22]. To this end, we conduct numerical experiments with $d = 100$, $\theta_0 = (2, 2, \ldots, 2) \in \mathbb{R}^{100}$, $\lambda = 0.1$, $n = 4000$, and the number of independent aHOLA and mTULA Markov chians equal to 1000.

Figure 2 depicts Wasserstein-1 distances between π_β^{dw} and the law of aHOLA and mTULA algorithms. On the one hand, it shows that the Wasserstein-1 distance vanishes as n gets large. This indicates that aHOLA converges to π_β^{dw} in Wasserstein distance which supports our finding in Theorem 2.6 and which cannot be covered by the results in [26] for the HOLA algorithm. On the other hand, Figure 2 shows that, for large n, the Wasserstein-1 distance between π_β^{dw} and the law of aHOLA is smaller than that between π_β^{dw} and the law of mTULA. This empirically demonstrates that our proposed algorithm aHOLA generates samples from the distribution that is closer to π_β^{dw} compared to mTULA in Wasserstein-1 distance, which is consistent with Theorem 2.6 due to the higher convergence rate of aHOLA.

6. Proof of main results for aHOLA

In this section, we provide the proofs for Theorem 2.6 and Corollary 2.7. We first introduce auxiliary processes which we use throughout the convergence analysis. Then, we provide moment estimates for the newly introduced processes, which are followed by the detailed proofs for the main results. We postpone the proofs for the results presented in this section to Appendix A.3.

6.1. Auxiliary processes.

Fix $\beta > 0$. Consider the Langevin SDE $(Z_t)_{t\geq 0}$ given by

$$Z_0 := \theta_0, \quad \mathrm{d}Z_t = -h\left(Z_t\right)\mathrm{d}t + \sqrt{2\beta^{-1}}\mathrm{d}B_t, \tag{24}$$

where $(B_t)_{t\geq 0}$ is a d-dimensional Brownian motion on $(\Omega, \mathcal{F}, P)$ with its completed natural filtration denoted by $(\mathcal{F}_t)_{t\geq 0}$. Moreover, we assume that $(\mathcal{F}_t)_{t\geq 0}$ is independent of $\sigma(\theta_0)$. Under Assumptions 2 and 3, and by Remarks 2.4 and 2.5, we note that the Langevin SDE (24) admits a unique solution, which is adapted to $\mathcal{F}_t \vee \sigma(\theta_0)$, $t \geq 0$, due to [15, Theorem 1]. Its $2p$-th moment estimate with $p \in \mathbb{N}$ is provided in Lemma A.1, which can be used to further deduce the $2p$-th moment estimate of π_β (3).

For each $\lambda > 0$, recall $B_t^\lambda := B_{\lambda t}/\sqrt{\lambda}, t \geq 0$. Denote by $(\mathcal{F}_t^\lambda)_{t\geq 0}$ with $\mathcal{F}_t^\lambda := \mathcal{F}_{\lambda t}$, $t \geq 0$, its completed natural filtration, which is independent of $\sigma(\theta_0)$. Moreover, we denote by $Z_t^\lambda := Z_{\lambda t}$, $t \geq 0$, the time-changed version of Langevin SDE (24), which is given by

$$Z_0^\lambda := \theta_0, \quad \mathrm{d}Z_t^\lambda = -\lambda h(Z_t^\lambda)\,\mathrm{d}t + \sqrt{2\lambda\beta^{-1}}\,\mathrm{d}B_t^\lambda. \tag{25}$$

Furthermore, we denote by $(\widetilde{\theta}_t^\lambda)_{t\geq 0}$ the continuous-time interpolation of aHOLA (5)-(7) given by

$$\widetilde{\theta}_0^\lambda := \theta_0, \quad \mathrm{d}\widetilde{\theta}_t^\lambda = \lambda\phi^\lambda(\widetilde{\theta}_{\lfloor t\rfloor}^\lambda)\,\mathrm{d}t + \sqrt{2\lambda\beta^{-1}}\psi^\lambda(\widetilde{\theta}_{\lfloor t\rfloor}^\lambda)\,\mathrm{d}B_t^\lambda, \tag{26}$$

where ϕ^λ and ψ^λ are defined in (6) and (7), respectively.

Remark 6.1. *Similarly, denote by $(\overline{\theta}_t^\lambda)_{t\geq 0}$ the continuous-time interpolation of the order 1.5 scheme (9) given by*

$$\begin{aligned} \mathrm{d}\overline{\theta}_t^\lambda &= -\lambda h_\lambda(\overline{\theta}_{\lfloor t\rfloor}^\lambda)\,\mathrm{d}t + \lambda^2 \int_{\lfloor t\rfloor}^t \left(H_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda)h_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda) - \beta^{-1}\Upsilon_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda)\right)\mathrm{d}s\,\mathrm{d}t \\ &\quad - \lambda\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t H_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda)\,\mathrm{d}B_s^\lambda\,\mathrm{d}t + \sqrt{2\lambda\beta^{-1}}\,\mathrm{d}B_t^\lambda \end{aligned} \tag{27}$$

with $\overline{\theta}_0^\lambda := \theta_0$. We note that $\mathcal{L}(\widetilde{\theta}_n^\lambda) = \mathcal{L}(\theta_n^{\mathsf{aHOLA}}) = \mathcal{L}(\theta_n^\lambda) = \mathcal{L}(\overline{\theta}_n^\lambda)$, for each $n \in \mathbb{N}_0$.

Finally, for any $s \geq 0$, consider the continuous-time process $(\zeta_t^{s,v,\lambda})_{t\geq s}$ defined by

$$\zeta_s^{s,v,\lambda} := v \in \mathbb{R}^d, \quad \mathrm{d}\zeta_t^{s,v,\lambda} = -\lambda h(\zeta_t^{s,v,\lambda})\,\mathrm{d}t + \sqrt{2\lambda\beta^{-1}}\,\mathrm{d}B_t^\lambda. \tag{28}$$

Definition 6.2. *Fix $\lambda > 0$. Define $T \equiv T(\lambda) := \lfloor 1/\lambda \rfloor$. Then, for any $n \in \mathbb{N}_0$ and $t \geq nT$, define*

$$\overline{\zeta}_t^{\lambda,n} := \zeta_t^{nT,\overline{\theta}_{nT}^\lambda,\lambda}.$$

6.2. Moment estimates.

We first introduce the following Lyapunov functions: for each $p \in [2,\infty) \cap \mathbb{N}$, define $V_p(\theta) := (1+|\theta|^2)^{p/2}$, for all $\theta \in \mathbb{R}^d$, as well as $\mathrm{v}_p(w) := (1+w^2)^{p/2}$, for all $w \geq 0$. We observe that V_p is twice continuously differentiable and satisfies:

$$\sup_{\theta\in\mathbb{R}^d} |\nabla V_p(\theta)|/V_p(\theta) < \infty, \quad \lim_{|\theta|\to\infty} \nabla V_p(\theta)/V_p(\theta) = 0. \tag{29}$$

Furthermore, we denote by $\mathcal{P}_{V_p}(\mathbb{R}^d)$ the set of probability measures $\mu \in \mathcal{P}(\mathbb{R}^d)$ which satisfies $\int_{\mathbb{R}^d} V_p(\theta)\,\mu(\mathrm{d}\theta) < \infty$.

Next, we establish moment estimates for $(\widetilde{\theta}_t^\lambda)_{t\geq 0}$ given in (26). The results with explicit constants are provided below. We note that for any $p \in [2,\infty) \cap \mathbb{N}$ and $t \geq 0$, we have that $\mathbb{E}[|\widetilde{\theta}_t^\lambda|^{2p}] = \mathbb{E}[|\overline{\theta}_t^\lambda|^{2p}]$.

Lemma 6.3. *Let Assumptions 1, 2, and 3 hold. Then, we obtain the following estimates:*

(i) *For any* $0 < \lambda \leq \lambda_{\max}$, $n \in \mathbb{N}_0$, *and* $t \in (n, n+1]$,

$$\mathbb{E}\left[|\widetilde{\theta}_t^\lambda|^2\right] \leq (1 - \lambda(t-n)\mathsf{a}_\mathsf{D}\kappa)\,(1-\lambda\mathsf{a}_\mathsf{D}\kappa)^n\, \mathbb{E}\left[|\theta_0|^2\right] + \mathsf{c}_0\left(1 + 1/(\mathsf{a}_\mathsf{D}\kappa)\right),$$

where the constants c_0, κ *are given explicitly in (73). In particular, the above inequality implies* $\sup_{t\geq 0} \mathbb{E}\left[|\widetilde{\theta}_t^\lambda|^2\right] \leq \mathbb{E}\left[|\theta_0|^2\right] + \mathsf{c}_0(1 + 1/(\mathsf{a}_\mathsf{D}\kappa)) < \infty$.

(ii) *For any* $p \in [2,\infty) \cap \mathbb{N}$, $0 < \lambda \leq \lambda_{\max}$, $n \in \mathbb{N}_0$, *and* $t \in (n, n+1]$,

$$\mathbb{E}\left[|\widetilde{\theta}_t^\lambda|^{2p}\right] \leq (1 - \lambda(t-n)\mathsf{a}_\mathsf{D}\kappa)\,(1-\lambda\mathsf{a}_\mathsf{D}\kappa)^n\, \mathbb{E}\left[|\theta_0|^{2p}\right] + \mathsf{c}_p\left(1 + 1/(\mathsf{a}_\mathsf{D}\kappa)\right),$$

where κ *is given explicitly in (73) and* c_p *is given in (81). In particular, the above estimate implies* $\sup_{t\geq 0} \mathbb{E}\left[|\widetilde{\theta}_t^\lambda|^{2p}\right] \leq \mathbb{E}\left[|\theta_0|^{2p}\right] + \mathsf{c}_p\left(1 + 1/(\mathsf{a}_\mathsf{D}\kappa)\right) < \infty$.

Proof. See Appendix A.3. □

We provide below a drift condition for V_p (defined in the beginning of Section 6.2), which is key to obtain the moment estimates of $(\bar{\zeta}_t^{\lambda,n})_{t\geq nT}$ defined in Definition 6.2.

Lemma 6.4. *Let Assumptions 1, 2, and 3 hold. Then, for any* $p \in [2,\infty) \cap \mathbb{N}$, $\theta \in \mathbb{R}^d$, *we obtain*

$$\Delta V_p(\theta)/\beta - \langle \nabla V_p(\theta), h(\theta)\rangle \leq -\mathsf{c}_{V,1}(p) V_p(\theta) + \mathsf{c}_{V,2}(p),$$

where $\mathsf{c}_{V,1}(p) := \mathsf{a}_\mathsf{D} p/4$, $\mathsf{c}_{V,2}(p) := (3/4)\mathsf{a}_\mathsf{D} p \mathrm{v}_p(\mathsf{M}_V(p))$ *with*

$$\mathsf{M}_V(p) := (1/3 + 4\bar{\mathsf{b}}_\mathsf{D}/(3\mathsf{a}_\mathsf{D}) + 4d/(3\mathsf{a}_\mathsf{D}\beta) + 4(p-2)/(3\mathsf{a}_\mathsf{D}\beta))^{1/2}.$$

Proof. See [4, Lemma 3.5]. □

By applying Lemma 6.3 and 6.4, we obtain the second and the fourth moment estimates for $(\bar{\zeta}_t^{\lambda,n})_{t\geq nT}$.

Lemma 6.5. *Let Assumptions 1, 2, and 3 hold. Then, for any* $p \in \mathbb{N}$, $0 < \lambda \leq \lambda_{\max}$, $n \in \mathbb{N}_0$, *and* $t \geq nT$, *we obtain*

$$\mathbb{E}[V_{2p}(\bar{\zeta}_t^{\lambda,n})] \leq 2^{p-1} e^{-\lambda \mathsf{a}_\mathsf{D}\min\{\kappa,1/2\}t}\mathbb{E}[|\theta_0|^{2p}] + 2^{p-1}\left(\mathsf{c}_p\left(1 + 1/(\mathsf{a}_\mathsf{D}\kappa)\right) + 1\right) + 3\mathrm{v}_{2p}(\mathsf{M}_V(2p)),$$

where κ *and* c_p *are given in (73) and (81) (see also Lemma 6.3) and* $\mathsf{M}_V(2p)$ *is given in Lemma 6.4.*

Proof. See [22, Corollary 4.6]. □

6.3. **Proof of main results.** In this section, we present key results used to obtain Theorem 2.6. We consider establishing a non-asymptotic error bound in Wasserstein-1 distance between the law of $\bar{\theta}_t^\lambda$ given in (27) and π_β, i.e., $W_1(\mathcal{L}(\bar{\theta}_t^\lambda), \pi_\beta)$, for any $n \in \mathbb{N}_0$ and $t \in (nT, (n+1)T]$. This and the fact that $W_1(\mathcal{L}(\theta_n^{\mathsf{aHOLA}}), \pi_\beta) = W_1(\mathcal{L}(\bar{\theta}_t^\lambda), \pi_\beta)$ holds at any grid point yields the desired result. To this end, we split $W_1(\mathcal{L}(\bar{\theta}_t^\lambda), \pi_\beta)$ by using the law of $\bar{\zeta}_t^{\lambda,n}$ defined in Definition 6.2 and Z_t^λ given in (25) as follows:

$$W_1(\mathcal{L}(\bar{\theta}_t^\lambda), \pi_\beta) \leq W_1(\mathcal{L}(\bar{\theta}_t^\lambda), \mathcal{L}(\bar{\zeta}_t^{\lambda,n})) + W_1(\mathcal{L}(\bar{\zeta}_t^{\lambda,n}), \mathcal{L}(Z_t^\lambda)) + W_1(\mathcal{L}(Z_t^\lambda), \pi_\beta). \tag{30}$$

In the following lemma, we provide a non-asymptotic estimate for $W_2(\mathcal{L}(\bar{\theta}_t^\lambda), \mathcal{L}(\bar{\zeta}_t^{\lambda,n}))$, which can be used to upper bound the first term on the RHS of (30).

Lemma 6.6. *Let Assumptions 1, 2, and 3 hold. Then, for any* $0 < \lambda \leq \lambda_{\max}$, $n \in \mathbb{N}_0$, *and* $t \in (nT, (n+1)T]$, *we obtain*

$$W_2(\mathcal{L}(\bar{\theta}_t^\lambda), \mathcal{L}(\bar{\zeta}_t^{\lambda,n})) \leq \lambda^{1+q/2}\left(e^{-\mathsf{a}_\mathsf{D}\min\{\kappa,1/2\}n/2}\mathsf{C}_0\mathbb{E}[|\theta_0|^{16(\rho+1)}] + \mathsf{C}_1\right)^{1/2},$$

where $\kappa, \mathsf{C}_0, \mathsf{C}_1$ *are given explicitly in (73) and (105).*

Proof. See Appendix A.3. □

For the last two terms on the RHS of (30), we observe that they can be viewed as Wasserstein-1 distances between distributions of Langevin processes starting from different initial points. Therefore, to obtain their upper bounds, we introduce a semi-metric which allows us to establish a contraction result for the Langevin SDE (24) under our assumptions.

We consider the following semi-metric: for any $p \in [2, \infty) \cap \mathbb{N}$, $\mu, \nu \in \mathcal{P}_{V_p}(\mathbb{R}^d)$, let

$$w_{1,p}(\mu, \nu) := \inf_{\zeta \in \mathcal{C}(\mu,\nu)} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} [1 \wedge |\theta - \theta'|](1 + V_p(\theta) + V_p(\theta'))\zeta(\mathrm{d}\theta, \mathrm{d}\theta'). \tag{31}$$

Then, we provide a result which states the contraction property of the Langevin SDE (24) in $w_{1,2}$.

Proposition 6.7. *Let Assumptions 1, 2, and 3 hold. Moreover, let $\theta'_0 \in L^2$, and let $(Z'_t)_{t\geq 0}$ be the solution of SDE (24) whose starting point $Z'_0 := \theta'_0$ is assumed to be independent of $\mathcal{F}_\infty := \sigma(\bigcup_{t\geq 0} \mathcal{F}_t)$. Then, we obtain*

$$w_{1,2}(\mathcal{L}(Z_t), \mathcal{L}(Z'_t)) \leq \hat{\mathsf{c}} e^{-\dot{\mathsf{c}}t} w_{1,2}(\mathcal{L}(\theta_0), \mathcal{L}(\theta'_0)), \tag{32}$$

where the explicit expressions for $\dot{\mathsf{c}}, \hat{\mathsf{c}}$ are given below.

The contraction constant $\dot{\mathsf{c}}$ is given by:

$$\dot{\mathsf{c}} := \min\{\bar{\phi}, \mathsf{c}_{V,1}(2), 4\mathsf{c}_{V,2}(2)\epsilon \mathsf{c}_{V,1}(2)\}/2,$$

where $\mathsf{c}_{V,1}(2) := \mathsf{a}_\mathsf{D}/2$, $\mathsf{c}_{V,2}(2) := 3\mathsf{a}_\mathsf{D}\mathrm{v}_2(\mathsf{M}_V(2))/2$ with $\mathsf{M}_V(2) := (1/3+4\bar{\mathsf{b}}_\mathsf{D}/(3\mathsf{a}_\mathsf{D})+4d/(3\mathsf{a}_\mathsf{D}\beta))^{1/2}$, the constant $\bar{\phi}$ is given by

$$\bar{\phi} := \left(\sqrt{8\pi/(\beta \mathsf{L}_{\mathsf{OS}})}\dot{\mathsf{c}}_0 \exp\left(\left(\dot{\mathsf{c}}_0\sqrt{\beta \mathsf{L}_{\mathsf{OS}}/8} + \sqrt{8/(\beta \mathsf{L}_{\mathsf{OS}})}\right)^2\right)\right)^{-1},$$

and $\epsilon > 0$ is chosen such that

$$\epsilon \leq 1 \wedge \left(4\mathsf{c}_{V,2}(2)\sqrt{2\beta\pi/\mathsf{L}_{\mathsf{OS}}} \int_0^{\dot{\mathsf{c}}_1} \exp\left(\left(s\sqrt{\beta \mathsf{L}_{\mathsf{OS}}/8} + \sqrt{8/(\beta \mathsf{L}_{\mathsf{OS}})}\right)^2\right) \mathrm{d}s\right)^{-1}$$

with $\dot{\mathsf{c}}_0 := 2(4\mathsf{c}_{V,2}(2)(1 + \mathsf{c}_{V,1}(2))/\mathsf{c}_{V,1}(2) - 1)^{1/2}$ and $\dot{\mathsf{c}}_1 := 2(2\mathsf{c}_{V,2}(2)/\mathsf{c}_{V,1}(2) - 1)^{1/2}$.

Moreover, the constant $\hat{\mathsf{c}}$ is given by:

$$\hat{\mathsf{c}} := 2(1 + \dot{\mathsf{c}}_0) \exp(\beta \mathsf{L}_{\mathsf{OS}}\dot{\mathsf{c}}_0^2/8 + 2\dot{\mathsf{c}}_0)/\epsilon.$$

Proof. We note that [11, Assumption 2.1] holds with $\kappa = \mathsf{L}_{\mathsf{OS}}$ due to Remark 2.5, [11, Assumption 2.2] holds with $V = V_2$ due to Remark 6.4, and [11, Assumptions 2.4 and 2.5] hold due to (29). Therefore, we can obtain (32) following the same arguments as in the proof of [4, Proposition 3.14] based on [11, Theorem 2.2, Corollary 2.3]. In addition, $\dot{\mathsf{c}}, \hat{\mathsf{c}}$ can be obtained following the arguments in the proof of [18, Proposition 4.6]. □

By using the above result and $W_1 \leq w_{1,2}$ (see [18, Lemma A.3]), we can establish a non-asymptotic error bound for the second term on the RHS of (30). The explicit statement is given below.

Lemma 6.8. *Let Assumptions 1, 2, and 3 hold. Then, for any $0 < \lambda \leq \lambda_{\max}$, $n \in \mathbb{N}_0$, and $t \in (nT, (n+1)T]$, we obtain*

$$W_1(\mathcal{L}(\bar{\zeta}_t^{\lambda,n}), \mathcal{L}(Z_t^\lambda)) \leq \lambda^{1+q/2}\left(e^{-\min\{\dot{\mathsf{c}},\mathsf{a}_\mathsf{D}\kappa,\mathsf{a}_\mathsf{D}/2\}n/4}\mathsf{C}_2\mathbb{E}[|\theta_0|^{16(\rho+1)}] + \mathsf{C}_3\right),$$

where

$$\begin{aligned}
\mathsf{C}_2 &:= \hat{\mathsf{c}}\left(1 + \frac{4}{\min\{\dot{\mathsf{c}}, \mathsf{a}_\mathsf{D}\kappa, \mathsf{a}_\mathsf{D}/2\}}\right) e^{\min\{\dot{\mathsf{c}},\mathsf{a}_\mathsf{D}\kappa,\mathsf{a}_\mathsf{D}/2\}/4}\left(\mathsf{C}_0 + 2^{8\rho+11}\right),\\
\mathsf{C}_3 &:= 2(\hat{\mathsf{c}}/\dot{\mathsf{c}})e^{\dot{\mathsf{c}}/2}\left(\mathsf{C}_1 + 15 + 2^{8\rho+11}(\mathsf{c}_{8(\rho+1)}(1 + 1/(\mathsf{a}_\mathsf{D}\kappa)) + 1)\right. \left. +18\mathrm{v}_{16(\rho+1)}(\mathsf{M}_V(16(\rho+1)))\right)
\end{aligned} \tag{33}$$

with $\dot{\mathsf{c}}, \hat{\mathsf{c}}$ given in Proposition 6.7, $\mathsf{C}_0, \mathsf{C}_1$ given in (105) (see also Lemma 6.6), κ, $\mathsf{c}_{8(\rho+1)}$ given in Lemma 6.3, and $\mathsf{M}_V(16(\rho+1))$ given in Lemma 6.5.

Proof. See [18, Lemma 4.7]. □

To obtain an upper bound for the last term on the RHS of (30), we observe that π_β is the invariant measure of the Langevin SDE (25). Thus, by applying Proposition 6.7, we have that

$$W_1(\mathcal{L}(Z_t^\lambda),\pi_\beta) \leq \hat{\mathsf{c}}e^{-\dot{\mathsf{c}}\lambda t}w_{1,2}(\mathcal{L}(\theta_0),\pi_\beta) \leq \hat{\mathsf{c}}e^{-\dot{\mathsf{c}}\lambda t}\left[1+\mathbb{E}[V_2(\theta_0)]+\int_{\mathbb{R}^d}V_2(\theta)\pi_\beta(\mathrm{d}\theta)\right]. \tag{34}$$

By using Lemma 6.6, 6.8 and (34), we can obtain an upper bound for each $W_2(\mathcal{L}(\theta_n^{\mathsf{aHOLA}}),\pi_\beta)$, $n\in\mathbb{N}_0$, as stated in Theorem 2.6.

Proof of Theorem 2.6. Substituting the results in Lemma 6.6, 6.8 and (34) into (30), for any $0<\lambda\leq\lambda_{\max}$, $n\in\mathbb{N}_0$, and $t\in(nT,(n+1)T]$, we have that

$$\begin{aligned}
W_1(\mathcal{L}(\overline{\theta}_t^\lambda),\pi_\beta) &\leq \lambda^{1+q/2}\left(e^{-\mathsf{a}_{\mathsf{D}}\min\{\kappa,1/2\}n/2}\mathsf{C}_0\mathbb{E}[|\theta_0|^{16(\rho+1)}]+\mathsf{C}_1\right)^{1/2}\\
&\quad+\lambda^{1+q/2}\left(e^{-\min\{\dot{\mathsf{c}},\mathsf{a}_{\mathsf{D}}\kappa,\mathsf{a}_{\mathsf{D}}/2\}n/4}\mathsf{C}_2\mathbb{E}[|\theta_0|^{16(\rho+1)}]+\mathsf{C}_3\right)\\
&\quad+\hat{\mathsf{c}}e^{-\dot{\mathsf{c}}\lambda t}\left[1+\mathbb{E}[V_2(\theta_0)]+\int_{\mathbb{R}^d}V_2(\theta)\pi_\beta(\mathrm{d}\theta)\right]\\
&\leq C_1e^{-C_0(n+1)}(\mathbb{E}[|\theta_0|^{16(\rho+1)}]+1)+C_2\lambda^{1+q/2},
\end{aligned}$$

where

$$\begin{aligned}
C_0 &:= \min\{\dot{\mathsf{c}},\mathsf{a}_{\mathsf{D}}\kappa,\mathsf{a}_{\mathsf{D}}/2\}/4,\\
C_1 &:= e^{\min\{\dot{\mathsf{c}},\mathsf{a}_{\mathsf{D}}\kappa,\mathsf{a}_{\mathsf{D}}/2\}/4}\left[\mathsf{C}_0^{1/2}+\mathsf{C}_2+\hat{\mathsf{c}}\left(3+\int_{\mathbb{R}^d}V_2(\theta)\pi_\beta(\mathrm{d}\theta)\right)\right],\\
C_2 &:= \mathsf{C}_1^{1/2}+\mathsf{C}_3
\end{aligned} \tag{35}$$

with $\dot{\mathsf{c}},\hat{\mathsf{c}}$ given in Proposition 6.7, κ given in Lemma 6.3, $\mathsf{C}_0,\mathsf{C}_1$ given in (105) (see also Lemma 6.6), $\mathsf{C}_2,\mathsf{C}_3$ given in (33) (see also Lemma 6.8). The above result implies that, for each $n\in\mathbb{N}_0$,

$$W_1(\mathcal{L}(\overline{\theta}_{nT}^\lambda),\pi_\beta)\leq C_1e^{-C_0n}(\mathbb{E}[|\theta_0|^{16(\rho+1)}]+1)+C_2\lambda^{1+q/2},$$

which further yields, by setting nT to n on the LHS and n to n/T on the RHS, that

$$W_1(\mathcal{L}(\overline{\theta}_n^\lambda),\pi_\beta)=W_1(\mathcal{L}(\theta_n^\lambda),\pi_\beta)=W_1(\mathcal{L}(\theta_n^{\mathsf{aHOLA}}),\pi_\beta)\leq C_1e^{-C_0\lambda n}(\mathbb{E}[|\theta_0|^{16(\rho+1)}]+1)+C_2\lambda^{1+q/2},$$

where the inequality holds due to $n\lambda\leq n/T$. This completes the proof. □

By using similar arguments as in the proof of Theorem 2.6, we can obtain the upper bound for $W_2(\mathcal{L}(\theta_n^{\mathsf{aHOLA}}),\pi_\beta)$, $n\in\mathbb{N}_0$, as stated in Corollary 2.7.

Proof of Corollary 2.7. To establish a non-asymptotic error bound for $W_2(\mathcal{L}(\theta_n^{\mathsf{aHOLA}}),\pi_\beta)$, we consider the following splitting: for any $0<\lambda\leq\lambda_{\max}$, $n\in\mathbb{N}_0$, and $t\in(nT,(n+1)T]$,

$$W_2(\mathcal{L}(\overline{\theta}_t^\lambda),\pi_\beta)\leq W_2(\mathcal{L}(\overline{\theta}_t^\lambda),\mathcal{L}(\overline{\zeta}_t^{\lambda,n}))+W_2(\mathcal{L}(\overline{\zeta}_t^{\lambda,n}),\mathcal{L}(Z_t^\lambda))+W_2(\mathcal{L}(Z_t^\lambda),\pi_\beta). \tag{36}$$

An upper bound for the first term on the RHS of (36) is provided in Lemma 6.6. To establish an estimate for the second term on the RHS of (36), we use $W_2\leq\sqrt{2w_{1,2}}$ (see [18, Lemma A.3] for the proof) and follow the same arguments as that in the proof of [18, Lemma 4.7]. Consequently, for any $0<\lambda\leq\lambda_{\max}$, $n\in\mathbb{N}_0$, and $t\in(nT,(n+1)T]$, we obtain that,

$$W_2(\mathcal{L}(\overline{\zeta}_t^{\lambda,n}),\mathcal{L}(Z_t^\lambda))\leq\lambda^{1/2+q/4}\left(e^{-\min\{\dot{\mathsf{c}},\mathsf{a}_{\mathsf{D}}\kappa,\mathsf{a}_{\mathsf{D}}/2\}n/8}\mathsf{C}_4\mathbb{E}^{1/2}[|\theta_0|^{16(\rho+1)}]+\mathsf{C}_5\right), \tag{37}$$

where

$$\begin{aligned}
\mathsf{C}_4 &:= \sqrt{\hat{\mathsf{c}}}\left(1+\frac{8}{\min\{\dot{\mathsf{c}},\mathsf{a}_{\mathsf{D}}\kappa,\mathsf{a}_{\mathsf{D}}/2\}}\right)e^{\min\{\dot{\mathsf{c}},\mathsf{a}_{\mathsf{D}}\kappa,\mathsf{a}_{\mathsf{D}}/2\}/8}\left(\sqrt{2}\mathsf{C}_0^{1/2}+2^{4\rho+5}\right),\\
\mathsf{C}_5 &:= 4(\sqrt{\hat{\mathsf{c}}}/\dot{\mathsf{c}})e^{\dot{\mathsf{c}}/4}\left(\sqrt{2}\mathsf{C}_1^{1/2}+3\sqrt{2}+2^{4\rho+5}(\mathsf{c}_{8(\rho+1)}(1+1/(\mathsf{a}_{\mathsf{D}}\kappa))+1)^{1/2}\right.\\
&\qquad\left.+\sqrt{6}\mathsf{v}_{16(\rho+1)}^{1/2}(\mathsf{M}_V(16(\rho+1)))\right)
\end{aligned} \tag{38}$$

with $\dot{\mathsf{c}}, \hat{\mathsf{c}}$ given in Proposition 6.7, $\mathsf{C}_0, \mathsf{C}_1$ given in (105) (see also Lemma 6.6), κ, $\mathsf{c}_{8(\rho+1)}$ given in Lemma 6.3, and $\mathsf{M}_V(16(\rho+1))$ given in Lemma 6.5. An upper bound for the last term on the RHS of (36) can be obtained by using $W_2 \leq \sqrt{2w_{1,2}}$ and Proposition 6.7:

$$
\begin{aligned}
W_2(\mathcal{L}(Z_t^\lambda), \pi_\beta) &\leq \sqrt{2\hat{\mathsf{c}}} e^{-\dot{\mathsf{c}}\lambda t/2} w_{1,2}^{1/2}(\mathcal{L}(\theta_0), \pi_\beta) \\
&\leq \sqrt{2\hat{\mathsf{c}}} e^{-\dot{\mathsf{c}}\lambda t/2} \left[1 + \mathbb{E}[V_2(\theta_0)] + \int_{\mathbb{R}^d} V_2(\theta) \pi_\beta(\mathrm{d}\theta)\right]^{1/2}. \qquad (39)
\end{aligned}
$$

Applying the results in Lemma 6.6, (37), (39) to (36) yields, for any $0 < \lambda \leq \lambda_{\max}$, $n \in \mathbb{N}_0$, and $t \in (nT, (n+1)T]$, that

$$
\begin{aligned}
W_2(\mathcal{L}(\overline{\theta}_t^\lambda), \pi_\beta) &\leq \lambda^{1+q/2} \left(e^{-\mathsf{a}_\mathsf{D} \min\{\kappa, 1/2\} n/2} \mathsf{C}_0 \mathbb{E}[|\theta_0|^{16(\rho+1)}] + \mathsf{C}_1\right)^{1/2} \\
&\quad + \lambda^{1/2+q/4} \left(e^{-\min\{\dot{\mathsf{c}}, \mathsf{a}_\mathsf{D}\kappa, \mathsf{a}_\mathsf{D}/2\} n/8} \mathsf{C}_4 \mathbb{E}^{1/2}[|\theta_0|^{16(\rho+1)}] + \mathsf{C}_5\right) \\
&\quad + \sqrt{2\hat{\mathsf{c}}} e^{-\dot{\mathsf{c}}\lambda t/2} \left[1 + \mathbb{E}[V_2(\theta_0)] + \int_{\mathbb{R}^d} V_2(\theta) \pi_\beta(\mathrm{d}\theta)\right]^{1/2} \\
&\leq C_4 e^{-C_3(n+1)} (\mathbb{E}[|\theta_0|^{16(\rho+1)}] + 1)^{1/2} + C_5 \lambda^{1/2+q/4},
\end{aligned}
$$

where

$$
\begin{aligned}
C_3 &:= \min\{\dot{\mathsf{c}}, \mathsf{a}_\mathsf{D}\kappa, \mathsf{a}_\mathsf{D}/2\}/8, \\
C_4 &:= e^{\min\{\dot{\mathsf{c}}, \mathsf{a}_\mathsf{D}\kappa, \mathsf{a}_\mathsf{D}/2\}/8} \left[\mathsf{C}_0^{1/2} + \mathsf{C}_4 + \sqrt{2\hat{\mathsf{c}}} \left(3 + \int_{\mathbb{R}^d} V_2(\theta) \pi_\beta(\mathrm{d}\theta)\right)^{1/2}\right], \qquad (40) \\
C_5 &:= \mathsf{C}_1^{1/2} + \mathsf{C}_5
\end{aligned}
$$

with $\dot{\mathsf{c}}, \hat{\mathsf{c}}$ given in Proposition 6.7, κ given in Lemma 6.3, $\mathsf{C}_0, \mathsf{C}_1$ given in (105) (see also Lemma 6.6), $\mathsf{C}_4, \mathsf{C}_5$ given in (38). This further implies that, for each $n \in \mathbb{N}_0$,

$$
\begin{aligned}
W_2(\mathcal{L}(\overline{\theta}_n^\lambda), \pi_\beta) = W_2(\mathcal{L}(\theta_n^\lambda), \pi_\beta) &= W_2(\mathcal{L}(\theta_n^{\mathsf{aHOLA}}), \pi_\beta) \\
&\leq C_4 e^{-C_3 \lambda n} (\mathbb{E}[|\theta_0|^{16(\rho+1)}] + 1)^{1/2} + C_5 \lambda^{1/2+q/4},
\end{aligned}
$$

which completes the proof. □

APPENDIX A. PROOF OF AUXILIARY RESULTS

A.1. Proof of auxiliary results in Section 2.

Proof of statements in Remark 2.2. We provide detailed proofs for inequalities (10)-(12), the other inequalities can be obtained by using similar arguments. By Assumption 2, for any $\theta, \overline{\theta} \in \mathbb{R}^d$, $i = 1, \dots, d$, we have that

$$
\begin{aligned}
|\nabla^2 h^{(i)}(\theta)| &\leq |\nabla^2 h^{(i)}(\theta) - \nabla^2 h^{(i)}(0)| + |\nabla^2 h^{(i)}(0)| \\
&\leq L(1+|\theta|)^{\rho-2}|\theta|^q + |\nabla^2 h^{(i)}(0)| \\
&\leq \mathsf{K}_0(1+|\theta|)^{\rho+q-2},
\end{aligned}
$$

where $\mathsf{K}_0 := 2^{1-q}\max\{L, |\nabla^2 h^{(1)}(0)|, \dots, |\nabla^2 h^{(d)}(0)|\}$. Moreover, fix $\theta, \overline{\theta} \in \mathbb{R}^d$, denote by $g(t) := \nabla h^{(i)}(t\theta + (1-t)\overline{\theta})$, $t \in [0,1]$. Then, by using the above inequality, we obtain that

$$
\begin{aligned}
|\nabla h^{(i)}(\theta) - \nabla h^{(i)}(\overline{\theta})| &= \left| \int_0^1 \nabla^2 h^{(i)}(t\theta + (1-t)\overline{\theta})(\theta - \overline{\theta}) \mathrm{d}t \right| \\
&\leq \int_0^1 \mathsf{K}_0(1 + |t\theta + (1-t)\overline{\theta}|)^{\rho+q-2} \mathrm{d}t |\theta - \overline{\theta}| \\
&\leq \mathsf{K}_0(1 + |\theta| + |\overline{\theta}|)^{\rho+q-2}|\theta - \overline{\theta}|.
\end{aligned}
$$

In addition, for any $\theta, \overline{\theta} \in \mathbb{R}^d$, by Assumption 2, we have that

$$
\begin{aligned}
|\Upsilon(\theta) - \Upsilon(\overline{\theta})| &= \left(\sum_{i=1}^d \left(\sum_{j=1}^d (\partial^2_{\theta^{(j)}} h^{(i)}(\theta) - \partial^2_{\theta^{(j)}} h^{(i)}(\overline{\theta})) \right)^2 \right)^{1/2} \\
&\leq \left(d \sum_{i=1}^d \sum_{j=1}^d \left(\partial^2_{\theta^{(j)}} h^{(i)}(\theta) - \partial^2_{\theta^{(j)}} h^{(i)}(\overline{\theta}) \right)^2 \right)^{1/2} \\
&\leq \left(d \sum_{i=1}^d |\nabla^2 h^{(i)}(\theta) - \nabla^2 h^{(i)}(\overline{\theta})|_{\mathsf{F}}^2 \right)^{1/2} \\
&\leq \left(d^2 \sum_{i=1}^d |\nabla^2 h^{(i)}(\theta) - \nabla^2 h^{(i)}(\overline{\theta})|^2 \right)^{1/2} \\
&\leq d^{3/2} L(1 + |\theta| + |\overline{\theta}|)^{\rho-2}|\theta - \overline{\theta}|^q,
\end{aligned}
$$

which completes the proof. □

Proof of statements in Remark 2.4. By Assumption 3, for any $\theta \in \mathbb{R}^d$, we have that

$$
\langle \theta, h(\theta) - h(0) \rangle \geq a|\theta|^2|\theta|^r - b|\theta|^2|\theta|^{\overline{r}},
$$

which implies that

$$
\begin{aligned}
\langle \theta, h(\theta) \rangle &\geq a|\theta|^{r+2} - b|\theta|^{\overline{r}+2} + \langle \theta, h(0) \rangle \\
&\geq a|\theta|^{r+2} - b|\theta|^{\overline{r}+2} - |\langle \theta, h(0) \rangle| \\
&\geq (a/2)|\theta|^{r+2} + ((a/4)|\theta|^{r+2} - b|\theta|^{\overline{r}+2}) + ((a/4)|\theta|^{r+2} - a|\theta|^2/2) - |h(0)|^2/(2a), \quad (41)
\end{aligned}
$$

where the last inequality holds due to the fact that $yz \leq \varepsilon y^2/2 + z^2/(2\varepsilon)$ for any $y, z \geq 0$ and $\varepsilon > 0$. Denote by $\mathsf{R}_\mathsf{D} := \max\{(4b/a)^{1/(r-\overline{r})}, 2^{1/r}\} > 1$. We observe that for any $|\theta| > \mathsf{R}_\mathsf{D}$,

$$
(a/4)|\theta|^{r+2} - b|\theta|^{\overline{r}+2} > 0, \quad (a/4)|\theta|^{r+2} - a|\theta|^2/2 > 0,
$$

and thus (41) becomes

$$
\langle \theta, h(\theta) \rangle > (a/2)|\theta|^{r+2} - |h(0)|^2/(2a). \quad (42)
$$

Moreover, for any $|\theta| \leq \mathsf{R}_\mathsf{D}$, (41) becomes

$$
\langle \theta, h(\theta) \rangle \geq (a/2)|\theta|^{r+2} - (b + a/2)\mathsf{R}_\mathsf{D}^{\overline{r}+2} - |h(0)|^2/(2a). \quad (43)
$$

Combining (42) and (43) yields

$$\langle\theta, h(\theta)\rangle \geq \mathsf{a}_{\mathsf{D}}|\theta|^{r+2} - \mathsf{b}_{\mathsf{D}}, \tag{44}$$

where $\mathsf{a}_{\mathsf{D}} := a/2$ and $\mathsf{b}_{\mathsf{D}} := (a/2+b)\mathsf{R}_{\mathsf{D}}^{\overline{r}+2} + |h(0)|^2/(2a)$ with $\mathsf{R}_{\mathsf{D}} := \max\{(4b/a)^{1/(r-\overline{r})}, 2^{1/r}\}$. Moreover, by (44), for any $\theta \in \mathbb{R}^d$, we observe that,

$$\langle\theta, h(\theta)\rangle \geq \mathsf{a}_{\mathsf{D}}|\theta|^{2} - \overline{\mathsf{b}}_{\mathsf{D}},$$

where $\overline{\mathsf{b}}_{\mathsf{D}} := \mathsf{a}_{\mathsf{D}} + \mathsf{b}_{\mathsf{D}}$. Indeed, for $|\theta| > 1$, it holds that

$$\langle\theta, h(\theta)\rangle \geq \mathsf{a}_{\mathsf{D}}|\theta|^{r+2} - \mathsf{b}_{\mathsf{D}} > \mathsf{a}_{\mathsf{D}}|\theta|^{2} - \mathsf{b}_{\mathsf{D}} > \mathsf{a}_{\mathsf{D}}|\theta|^{2} - \overline{\mathsf{b}}_{\mathsf{D}},$$

while for $|\theta| \leq 1$, we have that

$$\langle\theta, h(\theta)\rangle \geq \mathsf{a}_{\mathsf{D}}|\theta|^{r+2} - \mathsf{b}_{\mathsf{D}} \geq \mathsf{a}_{\mathsf{D}} - (\mathsf{a}_{\mathsf{D}} + \mathsf{b}_{\mathsf{D}}) \geq \mathsf{a}_{\mathsf{D}}|\theta|^{2} - \overline{\mathsf{b}}_{\mathsf{D}},$$

which completes the proof. □

Proof of statements in Remark 2.5. For any $\theta, \overline{\theta} \in \mathbb{R}^d$, we observe that the result holds trivially when $\theta = \overline{\theta}$, and thus we consider only the case where $\theta \neq \overline{\theta}$. Denote by $\mathsf{R}_{\mathsf{OS}} := (b/a)^{1/(r-\overline{r})}$. By Assumption 3, for any $|\theta|, |\overline{\theta}| > \mathsf{R}_{\mathsf{OS}}$, we have that

$$\langle\theta - \overline{\theta}, h(\theta) - h(\overline{\theta})\rangle > 0. \tag{45}$$

Furthermore, we note that by Remark 2.2,

$$|h(\theta) - h(\overline{\theta})| \leq \sqrt{d}\mathsf{K}_1(1 + |\theta| + |\overline{\theta}|)^{\rho+q-1}|\theta - \overline{\theta}|.$$

Thus, for any $|\theta|, |\overline{\theta}| \leq \mathsf{R}_{\mathsf{OS}}$, we have that

$$-\langle\theta - \overline{\theta}, h(\theta) - h(\overline{\theta})\rangle \leq |\theta - \overline{\theta}||h(\theta) - h(\overline{\theta})| \leq \mathsf{L}_{\mathsf{OS}}|\theta - \overline{\theta}|^2, \tag{46}$$

where $\mathsf{L}_{\mathsf{OS}} := \sqrt{d}\mathsf{K}_1(1 + 2\mathsf{R}_{\mathsf{OS}})^{\rho+q-1} > 0$. In addition, for $|\theta| \leq \mathsf{R}_{\mathsf{OS}}$, $|\overline{\theta}| > \mathsf{R}_{\mathsf{OS}}$, we consider the following two cases:

(i) For $|\theta| = \mathsf{R}_{\mathsf{OS}}$, $|\overline{\theta}| > \mathsf{R}_{\mathsf{OS}}$, we obtain (45).

(ii) For any $|\theta| < \mathsf{R}_{\mathsf{OS}}$, $|\overline{\theta}| > \mathsf{R}_{\mathsf{OS}}$, there is a unique $\hat{\theta} \in \mathbb{R}^d$ with $|\hat{\theta}| = |\mathsf{R}_{\mathsf{OS}}|$ such that

$$\theta - \hat{\theta} = \mathsf{c}_{\theta,\overline{\theta}}(\theta - \overline{\theta}), \quad \hat{\theta} - \overline{\theta} = (1 - \mathsf{c}_{\theta,\overline{\theta}})(\theta - \overline{\theta}),$$

where $\mathsf{c}_{\theta,\overline{\theta}} \in (0, 1)$. Then, we obtain that

$$\begin{aligned}
\langle\theta - \overline{\theta}, h(\theta) - h(\overline{\theta})\rangle &= \langle\theta - \overline{\theta}, h(\theta) - h(\hat{\theta})\rangle + \langle\theta - \hat{\theta}, h(\hat{\theta}) - h(\overline{\theta})\rangle + \langle\hat{\theta} - \overline{\theta}, h(\hat{\theta}) - h(\overline{\theta})\rangle \\
&= \langle\theta - \overline{\theta}, h(\theta) - h(\hat{\theta})\rangle + (\mathsf{c}_{\theta,\overline{\theta}}/(1 - \mathsf{c}_{\theta,\overline{\theta}}))\langle\hat{\theta} - \overline{\theta}, h(\hat{\theta}) - h(\overline{\theta})\rangle \\
&> (1/\mathsf{c}_{\theta,\overline{\theta}})\langle\theta - \hat{\theta}, h(\theta) - h(\hat{\theta})\rangle \qquad (47) \\
&\geq -(\mathsf{L}_{\mathsf{OS}}/\mathsf{c}_{\theta,\overline{\theta}})|\theta - \hat{\theta}|^2 \qquad (48) \\
&= -\mathsf{L}_{\mathsf{OS}}\mathsf{c}_{\theta,\overline{\theta}}|\theta - \hat{\theta}|^2 \\
&\geq -\mathsf{L}_{\mathsf{OS}}|\theta - \hat{\theta}|^2, \qquad (49)
\end{aligned}$$

where (47) holds due to (45), (48) holds due to (46), and (49) holds due to $\mathsf{c}_{\theta,\overline{\theta}} \in (0, 1)$.

Thus, for $|\theta| \leq \mathsf{R}_{\mathsf{OS}}$, $|\overline{\theta}| > \mathsf{R}_{\mathsf{OS}}$, combining the two cases yield

$$\langle\theta - \overline{\theta}, h(\theta) - h(\overline{\theta})\rangle \geq -\mathsf{L}_{\mathsf{OS}}|\theta - \overline{\theta}|^2. \tag{50}$$

Finally, for $|\theta| > \mathsf{R}_{\mathsf{OS}}$, $|\overline{\theta}| \leq \mathsf{R}_{\mathsf{OS}}$, we obtain (50) by applying the same arguments above. Combining (45), (46), and (50) yields the desired result. □

A.2. **Proof of auxiliary results in Section 5.**

Proof of Proposition 5.1. We first show that the potentials given in (20) and (21) satisfy Assumptions 5 and 6.

(i) For U given in (20), we have that, for any $\theta \in \mathbb{R}^d$, $i = 1, \dots, d$,

$$h(\theta) = \theta, \quad H(\theta) = I_d, \quad \nabla^2 h^{(i)}(\theta) = 0.$$

Then, we observe that Assumption 5 holds with $\overline{L}_1 = 0$, $q = 1$, $\overline{L}_2 = 0$, and $\overline{L}_3 = 1$. Moreover, Assumption 6 holds with $\overline{a} = 1$, $\overline{b} = 0$ since

$$\langle \theta, h(\theta) \rangle = |\theta|^2.$$

(ii) For U given in (21), we have that, for any $\theta \in \mathbb{R}^d$, $i = 1, \dots, d$,

$$\begin{aligned} h(\theta) &= \theta - \hat{a} + \frac{2\hat{a}}{1 + \exp(2\langle \theta, \hat{a} \rangle)}, \\ H(\theta) &= I_d - \frac{4\hat{a}\hat{a}^\mathsf{T} \exp(2\langle \theta, \hat{a} \rangle)}{(1 + \exp(2\langle \theta, \hat{a} \rangle))^2}, \\ \nabla^2 h^{(i)}(\theta) &= \frac{8\hat{a}^{(i)}\hat{a}\hat{a}^\mathsf{T}(\exp(4\langle \theta, \hat{a} \rangle) - \exp(2\langle \theta, \hat{a} \rangle))}{(1 + \exp(2\langle \theta, \hat{a} \rangle))^3}. \end{aligned}$$

To show that Assumption 5 holds, we consider the following calculations. For any $\theta, \overline{\theta} \in \mathbb{R}^d$, we have that

$$\begin{aligned} &|h(\theta) - h(\overline{\theta})| \\ &= \left| \theta - \hat{a} + \frac{2\hat{a}}{1 + \exp(2\langle \theta, \hat{a} \rangle)} - \left(\overline{\theta} - \hat{a} + \frac{2\hat{a}}{1 + \exp(2\langle \overline{\theta}, \hat{a} \rangle)}, \right) \right| \\ &\le |\theta - \overline{\theta}| + 2|\hat{a}| \frac{|\exp(2\langle \overline{\theta}, \hat{a} \rangle)) - \exp(2\langle \theta, \hat{a} \rangle))|}{(1 + \exp(2\langle \theta, \hat{a} \rangle))(1 + \exp(2\langle \overline{\theta}, \hat{a} \rangle))} \\ &= |\theta - \overline{\theta}| + 2|\hat{a}| \frac{\max\{\exp(2\langle \theta, \hat{a} \rangle), \exp(2\langle \overline{\theta}, \hat{a} \rangle)\} |1 - \exp(-2|\langle \theta, \hat{a} \rangle - \langle \overline{\theta}, \hat{a} \rangle|)|}{(1 + \exp(2\langle \theta, \hat{a} \rangle))(1 + \exp(2\langle \overline{\theta}, \hat{a} \rangle))} \\ &\le |\theta - \overline{\theta}| + 4|\hat{a}||\langle \theta, \hat{a} \rangle - \langle \overline{\theta}, \hat{a} \rangle| \\ &\le (1 + 4|\hat{a}|^2)|\theta - \overline{\theta}|, \end{aligned} \tag{51}$$

where the second last inequality holds due to $1 - e^{-x} \le x$, for all $x \in \mathbb{R}$, and the last inequality holds due to Cauchy-Schwarz inequality. Similarly, we have that, for any $\theta, \overline{\theta} \in \mathbb{R}^d$,

$$\begin{aligned} &|H(\theta) - H(\overline{\theta})| \\ &= 4|\hat{a}|^2 \frac{|\exp(2\langle \overline{\theta}, \hat{a} \rangle)(1 + \exp(2\langle \theta, \hat{a} \rangle))^2 - \exp(2\langle \theta, \hat{a} \rangle)(1 + \exp(2\langle \overline{\theta}, \hat{a} \rangle))^2|}{(1 + \exp(2\langle \theta, \hat{a} \rangle))^2(1 + \exp(2\langle \overline{\theta}, \hat{a} \rangle))^2} \\ &\le 4|\hat{a}|^2 \frac{(1 + \exp(2\langle \overline{\theta} + \theta, \hat{a} \rangle))|\exp(2\langle \overline{\theta}, \hat{a} \rangle) - \exp(2\langle \theta, \hat{a} \rangle)|}{(1 + \exp(2\langle \theta, \hat{a} \rangle))^2(1 + \exp(2\langle \overline{\theta}, \hat{a} \rangle))^2} \\ &\le 8|\hat{a}|^3|\theta - \overline{\theta}|, \end{aligned}$$

where the last inequality holds due to the calculations in (51). Furthermore, we note that, for any $\theta, \overline{\theta} \in \mathbb{R}^d$,

$$\begin{aligned} &|\nabla^2 h^{(i)}(\theta) - \nabla^2 h^{(i)}(\overline{\theta})| \\ &\le 8|\hat{a}|^3 \left| \frac{\exp(4\langle \theta, \hat{a} \rangle) - \exp(2\langle \theta, \hat{a} \rangle)}{(1 + \exp(2\langle \theta, \hat{a} \rangle))^3} - \frac{\exp(4\langle \overline{\theta}, \hat{a} \rangle) - \exp(2\langle \overline{\theta}, \hat{a} \rangle)}{(1 + \exp(2\langle \overline{\theta}, \hat{a} \rangle))^3} \right| \\ &\le 8|\hat{a}|^3 \frac{(1 + \exp(2\langle \overline{\theta} + \theta, \hat{a} \rangle))|\exp(4\langle \overline{\theta}, \hat{a} \rangle) - \exp(4\langle \theta, \hat{a} \rangle)|}{(1 + \exp(2\langle \theta, \hat{a} \rangle))^3(1 + \exp(2\langle \overline{\theta}, \hat{a} \rangle))^3} \\ &\quad + 8|\hat{a}|^3 \frac{(1 + 6\exp(2\langle \overline{\theta} + \theta, \hat{a} \rangle) + \exp(4\langle \overline{\theta} + \theta, \hat{a} \rangle))|\exp(2\langle \overline{\theta}, \hat{a} \rangle) - \exp(2\langle \theta, \hat{a} \rangle)|}{(1 + \exp(2\langle \theta, \hat{a} \rangle))^3(1 + \exp(2\langle \overline{\theta}, \hat{a} \rangle))^3} \\ &\le 56|\hat{a}|^4|\theta - \overline{\theta}|, \end{aligned}$$

where the last inequality holds due to the calculations in (51). Consequently, Assumption 5 holds with $\overline{L}_1 = 56|\hat{a}|^4$, $q = 1$, $\overline{L}_2 = 8|\hat{a}|^3$, and $\overline{L}_3 = 1 + 4|\hat{a}|^2$. Now, we show Assumption 6 holds with $\overline{a} = 1/2$, $\overline{b} = 1/2$. Indeed, we have, for any $\theta \in \mathbb{R}^d$, that

$$\langle \theta, h(\theta) \rangle = |\theta|^2 + \frac{1 - \exp(2\langle \theta, \hat{a}\rangle)}{1 + \exp(2\langle \theta, \hat{a}\rangle)} \langle \theta, \hat{a}\rangle \geq |\theta|^2 - \frac{1}{2}(|\theta|^2 + |\hat{a}|^2) = \frac{1}{2}(|\theta|^2 - |\hat{a}|^2).$$

Finally, we show that the potential (22) satisfies Assumptions 2 and 3.

(iii) For U given in (22), by Remark 2.3, we have that Assumption 2 is satisfied with $\rho = 2$, $L = 6$, $q = 1$, $K_h = 2$, and $K_H = 3$. In addition, Assumption 3 is satisfied with $a = 1/2$, $b = 1$, $r = 2$, and $\overline{r} = 0$. Indeed, for any $\theta \in \mathbb{R}^d$, by noticing

$$h(\theta) = \theta(|\theta|^2 - 1),$$

we obtain that

$$\begin{aligned}
\langle \theta - \overline{\theta}, h(\theta) - h(\overline{\theta}) \rangle &= \langle \theta - \overline{\theta}, \theta|\theta|^2 - \theta - (\overline{\theta}|\overline{\theta}|^2 - \overline{\theta}) \rangle \\
&= \frac{1}{2}\langle \theta - \overline{\theta}, \theta|\theta|^2 - \theta|\overline{\theta}|^2 + \theta|\overline{\theta}|^2 - \overline{\theta}|\overline{\theta}|^2 \rangle \\
&\quad + \frac{1}{2}\langle \theta - \overline{\theta}, \theta|\theta|^2 - \overline{\theta}|\theta|^2 + \overline{\theta}|\theta|^2 - \overline{\theta}|\overline{\theta}|^2 \rangle - |\theta - \overline{\theta}|^2 \\
&= \frac{1}{2}(|\theta|^2 + |\overline{\theta}|^2)|\theta - \overline{\theta}|^2 + \frac{1}{2}(|\theta|^2 - |\overline{\theta}|^2)^2 - |\theta - \overline{\theta}|^2 \\
&\geq \frac{1}{2}(|\theta|^2 + |\overline{\theta}|^2)|\theta - \overline{\theta}|^2 - |\theta - \overline{\theta}|^2.
\end{aligned}$$

This completes the proof. □

A.3. Proof of auxiliary results in Section 6.

Lemma A.1. *Let Assumptions 1, 2, and 3 hold. Then, for any $p \in \mathbb{N}, t \geq 0$, we obtain that*

$$\mathbb{E}[|Z_t^\lambda|^{2p}] \leq e^{-\lambda p \mathsf{a}_\mathsf{D} t}\mathbb{E}[|\theta_0|^{2p}] + 2(\overline{\mathsf{b}}_\mathsf{D} + \beta^{-1}(d + 2(p-1)))\mathsf{M}_0^{2p-2}/\mathsf{a}_\mathsf{D} < \infty,$$

where $\mathsf{M}_0 := (2(\overline{\mathsf{b}}_\mathsf{D} + \beta^{-1}(d + 2(p-1)))/\mathsf{a}_\mathsf{D})^{1/2}$.

Proof. See [18, Lemma A.1]. □

Proof of Lemma 6.3-(i). For any $0 < \lambda \leq \lambda_{\max} \leq 1$ with $\lambda_{\max}$ given in (14), $t \in (n, n+1]$, $n \in \mathbb{N}_0$, we define

$$\Delta_{n,t}^\lambda := \widetilde{\theta}_n^\lambda + \lambda\phi^\lambda(\widetilde{\theta}_n^\lambda)(t - n), \quad \Xi_{n,t}^\lambda := \sqrt{2\lambda\beta^{-1}}\psi^\lambda(\widetilde{\theta}_n^\lambda)(B_t^\lambda - B_n^\lambda), \tag{52}$$

where, for all $\theta \in \mathbb{R}^d$,

$$\phi^\lambda(\theta) := -h_\lambda(\theta) + (\lambda/2)\left(H_\lambda(\theta)h_\lambda(\theta) - \beta^{-1}\Upsilon_\lambda(\theta)\right), \tag{53}$$

and

$$\psi^\lambda(\theta) := \sqrt{\mathsf{I}_d - \lambda H_\lambda(\theta) + (\lambda^2/3)(H_\lambda(\theta))^2}. \tag{54}$$

By using (26) together with (52) – (54), and by noticing $\mathbb{E}\left[\langle \Delta_{n,t}^\lambda, \Xi_{n,t}^\lambda\rangle \middle| \widetilde{\theta}_n^\lambda\right] = 0$, we obtain that

$$\mathbb{E}\left[|\widetilde{\theta}_t^\lambda|^2 \middle| \widetilde{\theta}_n^\lambda\right] = |\Delta_{n,t}^\lambda|^2 + \mathbb{E}\left[|\Xi_{n,t}^\lambda|^2 \middle| \widetilde{\theta}_n^\lambda\right]. \tag{55}$$

Further calculations yield the following upper bound for the second term on the RHS of (55):

$$\begin{aligned}
\mathbb{E}\left[|\Xi_{n,t}^\lambda|^2 \middle| \widetilde{\theta}_n^\lambda\right] &= 2\lambda\beta^{-1}\mathbb{E}\left[\left\langle \psi^\lambda(\widetilde{\theta}_n^\lambda)(B_t^\lambda - B_n^\lambda), \psi^\lambda(\widetilde{\theta}_n^\lambda)(B_t^\lambda - B_n^\lambda)\right\rangle \middle| \widetilde{\theta}_n^\lambda\right] \\
&= 2\lambda\beta^{-1}\mathbb{E}\left[\left\langle B_t^\lambda - B_n^\lambda, \left(\mathsf{I}_d - \lambda H_\lambda(\widetilde{\theta}_n^\lambda) + (\lambda^2/3)(H_\lambda(\widetilde{\theta}_n^\lambda))^2\right)(B_t^\lambda - B_n^\lambda)\right\rangle \middle| \widetilde{\theta}_n^\lambda\right] \\
&\leq 2\lambda\beta^{-1}\left(d(t-n) + \lambda\sum_{i=1}^d |H_\lambda^{(i,i)}(\widetilde{\theta}_n^\lambda)|(t-n) + (\lambda^2/3)|H_\lambda(\widetilde{\theta}_n^\lambda)|_\mathsf{F}^2(t-n)\right) \\
&\leq 2\lambda\beta^{-1}(t-n)\left(d + \lambda\sqrt{d}|H_\lambda(\widetilde{\theta}_n^\lambda)|_\mathsf{F} + (\lambda^2/3)|H_\lambda(\widetilde{\theta}_n^\lambda)|_\mathsf{F}^2\right).
\end{aligned} \tag{56}$$

By Remark 2.2, we have that, for all $\theta \in \mathbb{R}^d$,

$$|H(\theta)|_{\mathsf{F}} = \left(\sum_{i=1}^{d} |\nabla h^{(i)}(\theta)|^2 \right)^{1/2} \leq \left(\sum_{i=1}^{d} \mathsf{K}_1^2 (1+|\theta|)^{2(\rho+q-1)} \right)^{1/2} = \sqrt{d}\mathsf{K}_1 (1+|\theta|)^{\rho+q-1},$$

which implies that, for any $\theta \in \mathbb{R}^d$,

$$\begin{aligned} |H_\lambda(\theta)|_{\mathsf{F}} &= \frac{|H(\theta)|_{\mathsf{F}}}{(1+\lambda^{3/2}|\theta|^{3(\rho+q-1)})^{1/3}} \leq \frac{4^{1/3}|H(\theta)|_{\mathsf{F}}}{1+\lambda^{1/2}|\theta|^{(\rho+q-1)}} \leq \frac{4^{1/3}\lambda^{-1/2}|H(\theta)|_{\mathsf{F}}}{1+|\theta|^{(\rho+q-1)}} \\ &\leq \frac{4^{1/3}\lambda^{-1/2}\sqrt{d}\mathsf{K}_1(1+|\theta|)^{\rho+q-1}}{1+|\theta|^{(\rho+q-1)}} \leq 2^{\rho+q-1}\lambda^{-1/2}\sqrt{d}\mathsf{K}_1. \end{aligned} \tag{57}$$

Substituting (57) into (56) yields

$$\begin{aligned} \mathbb{E}\left[|\Xi_{n,t}^\lambda|^2 \middle| \widetilde{\theta}_n^\lambda \right] &\leq 2\lambda\beta^{-1}(t-n)\left(d + 2^{\rho+q-1} d\mathsf{K}_1 + 2^{2(\rho+q-1)} d\mathsf{K}_1^2 \right) \\ &\leq 2^{2(\rho+q)}\lambda\beta^{-1}(t-n)d(1+\mathsf{K}_1)^2, \end{aligned} \tag{58}$$

where we use $1 + 2^{\rho+q-1} + 2^{2(\rho+q-1)} \leq 2^{2\rho+2q-3} + 2^{2\rho+2q-3} + 2^{2(\rho+q-1)} \leq 2^{2(\rho+q)-1}$ with $\rho \in [2,\infty)\cap\mathbb{N}$ and $q \in (0,1]$ to obtain the last inequality. To upper bound the first term on the RHS of (55), we use (52) to obtain

$$|\Delta_{n,t}^\lambda|^2 = |\widetilde{\theta}_n^\lambda|^2 + 2\lambda(t-n)\left\langle \widetilde{\theta}_n^\lambda, \phi^\lambda(\widetilde{\theta}_n^\lambda) \right\rangle + \lambda^2(t-n)^2|\phi^\lambda(\widetilde{\theta}_n^\lambda)|^2. \tag{59}$$

By using (53), the second term on the RHS of (59) can be estimated as follows:

$$\begin{aligned} \left\langle \widetilde{\theta}_n^\lambda, \phi^\lambda(\widetilde{\theta}_n^\lambda) \right\rangle &= -\frac{\left\langle \widetilde{\theta}_n^\lambda, h(\widetilde{\theta}_n^\lambda) \right\rangle}{\left(1+\lambda^{3/2}|\widetilde{\theta}_n^\lambda|^{3(\rho+q-1)}\right)^{1/3}} + \frac{(\lambda/2)\left\langle \widetilde{\theta}_n^\lambda, H(\widetilde{\theta}_n^\lambda)h(\widetilde{\theta}_n^\lambda) \right\rangle}{\left(1+\lambda^{3/2}|\widetilde{\theta}_n^\lambda|^{3(\rho+q-1)}\right)^{2/3}} \\ &\quad - \frac{(\lambda/2)\beta^{-1}\left\langle \widetilde{\theta}_n^\lambda, \Upsilon(\widetilde{\theta}_n^\lambda) \right\rangle}{\left(1+\lambda^{3/2}|\widetilde{\theta}_n^\lambda|^{3(\rho+q-1)}\right)^{1/3}} \\ &\leq -\frac{\mathsf{a}_{\mathsf{D}}|\widetilde{\theta}_n^\lambda|^{\rho+q+1} - \mathsf{b}_{\mathsf{D}}}{\left(1+\lambda^{3/2}|\widetilde{\theta}_n^\lambda|^{3(\rho+q-1)}\right)^{1/3}} + \frac{2\lambda K_h K_H\left(1+|\widetilde{\theta}_n^\lambda|^{2(\rho+q)}\right)}{\left(1+\lambda^{3/2}|\widetilde{\theta}_n^\lambda|^{3(\rho+q-1)}\right)^{2/3}} \\ &\quad + \frac{\lambda\beta^{-1}2^{\rho+q-3}\mathsf{K}_{3,d}\left(1+|\widetilde{\theta}_n^\lambda|^{\rho+q-1}\right)}{\left(1+\lambda^{3/2}|\widetilde{\theta}_n^\lambda|^{3(\rho+q-1)}\right)^{1/3}} \\ &\leq -\frac{\mathsf{a}_{\mathsf{D}}|\widetilde{\theta}_n^\lambda|^{\rho+q+1}}{\left(1+\lambda^{3/2}|\widetilde{\theta}_n^\lambda|^{3(\rho+q-1)}\right)^{1/3}} + \frac{2\lambda K_h K_H|\widetilde{\theta}_n^\lambda|^{2(\rho+q)}}{\left(1+\lambda^{3/2}|\widetilde{\theta}_n^\lambda|^{3(\rho+q-1)}\right)^{2/3}} \\ &\quad + \mathsf{b}_{\mathsf{D}} + 2K_hK_H + \beta^{-1}2^{\rho+q-2}\mathsf{K}_{3,d}, \end{aligned} \tag{60}$$

where the first inequality holds due to Remark 2.4 with $r = \rho+q-1$ and the following results: for all $\theta \in \mathbb{R}^d$,

$$\begin{aligned} |\theta||H(\theta)||h(\theta)| &\leq K_H K_h |\theta|(1+|\theta|^{\rho+q-1})(1+|\theta|^{\rho+q}) \\ &= K_H K_h(|\theta| + |\theta|^{\rho+q} + |\theta|^{\rho+q+1} + |\theta|^{2(\rho+q)}) \leq 4K_hK_H(1+|\theta|^{2(\rho+q)}), \\ |\theta||\Upsilon(\theta)| &\leq \mathsf{K}_{3,d}|\theta|(1+|\theta|)^{\rho+q-2} \leq \mathsf{K}_{3,d}(1+|\theta|)^{\rho+q-1} \leq 2^{\rho+q-2}\mathsf{K}_{3,d}(1+|\theta|^{\rho+q-1}), \end{aligned}$$

where we recall from our assumptions that $\rho \in [2,\infty)\cap\mathbb{N}$ and $q \in (0,1]$, and where the last inequality holds since, for all $\theta \in \mathbb{R}^d$,

$$\left(1+\lambda^{3/2}|\theta|^{3(\rho+q-1)}\right)^{1/3} \geq 4^{-1/3}(1+\lambda^{1/2}|\theta|^{\rho+q-1}) \geq 4^{-1/3}\lambda^{1/2}(1+|\theta|^{\rho+q-1}).$$

Similarly, the third term on the RHS of (59) can be upper bounded as follows:

$$|\phi^\lambda(\widetilde{\theta}_n^\lambda)|^2 = |-h_\lambda(\widetilde{\theta}_n^\lambda) + (\lambda/2)H_\lambda(\widetilde{\theta}_n^\lambda)h_\lambda(\widetilde{\theta}_n^\lambda) - (\lambda/2)\beta^{-1}\Upsilon_\lambda(\widetilde{\theta}_n^\lambda)|^2$$

$$
\begin{aligned}
&= |h_\lambda(\widetilde{\theta}^\lambda_n)|^2 + (\lambda^2/4)|H_\lambda(\widetilde{\theta}^\lambda_n)h_\lambda(\widetilde{\theta}^\lambda_n)|^2 + (\lambda^2/4)\beta^{-2}|\Upsilon_\lambda(\widetilde{\theta}^\lambda_n)|^2 \\
&\quad - \lambda\left\langle h_\lambda(\widetilde{\theta}^\lambda_n), H_\lambda(\widetilde{\theta}^\lambda_n)h_\lambda(\widetilde{\theta}^\lambda_n)\right\rangle + \lambda\beta^{-1}\left\langle h_\lambda(\widetilde{\theta}^\lambda_n), \Upsilon_\lambda(\widetilde{\theta}^\lambda_n)\right\rangle \\
&\quad - (\lambda^2/2)\beta^{-1}\left\langle H_\lambda(\widetilde{\theta}^\lambda_n)h_\lambda(\widetilde{\theta}^\lambda_n), \Upsilon_\lambda(\widetilde{\theta}^\lambda_n)\right\rangle. \qquad (61)
\end{aligned}
$$

By using Assumption 2, Remark 2.2, and the inequalities $(u+w)^v \geq 2^{v-1}(u^v+w^v)$, $(u+w)^\nu \leq 2^{\nu-1}(u^\nu+w^\nu)$, $u,w\geq 0$, $0<v\leq 1$, $\nu\geq 1$, $|\theta|^k \leq 1+|\theta|^l$, $\theta\in\mathbb{R}^d$, $k\leq l$, we obtain

$$
\begin{aligned}
|h_\lambda(\widetilde{\theta}^\lambda_n)|^2 &\leq \frac{K_h^2\left(1+|\widetilde{\theta}^\lambda_n|^{\rho+q}\right)^2}{\left(1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}\right)^{2/3}} \leq \frac{2K_h^2\left(1+|\widetilde{\theta}^\lambda_n|^{2(\rho+q)}\right)}{\left(1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}\right)^{2/3}} \\
&\leq \frac{2K_h^2|\widetilde{\theta}^\lambda_n|^{2(\rho+q)}}{\left(1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}\right)^{2/3}} + 2K_h^2, \qquad (62) \\
(\lambda^2/4)|H_\lambda(\widetilde{\theta}^\lambda_n)h_\lambda(\widetilde{\theta}^\lambda_n)|^2 &\leq \frac{(\lambda^2/4)K_H^2K_h^2\left(1+|\widetilde{\theta}^\lambda_n|^{\rho+q-1}\right)^2\left(1+|\widetilde{\theta}^\lambda_n|^{\rho+q}\right)^2}{\left(1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}\right)^{4/3}} \\
&\leq \frac{\lambda^2K_H^2K_h^2\left(1+|\widetilde{\theta}^\lambda_n|^{2(\rho+q-1)}+|\widetilde{\theta}^\lambda_n|^{2(\rho+q)}+|\widetilde{\theta}^\lambda_n|^{4(\rho+q)-2}\right)}{\left(1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}\right)^{4/3}} \\
&\leq \frac{3\lambda^2K_H^2K_h^2\left(1+|\widetilde{\theta}^\lambda_n|^{4(\rho+q)-2}\right)}{\left(1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}\right)^{4/3}} \\
&\leq \frac{3\lambda^2K_H^2K_h^2|\widetilde{\theta}^\lambda_n|^{4(\rho+q)-2}}{\left(1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}\right)^{4/3}} + 3K_H^2K_h^2, \qquad (63) \\
(\lambda^2/4)\beta^{-2}|\Upsilon_\lambda(\widetilde{\theta}^\lambda_n)|^2 &\leq \frac{(\lambda^2/4)\beta^{-2}\mathsf{K}_{3,d}^2\left(1+|\widetilde{\theta}^\lambda_n|\right)^{2(\rho+q-2)}}{\left(1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}\right)^{2/3}} \\
&\leq \frac{2^{2(\rho+q)-7}\lambda^2\beta^{-2}\mathsf{K}_{3,d}^2\left(1+|\widetilde{\theta}^\lambda_n|^{2(\rho+q-2)}\right)}{\left(1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}\right)^{2/3}} \\
&\leq \frac{2^{2(\rho+q)-6}\lambda^2\beta^{-2}\mathsf{K}_{3,d}^2\left(1+|\widetilde{\theta}^\lambda_n|^{2(\rho+q-2)}\right)}{1+\lambda|\widetilde{\theta}^\lambda_n|^{2(\rho+q-1)}} \\
&\leq 2^{2(\rho+q)-6}\beta^{-2}\mathsf{K}_{3,d}^2, \qquad (64) \\
-\lambda\left\langle h_\lambda(\widetilde{\theta}^\lambda_n), H_\lambda(\widetilde{\theta}^\lambda_n)h_\lambda(\widetilde{\theta}^\lambda_n)\right\rangle &\leq \frac{\lambda K_h^2K_H\left(1+|\widetilde{\theta}^\lambda_n|^{\rho+q}\right)^2\left(1+|\widetilde{\theta}^\lambda_n|^{\rho+q-1}\right)}{1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}} \\
&\leq \frac{2\lambda K_h^2K_H\left(1+|\widetilde{\theta}^\lambda_n|^{2(\rho+q)}+|\widetilde{\theta}^\lambda_n|^{\rho+q-1}+|\widetilde{\theta}^\lambda_n|^{3(\rho+q)-1}\right)}{1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}} \\
&\leq \frac{6\lambda K_h^2K_H\left(1+|\widetilde{\theta}^\lambda_n|^{3(\rho+q)-1}\right)}{1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}} \\
&\leq \frac{6\lambda K_h^2K_H|\widetilde{\theta}^\lambda_n|^{3(\rho+q)-1}}{1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}} + 6K_h^2K_H, \qquad (65)
\end{aligned}
$$

$$\begin{aligned}
\lambda\beta^{-1}\left\langle h_\lambda(\widetilde{\theta}^\lambda_n), \Upsilon_\lambda(\widetilde{\theta}^\lambda_n)\right\rangle &\leq \frac{\lambda\beta^{-1}K_h\mathsf{K}_{3,d}\left(1+|\widetilde{\theta}^\lambda_n|^{\rho+q}\right)\left(1+|\widetilde{\theta}^\lambda_n|\right)^{\rho+q-2}}{\left(1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}\right)^{2/3}} \\
&\leq \frac{\lambda\beta^{-1}K_h\mathsf{K}_{3,d}2^{\rho+q-3}\left(1+|\widetilde{\theta}^\lambda_n|^{\rho+q}+|\widetilde{\theta}^\lambda_n|^{\rho+q-2}+|\widetilde{\theta}^\lambda_n|^{2(\rho+q-1)}\right)}{\left(1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}\right)^{2/3}} \\
&\leq \frac{3\lambda\beta^{-1}K_h\mathsf{K}_{3,d}2^{\rho+q-2}\left(1+|\widetilde{\theta}^\lambda_n|^{2(\rho+q-1)}\right)}{1+\lambda|\widetilde{\theta}^\lambda_n|^{2(\rho+q-1)}} \\
&\leq \beta^{-1}K_h\mathsf{K}_{3,d}2^{\rho+q}, \qquad (66)
\end{aligned}$$

$$\begin{aligned}
&-(\lambda^2/2)\beta^{-1}\left\langle H_\lambda(\widetilde{\theta}^\lambda_n)h_\lambda(\widetilde{\theta}^\lambda_n), \Upsilon_\lambda(\widetilde{\theta}^\lambda_n)\right\rangle \\
&\quad\leq \frac{(\lambda^2/2)\beta^{-1}K_HK_h\mathsf{K}_{3,d}\left(1+|\widetilde{\theta}^\lambda_n|^{\rho+q-1}\right)\left(1+|\widetilde{\theta}^\lambda_n|^{\rho+q}\right)\left(1+|\widetilde{\theta}^\lambda_n|\right)^{\rho+q-2}}{1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}} \\
&\quad\leq \frac{\lambda^2\beta^{-1}K_HK_h\mathsf{K}_{3,d}2^{\rho+q-4}\left(1+|\widetilde{\theta}^\lambda_n|^{\rho+q-1}+|\widetilde{\theta}^\lambda_n|^{\rho+q}+|\widetilde{\theta}^\lambda_n|^{2(\rho+q)-1}\right)\left(1+|\widetilde{\theta}^\lambda_n|^{\rho+q-2}\right)}{1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}} \\
&\quad\leq \frac{7\lambda^2\beta^{-1}K_HK_h\mathsf{K}_{3,d}2^{\rho+q-4}\left(1+|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}\right)}{1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}} \leq \beta^{-1}K_HK_h\mathsf{K}_{3,d}2^{\rho+q-1}. \qquad (67)
\end{aligned}$$

Substituting (62) – (67) into (61) yields

$$\begin{aligned}
|\phi^\lambda(\widetilde{\theta}^\lambda_n)|^2 \leq{}& \frac{2K_h^2|\widetilde{\theta}^\lambda_n|^{2(\rho+q)}}{\left(1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}\right)^{2/3}}+2K_h^2+\frac{3\lambda^2K_H^2K_h^2|\widetilde{\theta}^\lambda_n|^{4(\rho+q)-2}}{\left(1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}\right)^{4/3}}+3K_H^2K_h^2 \\
&+\frac{6\lambda K_h^2K_H|\widetilde{\theta}^\lambda_n|^{3(\rho+q)-1}}{1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}}+6K_h^2K_H+2^{2(\rho+q)-6}\beta^{-2}\mathsf{K}_{3,d}^2+\beta^{-1}K_h\mathsf{K}_{3,d}2^{\rho+q} \qquad (68)\\
&+\beta^{-1}K_HK_h\mathsf{K}_{3,d}2^{\rho+q-1}.
\end{aligned}$$

Combining the results in (60) and (68), we obtain the following upper bound for (59):

$$\begin{aligned}
|\Delta^\lambda_{n,t}|^2 \leq{}& |\widetilde{\theta}^\lambda_n|^2-\frac{2\lambda(t-n)\mathsf{a}_\mathsf{D}|\widetilde{\theta}^\lambda_n|^{\rho+q+1}}{\left(1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}\right)^{1/3}}+\frac{4\lambda^2(t-n)K_hK_H|\widetilde{\theta}^\lambda_n|^{2(\rho+q)}}{\left(1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}\right)^{2/3}} \\
&+\frac{\lambda^2(t-n)2K_h^2|\widetilde{\theta}^\lambda_n|^{2(\rho+q)}}{\left(1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}\right)^{2/3}}+\frac{\lambda^4(t-n)3K_H^2K_h^2|\widetilde{\theta}^\lambda_n|^{4(\rho+q)-2}}{\left(1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}\right)^{4/3}} \\
&+\frac{\lambda^3(t-n)6K_h^2K_H|\widetilde{\theta}^\lambda_n|^{3(\rho+q)-1}}{1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}}+\lambda(t-n)\Big(2\mathsf{b}_\mathsf{D}+4K_hK_H \\
&+\beta^{-1}2^{\rho+q-1}\mathsf{K}_{3,d}+2K_h^2+3K_H^2K_h^2+6K_h^2K_H \\
&+2^{2(\rho+q)-6}\beta^{-2}\mathsf{K}_{3,d}^2+\beta^{-1}K_h\mathsf{K}_{3,d}2^{\rho+q}+\beta^{-1}K_HK_h\mathsf{K}_{3,d}2^{\rho+q-1}\Big) \\
={}& \left(1-\frac{\lambda(t-n)\mathsf{a}_\mathsf{D}|\widetilde{\theta}^\lambda_n|^{\rho+q-1}}{\left(1+\lambda^{3/2}|\widetilde{\theta}^\lambda_n|^{3(\rho+q-1)}\right)^{1/3}}\right)|\widetilde{\theta}^\lambda_n|^2-\lambda(t-n)\mathfrak{I}^\lambda_1(\widetilde{\theta}^\lambda_n)+\lambda(t-n)\mathsf{c}_1, \qquad (69)
\end{aligned}$$

where for any $\theta\in\mathbb{R}^d$

$$\mathfrak{I}^\lambda_1(\theta) := \frac{\mathsf{a}_\mathsf{D}|\theta|^{\rho+q+1}}{(1+\lambda^{3/2}|\theta|^{3(\rho+q-1)})^{1/3}}-\frac{\lambda(4K_hK_H+2K_h^2)|\theta|^{2(\rho+q)}}{(1+\lambda^{3/2}|\theta|^{3(\rho+q-1)})^{2/3}}$$

$$
- \frac{\lambda^3 3K_H^2 K_h^2 |\theta|^{4(\rho+q)-2}}{\left(1+\lambda^{3/2}|\theta|^{3(\rho+q-1)}\right)^{4/3}} - \frac{\lambda^2 6K_h^2 K_H |\widetilde{\theta}_n^\lambda|^{3(\rho+q)-1}}{1+\lambda^{3/2}|\widetilde{\theta}_n^\lambda|^{3(\rho+q-1)}},
$$

and where

$$
\begin{aligned}
\mathsf{c}_1 := \Big(& 2\mathsf{b}_\mathsf{D} + 4K_h K_H + \beta^{-1}2^{\rho+q-1}\mathsf{K}_{3,d} + 2K_h^2 + 3K_H^2K_h^2 + 6K_h^2K_H \\
&+ 2^{2(\rho+q)-6}\beta^{-2}\mathsf{K}_{3,d}^2 + \beta^{-1}K_h\mathsf{K}_{3,d}2^{\rho+q} + \beta^{-1}K_HK_h\mathsf{K}_{3,d}2^{\rho+q-1}\Big).
\end{aligned}
$$

We note that, for $0 < \lambda \leq \lambda_{\max} \leq \min\{(19\mathsf{a}_\mathsf{D}/240K_hK_H)^2, (19\mathsf{a}_\mathsf{D}/240K_h^2)^2, (\mathsf{a}_\mathsf{D}/120K_h^2K_H^2)^{2/3}, \mathsf{a}_\mathsf{D}/(480K_h^2K_H)\}$, and for all $\theta \in \mathbb{R}^d$,

$$
\mathfrak{I}_1^\lambda(\theta) \geq 0. \tag{70}
$$

Indeed, by using the expression of $\mathfrak{I}_1^\lambda$, we have that, for all $\theta \in \mathbb{R}^d$,

$$
\begin{aligned}
\mathfrak{I}_1^\lambda(\theta) = &\frac{(38\mathsf{a}_\mathsf{D}/60 + 19\mathsf{a}_\mathsf{D}/60)|\theta|^{\rho+q+1}\left(1+\lambda^{3/2}|\theta|^{3(\rho+q-1)}\right)^{1/3} - \lambda(4K_hK_H+2K_h^2)|\theta|^{2(\rho+q)}}{\left(1+\lambda^{3/2}|\theta|^{3(\rho+q-1)}\right)^{2/3}} \\
&+ \frac{(\mathsf{a}_\mathsf{D}/40)|\theta|^{\rho+q+1}\left(1+\lambda^{3/2}|\theta|^{3(\rho+q-1)}\right) - \lambda^3 3K_H^2K_h^2|\theta|^{4(\rho+q)-2}}{\left(1+\lambda^{3/2}|\theta|^{3(\rho+q-1)}\right)^{4/3}} \\
&+ \frac{(\mathsf{a}_\mathsf{D}/40)|\theta|^{\rho+q+1}\left(1+\lambda^{3/2}|\theta|^{3(\rho+q-1)}\right)^{2/3} - \lambda^2 6K_h^2K_H|\theta|^{3(\rho+q)-1}}{1+\lambda^{3/2}|\theta|^{3(\rho+q-1)}} \\
\geq &\frac{(19\mathsf{a}_\mathsf{D}/60 + 19\mathsf{a}_\mathsf{D}/120)\lambda^{1/2}|\theta|^{2(\rho+q)} - \lambda(4K_hK_H+2K_h^2)|\theta|^{2(\rho+q)}}{\left(1+\lambda^{3/2}|\theta|^{3(\rho+q-1)}\right)^{2/3}} \\
&+ \frac{(\mathsf{a}_\mathsf{D}/40)\lambda^{3/2}|\theta|^{4(\rho+q)-2} - \lambda^3 3K_H^2K_h^2|\theta|^{4(\rho+q)-2}}{\left(1+\lambda^{3/2}|\theta|^{3(\rho+q-1)}\right)^{4/3}} \\
&+ \frac{(\mathsf{a}_\mathsf{D}/80)\lambda|\theta|^{3(\rho+q)-1} - \lambda^2 6K_h^2K_H|\theta|^{3(\rho+q)-1}}{1+\lambda^{3/2}|\theta|^{3(\rho+q-1)}} \\
\geq &\, 0,
\end{aligned}
$$

where the last inequality holds due to $0 < \lambda \leq \lambda_{\max}$. Substituting (70) into (69) yields

$$
|\Delta_{n,t}^\lambda|^2 \leq \left(1 - \frac{\lambda(t-n)\mathsf{a}_\mathsf{D}|\widetilde{\theta}_n^\lambda|^{\rho+q-1}}{\left(1+\lambda^{3/2}|\widetilde{\theta}_n^\lambda|^{3(\rho+q-1)}\right)^{1/3}}\right)|\widetilde{\theta}_n^\lambda|^2 + \lambda(t-n)\mathsf{c}_1.
$$

In addition, since $f(s) = s/(1+\lambda^{3/2}s^3)^{1/3}$ is non-decreasing for all $s \geq 0$, we obtain that, for all $|\theta| > \mathsf{M}_1 > 0$,

$$
\frac{|\theta|^{\rho+q-1}}{\left(1+\lambda^{3/2}|\theta|^{3(\rho+q-1)}\right)^{1/3}} \geq \frac{\mathsf{M}_1^{\rho+q-1}}{\left(1+\lambda^{3/2}\mathsf{M}_1^{3(\rho+q-1)}\right)^{1/3}} \geq \frac{\mathsf{M}_1^{\rho+q-1}}{\left(1+\mathsf{M}_1^{3(\rho+q-1)}\right)^{1/3}} =: \kappa.
$$

Denote by $\mathsf{S}_{n,\mathsf{M}_1} := \{\omega \in \Omega : |\widetilde{\theta}_n^\lambda(\omega)| > \mathsf{M}_1\}$. The above inequality further implies that,

$$
|\Delta_{n,t}^\lambda|^2 \mathbb{1}_{\mathsf{S}_{n,\mathsf{M}_1}} \leq (1-\lambda(t-n)\mathsf{a}_\mathsf{D}\kappa)\,|\widetilde{\theta}_n^\lambda|^2\mathbb{1}_{\mathsf{S}_{n,\mathsf{M}_1}} + \lambda(t-n)\mathsf{c}_1\mathbb{1}_{\mathsf{S}_{n,\mathsf{M}_1}}.
$$

Similarly, we have that

$$
|\Delta_{n,t}^\lambda|^2 \mathbb{1}_{\mathsf{S}^\mathsf{c}_{n,\mathsf{M}_1}} \leq (1-\lambda(t-n)\mathsf{a}_\mathsf{D}\kappa)\,|\widetilde{\theta}_n^\lambda|^2\mathbb{1}_{\mathsf{S}^\mathsf{c}_{n,\mathsf{M}_1}} + \lambda(t-n)(\mathsf{a}_\mathsf{D}\kappa\mathsf{M}_1^2 + \mathsf{c}_1)\mathbb{1}_{\mathsf{S}^\mathsf{c}_{n,\mathsf{M}_1}}.
$$

Combining the two cases yields

$$
|\Delta_{n,t}^\lambda|^2 \leq (1-\lambda(t-n)\mathsf{a}_\mathsf{D}\kappa)\,|\widetilde{\theta}_n^\lambda|^2 + \lambda(t-n)(\mathsf{a}_\mathsf{D}\kappa\mathsf{M}_1^2 + \mathsf{c}_1). \tag{71}
$$

Finally, by substituting (58) and (71) into (55), we obtain

$$
\mathbb{E}\left[|\widetilde{\theta}_t^\lambda|^2 \middle| \widetilde{\theta}_n^\lambda\right] \leq (1-\lambda(t-n)\mathsf{a}_\mathsf{D}\kappa)\,|\widetilde{\theta}_n^\lambda|^2 + \lambda(t-n)\mathsf{c}_0, \tag{72}
$$

where

$$\begin{aligned}
\kappa &:= \mathsf{M}_1^{\rho+q-1}/\left(1+\mathsf{M}_1^{3(\rho+q-1)}\right)^{1/3} \quad (\text{with } \mathsf{M}_1>0),\\
\mathsf{c}_0 &:= \mathsf{a}_\mathsf{D}\kappa\mathsf{M}_1^2+\mathsf{c}_1+2^{2(\rho+q)}\beta^{-1}d(1+\mathsf{K}_1)^2,\\
\mathsf{c}_1 &:= \Big(2\mathsf{b}_\mathsf{D}+4K_hK_H+\beta^{-1}2^{\rho+q-1}\mathsf{K}_{3,d}+2K_h^2+3K_H^2K_h^2+6K_h^2K_H\\
&\qquad +2^{2(\rho+q)-6}\beta^{-2}\mathsf{K}_{3,d}^2+\beta^{-1}K_h\mathsf{K}_{3,d}2^{\rho+q}+\beta^{-1}K_HK_h\mathsf{K}_{3,d}2^{\rho+q-1}\Big).
\end{aligned} \tag{73}$$

We observe that, for $0<\lambda\le\lambda_{\max}\le 1/\mathsf{a}_\mathsf{D}$,

$$1>1-\lambda(t-n)\mathsf{a}_\mathsf{D}\kappa>1-\lambda\mathsf{a}_\mathsf{D}\ge 0,$$

then, by induction, (72) implies, for $t\in(n,n+1]$, $n\in\mathbb{N}_0$, $0<\lambda\le\lambda_{\max}\le 1$, that,

$$\begin{aligned}
\mathbb{E}\left[|\widetilde{\theta}_t^\lambda|^2\right] &\le (1-\lambda(t-n)\mathsf{a}_\mathsf{D}\kappa)\,\mathbb{E}\left[|\widetilde{\theta}_n^\lambda|^2\right]+\lambda(t-n)\mathsf{c}_0\\
&\le (1-\lambda(t-n)\mathsf{a}_\mathsf{D}\kappa)(1-\lambda\mathsf{a}_\mathsf{D}\kappa)\,\mathbb{E}\left[|\widetilde{\theta}_{n-1}^\lambda|^2\right]+\mathsf{c}_0+\lambda\mathsf{c}_0\\
&\le (1-\lambda(t-n)\mathsf{a}_\mathsf{D}\kappa)(1-\lambda\mathsf{a}_\mathsf{D}\kappa)^2\,\mathbb{E}\left[|\widetilde{\theta}_{n-2}^\lambda|^2\right]+\mathsf{c}_0+\lambda\mathsf{c}_0\left(1+(1-\lambda\mathsf{a}_\mathsf{D}\kappa)\right)\\
&\le \dots\\
&\le (1-\lambda(t-n)\mathsf{a}_\mathsf{D}\kappa)(1-\lambda\mathsf{a}_\mathsf{D}\kappa)^n\,\mathbb{E}\left[|\theta_0|^2\right]+\mathsf{c}_0\left(1+1/(\mathsf{a}_\mathsf{D}\kappa)\right),
\end{aligned}$$

which completes the proof. □

Proof of Lemma 6.3-(ii). For any $p\in[2,\infty)\cap\mathbb{N}$, $0<\lambda\le\lambda_{\max}\le 1$ with $\lambda_{\max}$ given in (14), $t\in(n,n+1]$, $n\in\mathbb{N}_0$, by using the same arguments as in the proof of [18, Lemma 4.2-(ii)] up to the inequality before [18, Eq. (134)] and by using (26) with (52), we obtain that

$$\begin{aligned}
\mathbb{E}\left[|\widetilde{\theta}_t^\lambda|^{2p}\middle|\widetilde{\theta}_n^\lambda\right] &\le |\Delta_{n,t}^\lambda|^{2p}+2^{2p-3}p(2p-1)|\Delta_{n,t}^\lambda|^{2p-2}\mathbb{E}\left[|\Xi_{n,t}^\lambda|^2\middle|\widetilde{\theta}_n^\lambda\right]\\
&\quad+2^{2p-3}p(2p-1)\mathbb{E}\left[|\Xi_{n,t}^\lambda|^{2p}\middle|\widetilde{\theta}_n^\lambda\right].
\end{aligned} \tag{74}$$

Then, by using (52) and (54), we can obtain an upper estimate for the last term in (74) as follows:

$$\begin{aligned}
\mathbb{E}\left[|\Xi_{n,t}^\lambda|^{2p}\middle|\widetilde{\theta}_n^\lambda\right] &= (2\lambda\beta^{-1})^p\mathbb{E}\left[\left\langle (B_t^\lambda-B_n^\lambda),\left(\psi^\lambda(\widetilde{\theta}_n^\lambda)\right)^2(B_t^\lambda-B_n^\lambda)\right\rangle^p\middle|\widetilde{\theta}_n^\lambda\right]\\
&\le (2\lambda\beta^{-1})^p\left|\mathsf{I}_d-\lambda H_\lambda(\widetilde{\theta}_n^\lambda)+(\lambda^2/3)(H_\lambda(\widetilde{\theta}_n^\lambda))^2\right|^p\mathbb{E}\left[|B_t^\lambda-B_n^\lambda|^{2p}\right]\\
&\le (2\lambda\beta^{-1}dp(2p-1)(t-n))^p3^{p-1}\left(1+\lambda^p|H_\lambda(\widetilde{\theta}_n^\lambda)|_\mathsf{F}^p+(\lambda^2/3)^p|H_\lambda(\widetilde{\theta}_n^\lambda)|_\mathsf{F}^{2p}\right)\\
&\le \lambda(t-n)(2\beta^{-1}dp(2p-1))^p3^{p-1}\left(1+\lambda^p|H_\lambda(\widetilde{\theta}_n^\lambda)|_\mathsf{F}^p\right)^2\\
&\le \lambda(t-n)(2\beta^{-1}dp(2p-1))^p3^{p-1}\left(1+2^{p(\rho+q-1)}d^{p/2}\mathsf{K}_1^p\right)^2,
\end{aligned} \tag{75}$$

where the last inequality holds due to (57). Substituting (58) and (75) into (74) yields

$$\begin{aligned}
\mathbb{E}\left[|\widetilde{\theta}_t^\lambda|^{2p}\middle|\widetilde{\theta}_n^\lambda\right] &\le |\Delta_{n,t}^\lambda|^{2p}+\lambda(t-n)2^{2(p+\rho+q)-3}p(2p-1)\beta^{-1}d(1+\mathsf{K}_1)^2|\Delta_{n,t}^\lambda|^{2p-2}\\
&\quad+\lambda(t-n)\mathsf{c}_\Xi(p),
\end{aligned} \tag{76}$$

where $\mathsf{c}_\Xi(p):=2^{2p-4}(\beta^{-1}d)^p(2p(2p-1))^{p+1}3^{p-1}\left(1+2^{p(\rho+q-1)}d^{p/2}\mathsf{K}_1^p\right)^2$. Next, we apply (71) to obtain

$$\begin{aligned}
|\Delta_{n,t}^\lambda|^{2p} &\le \left((1-\lambda(t-n)\mathsf{a}_\mathsf{D}\kappa)|\widetilde{\theta}_n^\lambda|^2+\lambda(t-n)(\mathsf{a}_\mathsf{D}\kappa\mathsf{M}_1^2+\mathsf{c}_1)\right)^p\\
&\le (1+\lambda(t-n)\mathsf{a}_\mathsf{D}\kappa/2)^{p-1}(1-\lambda(t-n)\mathsf{a}_\mathsf{D}\kappa)^p|\widetilde{\theta}_n^\lambda|^{2p}\\
&\quad+(1+2/(\lambda(t-n)\mathsf{a}_\mathsf{D}\kappa))^{p-1}(\lambda(t-n))^p(\mathsf{a}_\mathsf{D}\kappa\mathsf{M}_1^2+\mathsf{c}_1)^p\\
&\le (1-\lambda(t-n)\mathsf{a}_\mathsf{D}\kappa/2)^{p-1}(1-\lambda(t-n)\mathsf{a}_\mathsf{D}\kappa)|\widetilde{\theta}_n^\lambda|^{2p}\\
&\quad+\lambda(t-n)(1+2/(\mathsf{a}_\mathsf{D}\kappa))^{p-1}(\mathsf{a}_\mathsf{D}\kappa\mathsf{M}_1^2+\mathsf{c}_1)^p
\end{aligned}$$

$$= \bar{\mathsf{c}}_{n,t}^{\lambda}(p)|\widetilde{\theta}_n^{\lambda}|^{2p} + \widetilde{\mathsf{c}}_{n,t}^{\lambda}(p), \tag{77}$$

where the second inequality holds due to $(u+v)^p \leq (1+\varepsilon)^{p-1}u^p + (1+\varepsilon^{-1})^{p-1}v^p$, $u,v \geq 0$, $\varepsilon > 0$ with $\varepsilon = \lambda(t-n)\mathsf{a}_{\mathsf{D}}\kappa/2$, and where

$$\bar{\mathsf{c}}_{n,t}^{\lambda}(p) := (1-\lambda(t-n)\mathsf{a}_{\mathsf{D}}\kappa/2)^{p-1}\left(1-\lambda(t-n)\mathsf{a}_{\mathsf{D}}\kappa\right),$$
$$\widetilde{\mathsf{c}}_{n,t}^{\lambda}(p) := \lambda(t-n)\left(1+2/(\mathsf{a}_{\mathsf{D}}\kappa)\right)^{p-1}\left(\mathsf{a}_{\mathsf{D}}\kappa\mathsf{M}_1^2 + \mathsf{c}_1\right)^p.$$

In addition, we observe that by (77),

$$|\Delta_{n,t}^{\lambda}|^{2p-2} \leq \bar{\mathsf{c}}_{n,t}^{\lambda}(p-1)|\widetilde{\theta}_n^{\lambda}|^{2p-2} + \widetilde{\mathsf{c}}_{n,t}^{\lambda}(p-1), \tag{78}$$

and, in particular, when $p=2$, (78) yields $|\Delta_{n,t}^{\lambda}|^2 \leq \bar{\mathsf{c}}_{n,t}^{\lambda}(1)|\widetilde{\theta}_n^{\lambda}|^2 + \widetilde{\mathsf{c}}_{n,t}^{\lambda}(1)$ which is exactly the upper bound (71). By substituting (77) and (78) into (76), we have that

$$\begin{aligned}\mathbb{E}\left[|\widetilde{\theta}_t^{\lambda}|^{2p}\middle|\widetilde{\theta}_n^{\lambda}\right] &\leq \bar{\mathsf{c}}_{n,t}^{\lambda}(p)|\widetilde{\theta}_n^{\lambda}|^{2p} + \widetilde{\mathsf{c}}_{n,t}^{\lambda}(p) + \lambda(t-n)\mathsf{c}_{\Xi}(p)\\ &\quad + \lambda(t-n)2^{2(p+\rho+q)-3}p(2p-1)\beta^{-1}d(1+\mathsf{K}_1)^2\\ &\quad \times\left(\bar{\mathsf{c}}_{n,t}^{\lambda}(p-1)|\widetilde{\theta}_n^{\lambda}|^{2p-2} + \widetilde{\mathsf{c}}_{n,t}^{\lambda}(p-1)\right).\end{aligned} \tag{79}$$

Denote by $\mathsf{M}_2(p) := (2^{2(p+\rho+q)-1}p(2p-1)\beta^{-1}d(1+\mathsf{K}_1)^2/(\mathsf{a}_{\mathsf{D}}\kappa))^{1/2}$. For all $|\theta| > \mathsf{M}_2(p)$, we have that

$$(\lambda(t-n)\mathsf{a}_{\mathsf{D}}\kappa/4)|\theta|^{2p} > (\lambda(t-n)2^{2(p+\rho+q)-1}p(2p-1)\beta^{-1}d(1+\mathsf{K}_1)^2/4)|\theta|^{2p-2}.$$

Denote by $\mathsf{S}_{n,\mathsf{M}_2(p)} := \{\omega\in\Omega : |\widetilde{\theta}_n^{\lambda}(\omega)| > \mathsf{M}_2(p)\}$. By using the above inequality, (79) can be further bounded as follows:

$$\begin{aligned}&\mathbb{E}\left[|\widetilde{\theta}_t^{\lambda}|^{2p}\mathbb{1}_{\mathsf{S}_{n,\mathsf{M}_2(p)}}\middle|\widetilde{\theta}_n^{\lambda}\right]\\ &\leq (1-\lambda(t-n)\mathsf{a}_{\mathsf{D}}\kappa/4)\,\bar{\mathsf{c}}_{n,t}^{\lambda}(p-1)|\widetilde{\theta}_n^{\lambda}|^{2p}\mathbb{1}_{\mathsf{S}_{n,\mathsf{M}_2(p)}} + \left(\widetilde{\mathsf{c}}_{n,t}^{\lambda}(p) + \lambda(t-n)\mathsf{c}_{\Xi}(p)\right)\mathbb{1}_{\mathsf{S}_{n,\mathsf{M}_2(p)}}\\ &\quad + \lambda(t-n)2^{2(p+\rho+q)-3}p(2p-1)\beta^{-1}d(1+\mathsf{K}_1)^2\widetilde{\mathsf{c}}_{n,t}^{\lambda}(p-1)\mathbb{1}_{\mathsf{S}_{n,\mathsf{M}_2(p)}}\\ &\quad - (\lambda(t-n)\mathsf{a}_{\mathsf{D}}\kappa/4)\,\bar{\mathsf{c}}_{n,t}^{\lambda}(p-1)|\widetilde{\theta}_n^{\lambda}|^{2p}\mathbb{1}_{\mathsf{S}_{n,\mathsf{M}_2(p)}}\\ &\quad + (\lambda(t-n)2^{2(p+\rho+q)-1}p(2p-1)\beta^{-1}d(1+\mathsf{K}_1)^2/4)\bar{\mathsf{c}}_{n,t}^{\lambda}(p-1)|\widetilde{\theta}_n^{\lambda}|^{2p-2}\mathbb{1}_{\mathsf{S}_{n,\mathsf{M}_2(p)}}\\ &\leq (1-\lambda(t-n)\mathsf{a}_{\mathsf{D}}\kappa)\,|\widetilde{\theta}_n^{\lambda}|^{2p}\mathbb{1}_{\mathsf{S}_{n,\mathsf{M}_2(p)}} + \lambda(t-n)\left[(1+2/(\mathsf{a}_{\mathsf{D}}\kappa))^{p-1}\left(\mathsf{a}_{\mathsf{D}}\kappa\mathsf{M}_1^2+\mathsf{c}_1\right)^p + \mathsf{c}_{\Xi}(p)\right.\\ &\quad + 2^{2(p+\rho+q)-3}p(2p-1)\beta^{-1}d(1+\mathsf{K}_1)^2\\ &\quad\quad \left.\times\left((1+2/(\mathsf{a}_{\mathsf{D}}\kappa))^{p-2}\left(\mathsf{a}_{\mathsf{D}}\kappa\mathsf{M}_1^2+\mathsf{c}_1\right)^{p-1} + \mathsf{M}_2(p)^{2p-2}\right)\right]\mathbb{1}_{\mathsf{S}_{n,\mathsf{M}_2(p)}}\\ &\leq (1-\lambda(t-n)\mathsf{a}_{\mathsf{D}}\kappa)\,|\widetilde{\theta}_n^{\lambda}|^{2p}\mathbb{1}_{\mathsf{S}_{n,\mathsf{M}_2(p)}} + \lambda(t-n)\mathsf{c}_p\mathbb{1}_{\mathsf{S}_{n,\mathsf{M}_2(p)}},\end{aligned} \tag{80}$$

where

$$\begin{aligned}\mathsf{c}_p &:= (1+2/(\mathsf{a}_{\mathsf{D}}\kappa))^{p-1}\left(\mathsf{a}_{\mathsf{D}}\kappa\mathsf{M}_1^2+\mathsf{c}_1\right)^p + \mathsf{c}_{\Xi}(p) + 2^{2(p+\rho+q)-3}p(2p-1)\beta^{-1}d(1+\mathsf{K}_1)^2\\ &\quad\times\left((1+2/(\mathsf{a}_{\mathsf{D}}\kappa))^{p-2}\left(\mathsf{a}_{\mathsf{D}}\kappa\mathsf{M}_1^2+\mathsf{c}_1\right)^{p-1} + \mathsf{M}_2(p)^{2p-2}\right),\\ \mathsf{c}_{\Xi}(p) &:= 2^{2p-4}(\beta^{-1}d)^p(2p(2p-1))^{p+1}3^{p-1}\left(1+2^{p(\rho+q-1)}d^{p/2}\mathsf{K}_1^p\right)^2,\\ \mathsf{M}_2(p) &:= (2^{2(p+\rho+q)-1}p(2p-1)\beta^{-1}d(1+\mathsf{K}_1)^2/(\mathsf{a}_{\mathsf{D}}\kappa))^{1/2}\end{aligned} \tag{81}$$

with κ and c_1 given in (73). Similarly, by using (79), we have that

$$\mathbb{E}\left[|\widetilde{\theta}_t^{\lambda}|^{2p}\mathbb{1}_{\mathsf{S}^{\mathsf{c}}_{n,\mathsf{M}_2(p)}}\middle|\widetilde{\theta}_n^{\lambda}\right] \leq (1-\lambda(t-n)\mathsf{a}_{\mathsf{D}}\kappa)\,|\widetilde{\theta}_n^{\lambda}|^{2p}\mathbb{1}_{\mathsf{S}^{\mathsf{c}}_{n,\mathsf{M}_2(p)}} + \lambda(t-n)\mathsf{c}_p\mathbb{1}_{\mathsf{S}^{\mathsf{c}}_{n,\mathsf{M}_2(p)}}. \tag{82}$$

Combing (80) and (82) yields the desired result. □

Lemma A.2. *Let Assumptions 1, 2, and 3 hold. Then, for any* $0 < \lambda \leq \lambda_{\max}$, $p > 0, t \geq 0$, *we obtain*

$$\mathbb{E}\left[|\bar{\theta}_t^{\lambda} - \bar{\theta}_{\lfloor t\rfloor}^{\lambda}|^{2p}\right] \leq \lambda^p\left(e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t\rfloor}\overline{\mathsf{C}}_{\mathbf{A}0,p}\mathbb{E}[|\theta_0|^{4\lceil p\rceil(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}0,p}\right), \tag{83}$$

$$\mathbb{E}\left[|\bar{\zeta}_t^{\lambda,n} - \bar{\zeta}_{\lfloor t\rfloor}^{\lambda,n}|^{2p}\right] \leq \lambda^p\left(e^{-\lambda\mathsf{a}_{\mathsf{D}}\min\{\kappa,1/2\}\lfloor t\rfloor}\overline{\mathsf{C}}_{\mathbf{A}1,p}\mathbb{E}[|\theta_0|^{2\lceil p\rceil(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}1,p}\right), \tag{84}$$

where $\overline{\mathsf{C}}_{\mathbf{A}0,p}$ *and* $\widetilde{\mathsf{C}}_{\mathbf{A}0,p}$ *are given in* (85)*, and* $\overline{\mathsf{C}}_{\mathbf{A}1,p}$ *and* $\widetilde{\mathsf{C}}_{\mathbf{A}1,p}$ *are given in* (86).

Proof. To show that (83) holds, we use the definition of $(\overline{\theta}_t^\lambda)_{t\geq 0}$ given in (27) and obtain that, for any $t \geq 0$,

$$
\begin{aligned}
&\mathbb{E}\left[|\overline{\theta}_t^\lambda - \overline{\theta}_{\lfloor t\rfloor}^\lambda|^{2p}\right] \\
&= \mathbb{E}\left[\left|-\lambda\int_{\lfloor t\rfloor}^t h_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda)\,\mathrm{d}s + \lambda^2\int_{\lfloor t\rfloor}^t\int_{\lfloor s\rfloor}^s \left(H_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)h_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda) - \beta^{-1}\Upsilon_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)\right)\,\mathrm{d}r\,\mathrm{d}s\right.\right. \\
&\qquad\left.\left. -\lambda\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t\int_{\lfloor s\rfloor}^s H_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\,\mathrm{d}s + \sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t \mathrm{d}B_s^\lambda\right|^{2p}\right] \\
&\leq 5^{2p}\left(\lambda^{2p}\mathbb{E}\left[|h_\lambda(\overline{\theta}_{\lfloor t\rfloor}^\lambda)|^{2p}\right] + \lambda^{4p}\mathbb{E}\left[|H_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)|^{2p}|h_\lambda(\overline{\theta}_{\lfloor t\rfloor}^\lambda)|^{2p}\right]\right. \\
&\qquad + \lambda^{4p}\beta^{-2p}\mathbb{E}\left[|\Upsilon_\lambda(\overline{\theta}_{\lfloor t\rfloor}^\lambda)|^{2p}\right] + \lambda^p(2\beta^{-1})^p\mathbb{E}\left[\left|\int_{\lfloor t\rfloor}^t \mathrm{d}B_s^\lambda\right|^{2p}\right] \\
&\qquad\left. + \lambda^{3p}(2\beta^{-1})^p\mathbb{E}\left[|H_\lambda(\overline{\theta}_{\lfloor t\rfloor}^\lambda)|^{2p}\left|\int_{\lfloor t\rfloor}^t\int_{\lfloor s\rfloor}^s \mathrm{d}B_r^\lambda\,\mathrm{d}s\right|^{2p}\right]\right) \\
&\leq 5^{2p}\lambda^p\left(K_h^{2p}\mathbb{E}\left[(1+|\overline{\theta}_{\lfloor t\rfloor}^\lambda|^{\rho+q})^{2p}\right] + \beta^{-2p}\mathsf{K}_{3,d}^{2p}\mathbb{E}\left[(1+|\overline{\theta}_{\lfloor t\rfloor}^\lambda|)^{2p(\rho+q-2)}\right]\right. \\
&\qquad + K_H^{2p}K_h^{2p}\mathbb{E}\left[(1+|\overline{\theta}_{\lfloor t\rfloor}^\lambda|^{\rho+q-1})^{2p}(1+|\overline{\theta}_{\lfloor t\rfloor}^\lambda|^{\rho+q})^{2p}\right] + (2\beta^{-1}(p+1)(d+2p))^p \\
&\qquad\left. +(\beta^{-1}(4p+2)(d+4p))^pK_H^{2p}\left(\mathbb{E}\left[(1+|\overline{\theta}_{\lfloor t\rfloor}^\lambda|^{\rho+q-1})^{4p}\right]\right)^{1/2}\right) \\
&\leq \lambda^p5^{2p}\left(K_h^{2p} + \beta^{-2p}\mathsf{K}_{3,d}^{2p} + K_H^{2p}K_h^{2p} + (\beta^{-1}(4p+2)(d+4p))^p(1+K_H^{2p})\right) \\
&\qquad\times 2^{4\lceil p\rceil(\rho+1)}\left(\mathbb{E}\left[|\overline{\theta}_{\lfloor t\rfloor}^\lambda|^{4\lceil p\rceil(\rho+1)}\right]+1\right) \\
&\leq \lambda^p\left(e^{-\lambda\mathsf{a}_\mathsf{D}\kappa\lfloor t\rfloor}\overline{\mathsf{C}}_{\mathbf{A}0,p}\mathbb{E}[|\theta_0|^{4\lceil p\rceil(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}0,p}\right),
\end{aligned}
$$

where the first inequality holds due to $(\sum_{l=1}^v u_l)^w \leq v^w\sum_{l=1}^v u_l^w$, $v\in\mathbb{N}$, $u_l\geq 0$, $w>0$, the second inequality holds due to Remark 2.2, Cauchy-Schwarz inequality, and the following inequality:

$$
\mathbb{E}\left[\left|\int_{\lfloor t\rfloor}^t \mathrm{d}B_s^\lambda\right|^{2p}\right] \leq \max\{d^p, (p(d+2p-2))^p\} \leq ((p+1)(d+2p))^p,
$$

the fourth inequality holds due to Lemma 6.3, and where

$$
\begin{aligned}
\overline{\mathsf{C}}_{\mathbf{A}0,p} &:= 5^{2p}2^{4\lceil p\rceil(\rho+1)}\left(K_h^{2p} + \beta^{-2p}\mathsf{K}_{3,d}^{2p} + K_H^{2p}K_h^{2p} + (\beta^{-1}(4p+2)(d+4p))^p(1+K_H^{2p})\right), \\
\widetilde{\mathsf{C}}_{\mathbf{A}0,p} &:= \overline{\mathsf{C}}_{\mathbf{A}0,p}\left(\mathsf{c}_{2\lceil p\rceil(\rho+1)}\,(1+1/(\mathsf{a}_\mathsf{D}\kappa))+1\right).
\end{aligned}
\tag{85}
$$

The inequality (84) can be obtained by using similar arguments. More precisely, by using Definition 6.2 with (28), we obtain that, for any $t\geq 0$,

$$
\begin{aligned}
\mathbb{E}\left[|\overline{\zeta}_t^{\lambda,n} - \overline{\zeta}_{\lfloor t\rfloor}^{\lambda,n}|^{2p}\right] &= \mathbb{E}\left[\left|-\lambda\int_{\lfloor t\rfloor}^t h(\overline{\zeta}_s^{\lambda,n})\,\mathrm{d}s + \sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t \mathrm{d}B_s^\lambda\right|^{2p}\right] \\
&\leq 2^{2p}\left(\lambda^{2p}\mathbb{E}\left[\int_{\lfloor t\rfloor}^t |h(\overline{\zeta}_s^{\lambda,n})|^{2p}\,\mathrm{d}s\right] + \lambda^p(2\beta^{-1})^p\mathbb{E}\left[\left|\int_{\lfloor t\rfloor}^t \mathrm{d}B_s^\lambda\right|^{2p}\right]\right) \\
&\leq 2^{2p}\lambda^p\left(K_h^{2p}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[(1+|\overline{\zeta}_s^{\lambda,n}|^{\rho+q})^{2p}\right]\,\mathrm{d}s + (2\beta^{-1}(p+1)(d+2p))^p\right)
\end{aligned}
$$

$$\leq \lambda^p \left(2^{2p} K_h^{2p} \int_{\lfloor t \rfloor}^t \mathbb{E}\left[V_{2\lceil p \rceil(\rho+1)}(\overline{\zeta}_s^{\lambda,n}) \right] \mathrm{d}s + 2^{2p}(2\beta^{-1}(p+1)(d+2p))^p \right)$$
$$\leq \lambda^p \left(e^{-\lambda \mathsf{a_D} \min\{\kappa,1/2\}\lfloor t \rfloor} \overline{\mathsf{C}}_{\mathbf{A}1,p} \mathbb{E}[|\theta_0|^{2\lceil p \rceil(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}1,p} \right),$$

where the last inequality holds due to Lemma 6.5 and where

$$\begin{aligned} \overline{\mathsf{C}}_{\mathbf{A}1,p} &:= 2^{2p+\lceil p \rceil(\rho+1)-1} K_h^{2p}, \\ \widetilde{\mathsf{C}}_{\mathbf{A}1,p} &:= \overline{\mathsf{C}}_{\mathbf{A}1,p} \left(\mathsf{c}_{\lceil p \rceil(\rho+1)} \left(1 + 1/(\mathsf{a_D}\kappa)\right) + 1 \right) + 2^{2p} K_h^{2p} 3 \mathrm{v}_{2\lceil p \rceil(\rho+1)}(\mathsf{M}_V(2\lceil p \rceil(\rho+1))) \\ &\quad + 2^{2p}(2\beta^{-1}(p+1)(d+2p))^p. \end{aligned} \tag{86}$$

This completes the proof. □

Lemma A.3. *Let Assumptions 1, 2, and 3 hold. Then, for any* $0 < \lambda \leq \lambda_{\max}$, $n \in \mathbb{N}_0$, $, t \geq nT$, *we obtain the following inequalities:*

$$\mathbb{E}\left[\left| -\lambda \int_{\lfloor t \rfloor}^t \left(H(\overline{\theta}_s^\lambda) - H_\lambda(\overline{\theta}_{\lfloor s \rfloor}^\lambda) \right) h_\lambda(\overline{\theta}_{\lfloor s \rfloor}^\lambda) \,\mathrm{d}s \right|^2 \right] \leq \lambda^3 \left(e^{-\mathsf{a_D}\kappa n/2} \overline{\mathsf{C}}_{\mathbf{A}2} \mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2} \right),$$

$$\mathbb{E}\left[\left| \lambda^2 \int_{\lfloor t \rfloor}^t H(\overline{\theta}_s^\lambda) \int_{\lfloor s \rfloor}^s H_\lambda(\overline{\theta}_{\lfloor r \rfloor}^\lambda)) h_\lambda(\overline{\theta}_{\lfloor r \rfloor}^\lambda) \,\mathrm{d}r \,\mathrm{d}s \right|^2 \right] \leq \lambda^3 \left(e^{-\mathsf{a_D}\kappa n/2} \overline{\mathsf{C}}_{\mathbf{A}2} \mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2} \right),$$

$$\mathbb{E}\left[\left| -\lambda^2 \beta^{-1} \int_{\lfloor t \rfloor}^t H(\overline{\theta}_s^\lambda) \int_{\lfloor s \rfloor}^s \Upsilon_\lambda(\overline{\theta}_{\lfloor r \rfloor}^\lambda) \,\mathrm{d}r \,\mathrm{d}s \right|^2 \right] \leq \lambda^3 \left(e^{-\mathsf{a_D}\kappa n/2} \overline{\mathsf{C}}_{\mathbf{A}2} \mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2} \right),$$

$$\mathbb{E}\left[\left| -\lambda \sqrt{2\lambda\beta^{-1}} \int_{\lfloor t \rfloor}^t H(\overline{\theta}_s^\lambda) \int_{\lfloor s \rfloor}^s H_\lambda(\overline{\theta}_{\lfloor r \rfloor}^\lambda) \,\mathrm{d}B_r^\lambda \,\mathrm{d}s \right|^2 \right] \leq \lambda^3 \left(e^{-\mathsf{a_D}\kappa n/2} \overline{\mathsf{C}}_{\mathbf{A}2} \mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2} \right),$$

$$\mathbb{E}\left[\left| \sqrt{2\lambda\beta^{-1}} \int_{\lfloor t \rfloor}^t \left(H(\overline{\theta}_s^\lambda) - H_\lambda(\overline{\theta}_{\lfloor s \rfloor}^\lambda) \right) \,\mathrm{d}B_s^\lambda \right|^2 \right] \leq \lambda^2 \left(e^{-\mathsf{a_D}\kappa n/2} \overline{\mathsf{C}}_{\mathbf{A}2} \mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2} \right),$$

$$\begin{aligned} &\mathbb{E}\left[\left| \sqrt{2\lambda\beta^{-1}} \int_{\lfloor t \rfloor}^t \left(H(\overline{\zeta}_s^{\lambda,n}) - H(\overline{\zeta}_{\lfloor s \rfloor}^{\lambda,n}) \right) \,\mathrm{d}B_s^\lambda \right|^2 \right] \\ &\qquad \leq \lambda^2 \left(e^{-\mathsf{a_D}\min\{\kappa,1/2\} n/2} \overline{\mathsf{C}}_{\mathbf{A}2} \mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2} \right), \end{aligned}$$

$$\mathbb{E}\left[\left| \lambda\beta^{-1} \int_{\lfloor t \rfloor}^t \left(\Upsilon(\overline{\theta}_s^\lambda) - \Upsilon_\lambda(\overline{\theta}_{\lfloor s \rfloor}^\lambda) \right) \,\mathrm{d}s \right|^2 \right] \leq \lambda^{2+q} \left(e^{-\mathsf{a_D}\kappa n/2} \overline{\mathsf{C}}_{\mathbf{A}2} \mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2} \right),$$

$$\begin{aligned} &\mathbb{E}\left[\left| -\lambda \int_{\lfloor t \rfloor}^t H(\overline{\theta}_s^\lambda) h_\lambda(\overline{\theta}_{\lfloor s \rfloor}^\lambda) \,\mathrm{d}s \right|^2 + \left| \lambda\beta^{-1} \int_{\lfloor t \rfloor}^t \Upsilon(\overline{\theta}_s^\lambda) \mathrm{d}s \right|^2 \right] \\ &\qquad \leq \lambda^2 \left(e^{-\mathsf{a_D}\kappa n/2} \overline{\mathsf{C}}_{\mathbf{A}2} \mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2} \right), \end{aligned}$$

$$\begin{aligned} &\mathbb{E}\left[\left| -\lambda \int_{\lfloor t \rfloor}^t H(\overline{\zeta}_s^{\lambda,n}) h(\overline{\zeta}_s^{\lambda,n}) \,\mathrm{d}s \right|^2 + \left| \lambda\beta^{-1} \int_{\lfloor t \rfloor}^t \Upsilon(\overline{\zeta}_s^{\lambda,n}) \mathrm{d}s \right|^2 \right] \\ &\qquad \leq \lambda^2 \left(e^{-\mathsf{a_D}\min\{\kappa,1/2\} n/2} \overline{\mathsf{C}}_{\mathbf{A}2} \mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2} \right), \end{aligned}$$

where $\overline{\mathsf{C}}_{\mathbf{A}2}$ *and* $\widetilde{\mathsf{C}}_{\mathbf{A}2}$ *are given in* (88).

Proof. We note that, by using (4), for any $f : \mathbb{R}^d \to \mathbb{R}^{i \times j}$, $i.j \in \mathbb{N}$, $\theta \in \mathbb{R}^d$,

$$|f(\theta) - f_\lambda(\theta)| = \left| \frac{\left((1 + \lambda^{3/2}|\theta|^{3(\rho+q-1)})^{1/3} - 1 \right) f(\theta)}{(1 + \lambda^{3/2}|\theta|^{3(\rho+q-1)})^{1/3}} \right| \leq \lambda^{3/2}|\theta|^{3(\rho+q-1)}|f(\theta)|. \tag{87}$$

The inequalities can be obtained by using the following arguments:

(i) To show that the first inequality holds, by using Remark 2.2 and (87), we have that

$$\begin{aligned}
&\mathbb{E}\left[\left|-\lambda\int_{\lfloor t\rfloor}^{t}\left(H(\overline{\theta}^{\lambda}_{s})-H_{\lambda}(\overline{\theta}^{\lambda}_{\lfloor s\rfloor})\right)h_{\lambda}(\overline{\theta}^{\lambda}_{\lfloor s\rfloor})\,\mathrm{d}s\right|^{2}\right]\\
&\leq\lambda^{2}\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[|H(\overline{\theta}^{\lambda}_{s})-H_{\lambda}(\overline{\theta}^{\lambda}_{\lfloor s\rfloor})|^{2}|h_{\lambda}(\overline{\theta}^{\lambda}_{\lfloor s\rfloor})|^{2}\right]\,\mathrm{d}s\\
&\leq2\lambda^{2}\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[|H(\overline{\theta}^{\lambda}_{s})-H(\overline{\theta}^{\lambda}_{\lfloor s\rfloor})|_{\mathsf{F}}^{2}|h_{\lambda}(\overline{\theta}^{\lambda}_{\lfloor s\rfloor})|^{2}\right]\,\mathrm{d}s\\
&\quad+2\lambda^{2}\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[|H(\overline{\theta}^{\lambda}_{\lfloor s\rfloor})-H_{\lambda}(\overline{\theta}^{\lambda}_{\lfloor s\rfloor})|^{2}|h_{\lambda}(\overline{\theta}^{\lambda}_{\lfloor s\rfloor})|^{2}\right]\,\mathrm{d}s\\
&\leq2\lambda^{2}\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[d\mathsf{K}_{0}^{2}(1+|\overline{\theta}^{\lambda}_{s}|+|\overline{\theta}^{\lambda}_{\lfloor s\rfloor}|)^{2(\rho+q-2)}|\overline{\theta}^{\lambda}_{s}-\overline{\theta}^{\lambda}_{\lfloor s\rfloor}|^{2}K_{h}^{2}(1+|\overline{\theta}^{\lambda}_{\lfloor s\rfloor}|^{\rho+q})^{2}\right]\,\mathrm{d}s\\
&\quad+2\lambda^{2}\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[\lambda^{3}|\overline{\theta}^{\lambda}_{\lfloor s\rfloor}|^{6(\rho+q-1)}K_{H}^{2}(1+|\overline{\theta}^{\lambda}_{\lfloor s\rfloor}|^{\rho+q-1})^{2}K_{h}^{2}(1+|\overline{\theta}^{\lambda}_{\lfloor s\rfloor}|^{\rho+q})^{2}\right]\,\mathrm{d}s\\
&\leq2\lambda^{2}d\mathsf{K}_{0}^{2}K_{h}^{2}\int_{\lfloor t\rfloor}^{t}3^{8(\rho+1)}\left(\mathbb{E}\left[1+|\overline{\theta}^{\lambda}_{s}|^{16(\rho+1)}+|\overline{\theta}^{\lambda}_{\lfloor s\rfloor}|^{16(\rho+1)}\right]\right)^{1/2}\\
&\qquad\times\left(\mathbb{E}\left[|\overline{\theta}^{\lambda}_{s}-\overline{\theta}^{\lambda}_{\lfloor s\rfloor}|^{4}\right]\right)^{1/2}\,\mathrm{d}s\\
&\quad+\lambda^{5}K_{H}^{2}K_{h}^{2}\int_{\lfloor t\rfloor}^{t}2^{16(\rho+1)}\mathbb{E}\left[1+|\overline{\theta}^{\lambda}_{\lfloor s\rfloor}|^{16(\rho+1)}\right]\,\mathrm{d}s\\
&\leq\lambda^{2}3^{8(\rho+1)+1}d\mathsf{K}_{0}^{2}K_{h}^{2}\left(2e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t\rfloor}\mathbb{E}\left[|\theta_{0}|^{16(\rho+1)}\right]+2\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_{\mathsf{D}}\kappa)\right)+1\right)^{1/2}\\
&\qquad\times\lambda\left(e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t\rfloor}\overline{\mathsf{C}}_{\mathbf{A}0,2}\mathbb{E}[|\theta_{0}|^{8(\rho+1)}]+\widetilde{\mathsf{C}}_{\mathbf{A}0,2}\right)^{1/2}\\
&\quad+\lambda^{5}2^{16(\rho+1)}K_{H}^{2}K_{h}^{2}\left(e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t\rfloor}\mathbb{E}\left[|\theta_{0}|^{16(\rho+1)}\right]+\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_{\mathsf{D}}\kappa)\right)+1\right)\\
&\leq\lambda^{3}3^{8(\rho+1)+1}d\mathsf{K}_{0}^{2}K_{h}^{2}\left(e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t\rfloor}(2+\overline{\mathsf{C}}_{\mathbf{A}0,2})\mathbb{E}\left[|\theta_{0}|^{16(\rho+1)}\right]\right.\\
&\qquad\left.+2\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_{\mathsf{D}}\kappa)\right)+2\widetilde{\mathsf{C}}_{\mathbf{A}0,2}+1\right)\\
&\quad+\lambda^{3}2^{16(\rho+1)}K_{H}^{2}K_{h}^{2}\left(e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t\rfloor}\mathbb{E}\left[|\theta_{0}|^{16(\rho+1)}\right]+\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_{\mathsf{D}}\kappa)\right)+1\right)\\
&\leq\lambda^{3}\left(e^{-\mathsf{a}_{\mathsf{D}}\kappa n/2}\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}\left[|\theta_{0}|^{16(\rho+1)}\right]+\widetilde{\mathsf{C}}_{\mathbf{A}2}\right),
\end{aligned}$$

where the fifth inequality holds by applying Lemma 6.3 and A.2, the second last inequality holds by using $\overline{\mathsf{C}}_{\mathbf{A}0,2}<\widetilde{\mathsf{C}}_{\mathbf{A}0,2}$, the last inequality holds due to $\lambda\lfloor t\rfloor\geq\lambda nT\geq n/2$, and where κ is given in (73),

$$\begin{aligned}
\overline{\mathsf{C}}_{\mathbf{A}2}&:=3^{24(\rho+1)}d^{3}(1+\mathsf{K}_{0})^{2}(1+K_{h})^{2}(1+K_{H})^{4}(1+\beta^{-1})^{2}(1+L)^{2}(1+\mathsf{K}_{3,d})^{2}\\
&\qquad\times\left(2+\max\{\overline{\mathsf{C}}_{\mathbf{A}0,2},\overline{\mathsf{C}}_{\mathbf{A}0,2q},\overline{\mathsf{C}}_{\mathbf{A}1,2},\overline{\mathsf{C}}_{\mathbf{A}0,2+2q}\}\right)\\
\widetilde{\mathsf{C}}_{\mathbf{A}2}&:=3^{24(\rho+1)}d^{3}(1+\mathsf{K}_{0})^{2}(1+K_{h})^{2}(1+K_{H})^{4}(1+\beta^{-1})^{2}(1+L)^{2}(1+\mathsf{K}_{3,d})^{2}\\
&\qquad\times\left(2\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_{\mathsf{D}}\kappa)\right)+2\max\{\widetilde{\mathsf{C}}_{\mathbf{A}0,2},\widetilde{\mathsf{C}}_{\mathbf{A}0,2q},\widetilde{\mathsf{C}}_{\mathbf{A}1,2},\widetilde{\mathsf{C}}_{\mathbf{A}0,2+2q}\}+1\right.\\
&\qquad\left.+\,\mathrm{v}_{16(\rho+1)}(\mathsf{M}_{V}(16(\rho+1)))\right)
\end{aligned}\tag{88}$$

with $\overline{\mathsf{C}}_{\mathbf{A}0,p}$, $\widetilde{\mathsf{C}}_{\mathbf{A}0,p}$, $\overline{\mathsf{C}}_{\mathbf{A}1,p}$, $\widetilde{\mathsf{C}}_{\mathbf{A}1,p}$, $p>0$, given in (85) and (86), and c_{p}, $\mathsf{M}_{V}(p)$, $p\in[2,\infty)\cap\mathbb{N}$ given in (81) (see also Lemma 6.3) and Lemma 6.4.

(ii) To establish the second inequality, we apply Lemma 6.3 to obtain

$$\begin{aligned}
&\mathbb{E}\left[\left|\lambda^2\int_{\lfloor t\rfloor}^t H(\overline{\theta}_s^\lambda)\int_{\lfloor s\rfloor}^s H_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda))h_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}r\,\mathrm{d}s\right|^2\right]\\
&\leq \lambda^4\int_{\lfloor t\rfloor}^t \mathbb{E}\left[|H(\overline{\theta}_s^\lambda)|^2|H_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda))|^2|h_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda)|^2\right]\,\mathrm{d}s\\
&\leq \lambda^4\int_{\lfloor t\rfloor}^t \mathbb{E}\left[K_H^4(1+|\overline{\theta}_s^\lambda|^{\rho+q-1})^2(1+|\overline{\theta}_{\lfloor s\rfloor}^\lambda|^{\rho+q-1})^2K_h^2(1+|\overline{\theta}_{\lfloor s\rfloor}^\lambda|^{\rho+q})^2\right]\,\mathrm{d}s\\
&\leq \lambda^4K_H^4K_h^2\int_{\lfloor t\rfloor}^t 3^{16(\rho+1)}\mathbb{E}\left[1+|\overline{\theta}_s^\lambda|^{16(\rho+1)}+|\overline{\theta}_{\lfloor s\rfloor}^\lambda|^{16(\rho+1)}\right]\,\mathrm{d}s\\
&\leq \lambda^3 3^{16(\rho+1)}K_H^4K_h^2\left(2e^{-\lambda\mathsf{a}_\mathsf{D}\kappa\lfloor t\rfloor}\mathbb{E}\left[|\theta_0|^{16(\rho+1)}\right]+2\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_\mathsf{D}\kappa)\right)+1\right)\\
&\leq \lambda^3\left(e^{-\mathsf{a}_\mathsf{D}\kappa n/2}\overline{\mathsf{C}}_{\mathbf{A2}}\mathbb{E}[|\theta_0|^{16(\rho+1)}]+\widetilde{\mathsf{C}}_{\mathbf{A2}}\right),
\end{aligned}$$

where $\overline{\mathsf{C}}_{\mathbf{A2}}$ and $\widetilde{\mathsf{C}}_{\mathbf{A2}}$ are given in (88).

(iii) To obtain the third inequality, we apply Lemma 6.3 and write

$$\begin{aligned}
&\mathbb{E}\left[\left|-\lambda^2\beta^{-1}\int_{\lfloor t\rfloor}^t H(\overline{\theta}_s^\lambda)\int_{\lfloor s\rfloor}^s \Upsilon_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}r\,\mathrm{d}s\right|^2\right]\\
&\leq \lambda^4\beta^{-2}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[|H(\overline{\theta}_s^\lambda)|^2|\Upsilon_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda)|^2\right]\,\mathrm{d}s\\
&\leq \lambda^4\beta^{-2}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[K_H^2(1+|\overline{\theta}_s^\lambda|^{\rho+q-1})^2\mathsf{K}_{3,d}^2(1+|\overline{\theta}_{\lfloor s\rfloor}^\lambda|)^{2(\rho+q-2)}\right]\,\mathrm{d}s\\
&\leq \lambda^4\beta^{-2}K_H^2\mathsf{K}_{3,d}^2\int_{\lfloor t\rfloor}^t 3^{16(\rho+1)}\mathbb{E}\left[1+|\overline{\theta}_s^\lambda|^{16(\rho+1)}+|\overline{\theta}_{\lfloor s\rfloor}^\lambda|^{16(\rho+1)}\right]\,\mathrm{d}s\\
&\leq \lambda^3 3^{16(\rho+1)}\beta^{-2}K_H^2\mathsf{K}_{3,d}^2\left(2e^{-\lambda\mathsf{a}_\mathsf{D}\kappa\lfloor t\rfloor}\mathbb{E}\left[|\theta_0|^{16(\rho+1)}\right]+2\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_\mathsf{D}\kappa)\right)+1\right)\\
&\leq \lambda^3\left(e^{-\mathsf{a}_\mathsf{D}\kappa n/2}\overline{\mathsf{C}}_{\mathbf{A2}}\mathbb{E}[|\theta_0|^{16(\rho+1)}]+\widetilde{\mathsf{C}}_{\mathbf{A2}}\right),
\end{aligned}$$

where $\overline{\mathsf{C}}_{\mathbf{A2}}$ and $\widetilde{\mathsf{C}}_{\mathbf{A2}}$ are given in (88).

(iv) To obtain the fourth inequality, we use Cauchy-Schwarz inequality and Lemma 6.3:

$$\begin{aligned}
&\mathbb{E}\left[\left|-\lambda\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t H(\overline{\theta}_s^\lambda)\int_{\lfloor s\rfloor}^s H_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\,\mathrm{d}s\right|^2\right]\\
&\leq 2\lambda^3\beta^{-1}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[|H(\overline{\theta}_s^\lambda)|^2|H_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda)|^2\left|\int_{\lfloor s\rfloor}^s \mathrm{d}B_r^\lambda\right|^2\right]\,\mathrm{d}s\\
&\leq 2\lambda^3\beta^{-1}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[K_H^4(1+|\overline{\theta}_s^\lambda|^{\rho+q-1})^2(1+|\overline{\theta}_{\lfloor s\rfloor}^\lambda|^{\rho+q-1})^2\left|\int_{\lfloor s\rfloor}^s \mathrm{d}B_r^\lambda\right|^2\right]\,\mathrm{d}s\\
&\leq 2\lambda^3\beta^{-1}K_H^4\int_{\lfloor t\rfloor}^t 3^{8(\rho+1)}\left(\mathbb{E}\left[1+|\overline{\theta}_s^\lambda|^{16(\rho+1)}+|\overline{\theta}_{\lfloor s\rfloor}^\lambda|^{16(\rho+1)}\right]\right)^{1/2}\\
&\qquad\times\left(\mathbb{E}\left[\left|\int_{\lfloor s\rfloor}^s \mathrm{d}B_r^\lambda\right|^4\right]\right)^{1/2}\,\mathrm{d}s\\
&\leq \lambda^3\beta^{-1}K_H^4 3^{8(\rho+1)+2}(d+4)\left(2e^{-\lambda\mathsf{a}_\mathsf{D}\kappa\lfloor t\rfloor}\mathbb{E}\left[|\theta_0|^{16(\rho+1)}\right]+2\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_\mathsf{D}\kappa)\right)+1\right)\\
&\leq \lambda^3\left(e^{-\mathsf{a}_\mathsf{D}\kappa n/2}\overline{\mathsf{C}}_{\mathbf{A2}}\mathbb{E}[|\theta_0|^{16(\rho+1)}]+\widetilde{\mathsf{C}}_{\mathbf{A2}}\right),
\end{aligned}$$

where $\overline{\mathsf{C}}_{\mathbf{A}2}$ and $\widetilde{\mathsf{C}}_{\mathbf{A}2}$ are given in (88).

(v) To obtain the fifth inequality, we apply Remark 2.2, (87), and Cauchy-Schwarz inequality:

$$
\begin{aligned}
&\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^{t}\left(H(\overline{\theta}_s^{\lambda})-H_\lambda(\overline{\theta}_{\lfloor s\rfloor}^{\lambda})\right)\,\mathrm{d}B_s^{\lambda}\right|^2\right] \\
&= 2\lambda\beta^{-1}\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[|H(\overline{\theta}_s^{\lambda})-H_\lambda(\overline{\theta}_{\lfloor s\rfloor}^{\lambda})|_{\mathsf{F}}^2\right]\,\mathrm{d}s \\
&\leq 4\lambda\beta^{-1}\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[|H(\overline{\theta}_s^{\lambda})-H(\overline{\theta}_{\lfloor s\rfloor}^{\lambda})|_{\mathsf{F}}^2\right]\,\mathrm{d}s+4\lambda\beta^{-1}\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[|H(\overline{\theta}_{\lfloor s\rfloor}^{\lambda})-H_\lambda(\overline{\theta}_{\lfloor s\rfloor}^{\lambda})|_{\mathsf{F}}^2\right]\,\mathrm{d}s \\
&\leq 4\lambda\beta^{-1}\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[d\mathsf{K}_0^2(1+|\overline{\theta}_s^{\lambda}|+|\overline{\theta}_{\lfloor s\rfloor}^{\lambda}|)^{2(\rho+q-2)}|\overline{\theta}_s^{\lambda}-\overline{\theta}_{\lfloor s\rfloor}^{\lambda}|^2\right]\,\mathrm{d}s \\
&\quad+4\lambda\beta^{-1}\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[\lambda^3|\overline{\theta}_{\lfloor s\rfloor}^{\lambda}|^{6(\rho+q-1)}dK_H^2(1+|\overline{\theta}_{\lfloor s\rfloor}^{\lambda}|^{\rho+q-1})^2\right]\,\mathrm{d}s \\
&\leq 4\lambda\beta^{-1}d\mathsf{K}_0^2\int_{\lfloor t\rfloor}^{t}3^{8(\rho+1)}\left(\mathbb{E}\left[1+|\overline{\theta}_s^{\lambda}|^{16(\rho+1)}+|\overline{\theta}_{\lfloor s\rfloor}^{\lambda}|^{16(\rho+1)}\right]\right)^{1/2} \\
&\qquad\times\left(\mathbb{E}\left[|\overline{\theta}_s^{\lambda}-\overline{\theta}_{\lfloor s\rfloor}^{\lambda}|^4\right]\right)^{1/2}\,\mathrm{d}s \\
&\quad+4\lambda^4\beta^{-1}dK_H^2\int_{\lfloor t\rfloor}^{t}2^{16(\rho+1)}\mathbb{E}\left[1+|\overline{\theta}_{\lfloor s\rfloor}^{\lambda}|^{16(\rho+1)}\right]\,\mathrm{d}s \\
&\leq \lambda\beta^{-1}d\mathsf{K}_0^2 3^{8(\rho+1)+2}\left(2e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t\rfloor}\mathbb{E}\left[|\theta_0|^{16(\rho+1)}\right]+2\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_{\mathsf{D}}\kappa)\right)+1\right)^{1/2} \\
&\qquad\times\lambda\left(e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t\rfloor}\overline{\mathsf{C}}_{\mathbf{A}0,2}\mathbb{E}[|\theta_0|^{8(\rho+1)}]+\widetilde{\mathsf{C}}_{\mathbf{A}0,2}\right)^{1/2} \\
&\quad+\lambda^4\beta^{-1}dK_H^2 2^{16(\rho+1)+2}\left(e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t\rfloor}\mathbb{E}\left[|\theta_0|^{16(\rho+1)}\right]+\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_{\mathsf{D}}\kappa)\right)+1\right) \\
&\leq \lambda^2\beta^{-1}d\mathsf{K}_0^2 3^{8(\rho+1)+2}\left(e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t\rfloor}(2+\overline{\mathsf{C}}_{\mathbf{A}0,2})\mathbb{E}\left[|\theta_0|^{16(\rho+1)}\right]\right. \\
&\qquad\left.+2\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_{\mathsf{D}}\kappa)\right)+2\widetilde{\mathsf{C}}_{\mathbf{A}0,2}+1\right) \\
&\quad+\lambda^2\beta^{-1}dK_H^2 2^{16(\rho+1)+2}\left(e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t\rfloor}\mathbb{E}\left[|\theta_0|^{16(\rho+1)}\right]+\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_{\mathsf{D}}\kappa)\right)+1\right) \\
&\leq \lambda^2\left(e^{-\mathsf{a}_{\mathsf{D}}\kappa n/2}\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}]+\widetilde{\mathsf{C}}_{\mathbf{A}2}\right),
\end{aligned}
$$

where the fourth inequality holds by applying Lemma 6.3 and A.2, and where $\overline{\mathsf{C}}_{\mathbf{A}2}$ and $\widetilde{\mathsf{C}}_{\mathbf{A}2}$ are given in (88).

(vi) To obtain the sixth inequality, we use the same arguments as in (v):

$$
\begin{aligned}
&\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^{t}\left(H(\overline{\zeta}_s^{\lambda,n})-H(\overline{\zeta}_{\lfloor s\rfloor}^{\lambda,n})\right)\,\mathrm{d}B_s^{\lambda}\right|^2\right] \\
&= 2\lambda\beta^{-1}\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[|H(\overline{\zeta}_s^{\lambda,n})-H(\overline{\zeta}_{\lfloor s\rfloor}^{\lambda,n})|_{\mathsf{F}}^2\right]\,\mathrm{d}s \\
&\leq 2\lambda\beta^{-1}\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[d\mathsf{K}_0^2(1+|\overline{\zeta}_s^{\lambda,n}|+|\overline{\zeta}_{\lfloor s\rfloor}^{\lambda,n}|)^{2(\rho+q-2)}|\overline{\zeta}_s^{\lambda,n}-\overline{\zeta}_{\lfloor s\rfloor}^{\lambda,n}|^2\right]\,\mathrm{d}s \\
&\leq 2\lambda\beta^{-1}d\mathsf{K}_0^2\int_{\lfloor t\rfloor}^{t}3^{8(\rho+1)}\left(\mathbb{E}\left[1+|\overline{\zeta}_s^{\lambda,n}|^{16(\rho+1)}+|\overline{\zeta}_{\lfloor s\rfloor}^{\lambda,n}|^{16(\rho+1)}\right]\right)^{1/2} \\
&\qquad\times\left(\mathbb{E}\left[|\overline{\zeta}_s^{\lambda,n}-\overline{\zeta}_{\lfloor s\rfloor}^{\lambda,n}|^4\right]\right)^{1/2}\,\mathrm{d}s \\
&\leq \lambda\beta^{-1}d\mathsf{K}_0^2 3^{8(\rho+1)+1}\left(2^{8(\rho+1)}e^{-\lambda\mathsf{a}_{\mathsf{D}}\min\{\kappa,1/2\}\lfloor t\rfloor}\mathbb{E}[|\theta_0|^{16(\rho+1)}]\right.
\end{aligned}
$$

$$
\begin{aligned}
&\quad +2^{8(\rho+1)}\left(\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_{\mathsf{D}}\kappa)\right)+1\right)+6\mathrm{v}_{16(\rho+1)}(\mathsf{M}_V(16(\rho+1)))\Big)^{1/2}\\
&\quad \times\lambda\left(e^{-\lambda\mathsf{a}_{\mathsf{D}}\min\{\kappa,1/2\}\lfloor t\rfloor}\overline{\mathsf{C}}_{\mathbf{A}1,2}\mathbb{E}[|\theta_0|^{4(\rho+1)}]+\widetilde{\mathsf{C}}_{\mathbf{A}1,2}\right)^{1/2}\\
&\le \lambda^2\beta^{-1}d\mathsf{K}_0^2 3^{16(\rho+1)}\Big(e^{-\lambda\mathsf{a}_{\mathsf{D}}\min\{\kappa,1/2\}\lfloor t\rfloor}(1+\overline{\mathsf{C}}_{\mathbf{A}1,2})\mathbb{E}[|\theta_0|^{16(\rho+1)}]\\
&\quad +\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_{\mathsf{D}}\kappa)\right)+1+2\widetilde{\mathsf{C}}_{\mathbf{A}1,2}+\mathrm{v}_{16(\rho+1)}(\mathsf{M}_V(16(\rho+1)))\Big)\\
&\le \lambda^2\left(e^{-\mathsf{a}_{\mathsf{D}}\min\{\kappa,1/2\}n/2}\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}]+\widetilde{\mathsf{C}}_{\mathbf{A}2}\right),
\end{aligned}
$$

where the third inequality holds due to Lemma 6.5 and A.2, the fourth inequality holds due to $\overline{\mathsf{C}}_{\mathbf{A}1,2}<\widetilde{\mathsf{C}}_{\mathbf{A}1,2}$, and where $\overline{\mathsf{C}}_{\mathbf{A}2}$ and $\widetilde{\mathsf{C}}_{\mathbf{A}2}$ are given in (88).

(vii) To establish the seventh inequality, we apply Remark 2.2, (87), and Cauchy-Schwarz inequality to obtain

$$
\begin{aligned}
&\mathbb{E}\left[\left|\lambda\beta^{-1}\int_{\lfloor t\rfloor}^t\left(\Upsilon(\overline{\theta}_s^\lambda)-\Upsilon_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda)\right)\mathrm{d}s\right|^2\right]\\
&\le \lambda^2\beta^{-2}\int_{\lfloor t\rfloor}^t\mathbb{E}\left[|\Upsilon(\overline{\theta}_s^\lambda)-\Upsilon_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda)|^2\right]\mathrm{d}s\\
&\le 2\lambda^2\beta^{-2}\int_{\lfloor t\rfloor}^t\mathbb{E}\left[|\Upsilon(\overline{\theta}_s^\lambda)-\Upsilon(\overline{\theta}_{\lfloor s\rfloor}^\lambda)|^2\right]\mathrm{d}s+2\lambda^2\beta^{-2}\int_{\lfloor t\rfloor}^t\mathbb{E}\left[|\Upsilon(\overline{\theta}_{\lfloor s\rfloor}^\lambda)-\Upsilon_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda)|^2\right]\mathrm{d}s\\
&\le 2\lambda^2\beta^{-2}\int_{\lfloor t\rfloor}^t\mathbb{E}\left[d^3L^2(1+|\overline{\theta}_s^\lambda|+|\overline{\theta}_{\lfloor s\rfloor}^\lambda|)^{2(\rho-2)}|\overline{\theta}_s^\lambda-\overline{\theta}_{\lfloor s\rfloor}^\lambda|^{2q}\right]\mathrm{d}s\\
&\quad +2\lambda^2\beta^{-2}\int_{\lfloor t\rfloor}^t\mathbb{E}\left[\lambda^3|\overline{\theta}_{\lfloor s\rfloor}^\lambda|^{6(\rho+q-1)}\mathsf{K}_{3,d}^2(1+|\overline{\theta}_{\lfloor s\rfloor}^\lambda|)^{2(\rho+q-2)}\right]\mathrm{d}s\\
&\le 2\lambda^2\beta^{-2}d^3L^2\int_{\lfloor t\rfloor}^t 3^{8(\rho+1)}\left(\mathbb{E}\left[1+|\overline{\theta}_s^\lambda|^{16(\rho+1)}+|\overline{\theta}_{\lfloor s\rfloor}^\lambda|^{16(\rho+1)}\right]\right)^{1/2}\\
&\quad\times\left(\mathbb{E}\left[|\overline{\theta}_s^\lambda-\overline{\theta}_{\lfloor s\rfloor}^\lambda|^{4q}\right]\right)^{1/2}\mathrm{d}s\\
&\quad +2\lambda^5\beta^{-2}\mathsf{K}_{3,d}^2\int_{\lfloor t\rfloor}^t 2^{16(\rho+1)}\mathbb{E}\left[1+|\overline{\theta}_{\lfloor s\rfloor}^\lambda|^{16(\rho+1)}\right]\mathrm{d}s\\
&\le \lambda^2\beta^{-2}d^3L^2 3^{8(\rho+1)+1}\left(2e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t\rfloor}\mathbb{E}\left[|\theta_0|^{16(\rho+1)}\right]+2\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_{\mathsf{D}}\kappa)\right)+1\right)^{1/2}\\
&\quad\times\lambda^q\left(e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t\rfloor}\overline{\mathsf{C}}_{\mathbf{A}0,2q}\mathbb{E}[|\theta_0|^{4\lceil 2q\rceil(\rho+1)}]+\widetilde{\mathsf{C}}_{\mathbf{A}0,2q}\right)^{1/2}\\
&\quad +\lambda^5\beta^{-2}\mathsf{K}_{3,d}^2 2^{16(\rho+1)+1}\left(e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t\rfloor}\mathbb{E}\left[|\theta_0|^{16(\rho+1)}\right]+\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_{\mathsf{D}}\kappa)\right)+1\right)\\
&\le \lambda^{2+q}\beta^{-2}d^3L^2 3^{8(\rho+1)+1}\\
&\quad\times\left(e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t\rfloor}(2+\overline{\mathsf{C}}_{\mathbf{A}0,2q})\mathbb{E}\left[|\theta_0|^{16(\rho+1)}\right]+2\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_{\mathsf{D}}\kappa)\right)+2\widetilde{\mathsf{C}}_{\mathbf{A}0,2q}+1\right)\\
&\quad +\lambda^{2+q}\beta^{-2}\mathsf{K}_{3,d}^2 2^{16(\rho+1)+1}\left(e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t\rfloor}\mathbb{E}\left[|\theta_0|^{16(\rho+1)}\right]+\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_{\mathsf{D}}\kappa)\right)+1\right)\\
&\le \lambda^{2+q}\left(e^{-\mathsf{a}_{\mathsf{D}}\kappa n/2}\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}]+\widetilde{\mathsf{C}}_{\mathbf{A}2}\right),
\end{aligned}
$$

where the fourth inequality holds due to Lemma 6.3 and A.2, and $\overline{\mathsf{C}}_{\mathbf{A}2}$, $\widetilde{\mathsf{C}}_{\mathbf{A}2}$ are given in (88).

(viii) To obtain the eighth inequality, we apply Remark 2.2 and write

$$
\begin{aligned}
&\mathbb{E}\left[\left|-\lambda\int_{\lfloor t\rfloor}^t H(\overline{\theta}_s^\lambda)h_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda)\,\mathrm{d}s\right|^2+\left|\lambda\beta^{-1}\int_{\lfloor t\rfloor}^t\Upsilon(\overline{\theta}_s^\lambda)\,\mathrm{d}s\right|^2\right]\\
&\le\lambda^2\int_{\lfloor t\rfloor}^t\mathbb{E}\left[|H(\overline{\theta}_s^\lambda)|^2|h(\overline{\theta}_{\lfloor s\rfloor}^\lambda)|^2\right]\mathrm{d}s+\lambda^2\beta^{-2}\int_{\lfloor t\rfloor}^t\mathbb{E}\left[|\Upsilon(\overline{\theta}_s^\lambda)|^2\right]\mathrm{d}s
\end{aligned}
$$

$$
\begin{aligned}
&\leq \lambda^2 \int_{\lfloor t\rfloor}^t \mathbb{E}\left[K_H^2(1+|\overline{\theta}_s^\lambda|^{\rho+q-1})^2 K_h^2(1+|\overline{\theta}_{\lfloor s\rfloor}^\lambda|^{\rho+q})^2\right] \mathrm{d}s \\
&\quad + \lambda^2\beta^{-2} \int_{\lfloor t\rfloor}^t \mathbb{E}\left[\mathsf{K}_{3,d}^2(1+|\overline{\theta}_s^\lambda|)^{2(\rho+q-2)}\right] \mathrm{d}s \\
&\leq \lambda^2(1+K_H)^2(1+K_h)^2(1+\beta^{-1})^2(1+\mathsf{K}_{3,d})^2 3^{16(\rho+1)} \\
&\quad\quad \times \int_{\lfloor t\rfloor}^t \mathbb{E}\left[1+|\overline{\theta}_s^\lambda|^{16(\rho+1)}+|\overline{\theta}_{\lfloor s\rfloor}^\lambda|^{16(\rho+1)}\right] \mathrm{d}s \\
&\leq \lambda^2(1+K_H)^2(1+K_h)^2(1+\beta^{-1})^2(1+\mathsf{K}_{3,d})^2 3^{16(\rho+1)} \\
&\quad\quad \times \left(2e^{-\lambda \mathsf{a}_\mathsf{D}\kappa\lfloor t\rfloor}\mathbb{E}\left[|\theta_0|^{16(\rho+1)}\right]+2\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_\mathsf{D}\kappa)\right)+1\right) \\
&\leq \lambda^2\left(e^{-\mathsf{a}_\mathsf{D}\kappa n/2}\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}]+\widetilde{\mathsf{C}}_{\mathbf{A}2}\right),
\end{aligned}
$$

where the fourth inequality holds due to Lemma 6.3, and where $\overline{\mathsf{C}}_{\mathbf{A}2}$, $\widetilde{\mathsf{C}}_{\mathbf{A}2}$ are given in (88).

(ix) To establish the last inequality, we follow the arguments in (viii):

$$
\begin{aligned}
&\mathbb{E}\left[\left|-\lambda\int_{\lfloor t\rfloor}^t H(\overline{\zeta}_s^{\lambda,n})h(\overline{\zeta}_s^{\lambda,n})\,\mathrm{d}s\right|^2+\left|\lambda\beta^{-1}\int_{\lfloor t\rfloor}^t \Upsilon(\overline{\zeta}_s^{\lambda,n})\mathrm{d}s\right|^2\right] \\
&\leq \lambda^2(1+K_H)^2(1+K_h)^2(1+\beta^{-1})^2(1+\mathsf{K}_{3,d})^2 3^{16(\rho+1)} \\
&\quad \times \int_{\lfloor t\rfloor}^t \mathbb{E}\left[1+|\overline{\zeta}_s^{\lambda,n}|^{16(\rho+1)}+|\overline{\zeta}_{\lfloor s\rfloor}^{\lambda,n}|^{16(\rho+1)}\right] \mathrm{d}s \\
&\leq \lambda^2(1+K_H)^2(1+K_h)^2(1+\beta^{-1})^2(1+\mathsf{K}_{3,d})^2 3^{16(\rho+1)} \\
&\quad \times \Big(2^{8(\rho+1)}e^{-\lambda \mathsf{a}_\mathsf{D}\min\{\kappa,1/2\}\lfloor t\rfloor}\mathbb{E}[|\theta_0|^{16(\rho+1)}] \\
&\quad +2^{8(\rho+1)}\left(\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_\mathsf{D}\kappa)\right)+1\right)+6\mathrm{v}_{16(\rho+1)}(\mathsf{M}_V(16(\rho+1)))\Big) \\
&\leq \lambda^2\left(e^{-\mathsf{a}_\mathsf{D}\min\{\kappa,1/2\}n/2}\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}]+\widetilde{\mathsf{C}}_{\mathbf{A}2}\right),
\end{aligned}
$$

where the second inequality holds due to Lemma 6.5, and where $\overline{\mathsf{C}}_{\mathbf{A}2}$ and $\widetilde{\mathsf{C}}_{\mathbf{A}2}$ are given in (88).

This completes the proof. □

Corollary A.4. *Let Assumptions 1, 2, and 3 hold. Then, for any $0<\lambda\leq\lambda_{\max}$, $n\in\mathbb{N}_0$, , $t\geq nT$, we obtain the following inequalities:*

$$
\begin{aligned}
&\mathbb{E}\left[\left|h(\overline{\theta}_t^\lambda)-h(\overline{\theta}_{\lfloor t\rfloor}^\lambda)+\lambda\int_{\lfloor t\rfloor}^t\left(H_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda)h_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda)-\beta^{-1}\Upsilon_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda)\right)\mathrm{d}s\right.\right. \\
&\qquad\left.\left.-\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t H_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda)\,\mathrm{d}B_s^\lambda\right|^2\right]\leq\lambda^2\left(e^{-\mathsf{a}_\mathsf{D}\kappa n/2}36\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}]+36\widetilde{\mathsf{C}}_{\mathbf{A}2}\right), \\
&\mathbb{E}\left[\left|h(\overline{\zeta}_t^{\lambda,n})-h(\overline{\zeta}_{\lfloor t\rfloor}^{\lambda,n})-\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t H(\overline{\zeta}_{\lfloor s\rfloor}^{\lambda,n})\,\mathrm{d}B_s^\lambda-\left(h(\overline{\theta}_t^\lambda)-h(\overline{\theta}_{\lfloor t\rfloor}^\lambda)\right.\right.\right. \\
&\qquad\left.\left.\left.-\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t H(\overline{\theta}_{\lfloor s\rfloor}^\lambda)\,\mathrm{d}B_s^\lambda\right)\right|^2\right]\leq\lambda^2\left(e^{-\mathsf{a}_\mathsf{D}\min\{\kappa,1/2\}n/2}72\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}]+72\widetilde{\mathsf{C}}_{\mathbf{A}2}\right),
\end{aligned}
$$

where $\overline{\mathsf{C}}_{\mathbf{A}2}$ and $\widetilde{\mathsf{C}}_{\mathbf{A}2}$ are given in (88).

Proof. For any $t \geq nT$, by applying Itô's formula to $h(\overline{\theta}_t^\lambda)$, we obtain, almost surely

$$\begin{aligned} h(\overline{\theta}_t^\lambda) - h(\overline{\theta}_{\lfloor t \rfloor}^\lambda) = & -\lambda \int_{\lfloor t \rfloor}^t H(\overline{\theta}_s^\lambda) h_\lambda(\overline{\theta}_{\lfloor s \rfloor}^\lambda)\, \mathrm{d}s + \lambda^2 \int_{\lfloor t \rfloor}^t H(\overline{\theta}_s^\lambda) \int_{\lfloor s \rfloor}^s H_\lambda(\overline{\theta}_{\lfloor r \rfloor}^\lambda) h_\lambda(\overline{\theta}_{\lfloor r \rfloor}^\lambda)\, \mathrm{d}r\, \mathrm{d}s \\ & - \lambda^2 \beta^{-1} \int_{\lfloor t \rfloor}^t H(\overline{\theta}_s^\lambda) \int_{\lfloor s \rfloor}^s \Upsilon_\lambda(\overline{\theta}_{\lfloor r \rfloor}^\lambda)\, \mathrm{d}r\, \mathrm{d}s \\ & - \lambda \sqrt{2\lambda\beta^{-1}} \int_{\lfloor t \rfloor}^t H(\overline{\theta}_s^\lambda) \int_{\lfloor s \rfloor}^s H_\lambda(\overline{\theta}_{\lfloor r \rfloor}^\lambda)\, \mathrm{d}B_r^\lambda\, \mathrm{d}s \\ & + \sqrt{2\lambda\beta^{-1}} \int_{\lfloor t \rfloor}^t H(\overline{\theta}_s^\lambda)\, \mathrm{d}B_s^\lambda + \lambda\beta^{-1} \int_{\lfloor t \rfloor}^t \Upsilon(\overline{\theta}_s^\lambda) \mathrm{d}s. \end{aligned} \tag{89}$$

Similarly, applying Itô's formula to $h(\overline{\zeta}_t^{\lambda,n})$ yields, almost surely

$$\begin{aligned} h(\overline{\zeta}_t^{\lambda,n}) - h(\overline{\zeta}_{\lfloor t \rfloor}^{\lambda,n}) = & -\lambda \int_{\lfloor t \rfloor}^t H(\overline{\zeta}_s^{\lambda,n}) h(\overline{\zeta}_s^{\lambda,n})\, \mathrm{d}s + \sqrt{2\lambda\beta^{-1}} \int_{\lfloor t \rfloor}^t H(\overline{\zeta}_s^{\lambda,n})\, \mathrm{d}B_s^\lambda \\ & + \lambda\beta^{-1} \int_{\lfloor t \rfloor}^t \Upsilon(\overline{\zeta}_s^{\lambda,n}) \mathrm{d}s. \end{aligned} \tag{90}$$

(i) To obtain the first inequality, by using (89) with $(\sum_{l=1}^v u_l)^2 \leq v \sum_{l=1}^v u_l^2$, $v \in \mathbb{N}$, $u_l \geq 0$, we obtain that

$$\begin{aligned} & \mathbb{E}\left[\left| h(\overline{\theta}_t^\lambda) - h(\overline{\theta}_{\lfloor t \rfloor}^\lambda) + \lambda \int_{\lfloor t \rfloor}^t \left(H_\lambda(\overline{\theta}_{\lfloor s \rfloor}^\lambda) h_\lambda(\overline{\theta}_{\lfloor s \rfloor}^\lambda) - \beta^{-1} \Upsilon_\lambda(\overline{\theta}_{\lfloor s \rfloor}^\lambda) \right) \mathrm{d}s \right.\right. \\ & \qquad \left.\left. - \sqrt{2\lambda\beta^{-1}} \int_{\lfloor t \rfloor}^t H_\lambda(\overline{\theta}_{\lfloor s \rfloor}^\lambda)\, \mathrm{d}B_s^\lambda \right|^2 \right] \\ & \leq 6\mathbb{E}\left[\left| -\lambda \int_{\lfloor t \rfloor}^t \left(H(\overline{\theta}_s^\lambda) - H_\lambda(\overline{\theta}_{\lfloor s \rfloor}^\lambda) \right) h_\lambda(\overline{\theta}_{\lfloor s \rfloor}^\lambda)\, \mathrm{d}s \right|^2 \right] \\ & \quad + 6\mathbb{E}\left[\left| \lambda^2 \int_{\lfloor t \rfloor}^t H(\overline{\theta}_s^\lambda) \int_{\lfloor s \rfloor}^s H_\lambda(\overline{\theta}_{\lfloor r \rfloor}^\lambda)) h_\lambda(\overline{\theta}_{\lfloor r \rfloor}^\lambda)\, \mathrm{d}r\, \mathrm{d}s \right|^2 \right] \\ & \quad + 6\mathbb{E}\left[\left| -\lambda^2\beta^{-1} \int_{\lfloor t \rfloor}^t H(\overline{\theta}_s^\lambda) \int_{\lfloor s \rfloor}^s \Upsilon_\lambda(\overline{\theta}_{\lfloor r \rfloor}^\lambda)\, \mathrm{d}r\, \mathrm{d}s \right|^2 \right] \\ & \quad + 6\mathbb{E}\left[\left| -\lambda\sqrt{2\lambda\beta^{-1}} \int_{\lfloor t \rfloor}^t H(\overline{\theta}_s^\lambda) \int_{\lfloor s \rfloor}^s H_\lambda(\overline{\theta}_{\lfloor r \rfloor}^\lambda)\, \mathrm{d}B_r^\lambda\, \mathrm{d}s \right|^2 \right] \\ & \quad + 6\mathbb{E}\left[\left| \sqrt{2\lambda\beta^{-1}} \int_{\lfloor t \rfloor}^t \left(H(\overline{\theta}_s^\lambda) - H_\lambda(\overline{\theta}_{\lfloor s \rfloor}^\lambda) \right) \mathrm{d}B_s^\lambda \right|^2 \right] \\ & \quad + 6\mathbb{E}\left[\left| \lambda\beta^{-1} \int_{\lfloor t \rfloor}^t \left(\Upsilon(\overline{\theta}_s^\lambda) - \Upsilon_\lambda(\overline{\theta}_{\lfloor s \rfloor}^\lambda) \right) \mathrm{d}s \right|^2 \right] \\ & \leq \lambda^2 \left(e^{-\mathsf{a}_\mathsf{D}\kappa n/2} 36 \overline{\mathsf{C}}_{\mathbf{A2}} \mathbb{E}[|\theta_0|^{16(\rho+1)}] + 36 \widetilde{\mathsf{C}}_{\mathbf{A2}} \right), \end{aligned}$$

where the last inequality holds due to Lemma A.3.

(ii) To establish the second inequality, we use (89) and (90) to obtain

$$\mathbb{E}\left[\left| h(\overline{\zeta}_t^{\lambda,n}) - h(\overline{\zeta}_{\lfloor t \rfloor}^{\lambda,n}) - \sqrt{2\lambda\beta^{-1}} \int_{\lfloor t \rfloor}^t H(\overline{\zeta}_{\lfloor s \rfloor}^{\lambda,n})\, \mathrm{d}B_s^\lambda \right.\right.$$

$$
\begin{aligned}
&\quad - \left(h(\overline{\theta}_t^\lambda) - h(\overline{\theta}_{\lfloor t \rfloor}^\lambda) - \sqrt{2\lambda\beta^{-1}} \int_{\lfloor t \rfloor}^t H(\overline{\theta}_{\lfloor s \rfloor}^\lambda) \, \mathrm{d}B_s^\lambda \right) \Bigg|^2 \Bigg] \\
&\leq 2\mathbb{E}\left[\left| h(\overline{\zeta}_t^{\lambda,n}) - h(\overline{\zeta}_{\lfloor t \rfloor}^{\lambda,n}) - \sqrt{2\lambda\beta^{-1}} \int_{\lfloor t \rfloor}^t H(\overline{\zeta}_{\lfloor s \rfloor}^{\lambda,n}) \, \mathrm{d}B_s^\lambda \right|^2 \right] \\
&\quad + 2\mathbb{E}\left[\left| h(\overline{\theta}_t^\lambda) - h(\overline{\theta}_{\lfloor t \rfloor}^\lambda) - \sqrt{2\lambda\beta^{-1}} \int_{\lfloor t \rfloor}^t H(\overline{\theta}_{\lfloor s \rfloor}^\lambda) \, \mathrm{d}B_s^\lambda \right|^2 \right] \\
&\leq 6\mathbb{E}\left[\left| -\lambda \int_{\lfloor t \rfloor}^t H(\overline{\zeta}_s^{\lambda,n}) h(\overline{\zeta}_s^{\lambda,n}) \, \mathrm{d}s \right|^2 \right] + 6\mathbb{E}\left[\left| \lambda\beta^{-1} \int_{\lfloor t \rfloor}^t \Upsilon(\overline{\zeta}_s^{\lambda,n}) \mathrm{d}s \right|^2 \right] \\
&\quad + 6\mathbb{E}\left[\left| \sqrt{2\lambda\beta^{-1}} \int_{\lfloor t \rfloor}^t \left(H(\overline{\zeta}_s^{\lambda,n}) - H(\overline{\zeta}_{\lfloor s \rfloor}^{\lambda,n}) \right) \, \mathrm{d}B_s^\lambda \right|^2 \right] \\
&\quad + 12\mathbb{E}\left[\left| -\lambda \int_{\lfloor t \rfloor}^t H(\overline{\theta}_s^\lambda) h_\lambda(\overline{\theta}_{\lfloor s \rfloor}^\lambda) \, \mathrm{d}s \right|^2 \right] + 12\mathbb{E}\left[\left| \lambda\beta^{-1} \int_{\lfloor t \rfloor}^t \Upsilon(\overline{\theta}_s^\lambda) \mathrm{d}s \right|^2 \right] \\
&\quad + 12\mathbb{E}\left[\left| \lambda^2 \int_{\lfloor t \rfloor}^t H(\overline{\theta}_s^\lambda) \int_{\lfloor s \rfloor}^s H_\lambda(\overline{\theta}_{\lfloor r \rfloor}^\lambda) h_\lambda(\overline{\theta}_{\lfloor r \rfloor}^\lambda) \, \mathrm{d}r \, \mathrm{d}s \right|^2 \right] \\
&\quad + 12\mathbb{E}\left[\left| -\lambda^2\beta^{-1} \int_{\lfloor t \rfloor}^t H(\overline{\theta}_s^\lambda) \int_{\lfloor s \rfloor}^s \Upsilon_\lambda(\overline{\theta}_{\lfloor r \rfloor}^\lambda) \, \mathrm{d}r \, \mathrm{d}s \right|^2 \right] \\
&\quad + 12\mathbb{E}\left[\left| -\lambda\sqrt{2\lambda\beta^{-1}} \int_{\lfloor t \rfloor}^t H(\overline{\theta}_s^\lambda) \int_{\lfloor s \rfloor}^s H_\lambda(\overline{\theta}_{\lfloor r \rfloor}^\lambda) \, \mathrm{d}B_r^\lambda \, \mathrm{d}s \right|^2 \right] \\
&\quad + 12\mathbb{E}\left[\left| \sqrt{2\lambda\beta^{-1}} \int_{\lfloor t \rfloor}^t \left(H(\overline{\theta}_s^\lambda) - H(\overline{\theta}_{\lfloor s \rfloor}^\lambda) \right) \, \mathrm{d}B_s^\lambda \right|^2 \right] \\
&\leq 12\lambda^2 \left(e^{-\mathsf{a}_\mathsf{D} \min\{\kappa,1/2\}n/2} \overline{\mathsf{C}}_{\mathbf{A}2} \mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2} \right) \\
&\quad + 60\lambda^2 \left(e^{-\mathsf{a}_\mathsf{D} \min\{\kappa,1/2\}n/2} \overline{\mathsf{C}}_{\mathbf{A}2} \mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2} \right) \\
&\leq \lambda^2 \left(e^{-\mathsf{a}_\mathsf{D} \min\{\kappa,1/2\}n/2} 72\overline{\mathsf{C}}_{\mathbf{A}2} \mathbb{E}[|\theta_0|^{16(\rho+1)}] + 72\widetilde{\mathsf{C}}_{\mathbf{A}2} \right),
\end{aligned}
$$

where the second last inequality holds by using Lemma A.3 with the fact that

$$
\begin{aligned}
&\mathbb{E}\left[\left| \sqrt{2\lambda\beta^{-1}} \int_{\lfloor t \rfloor}^t \left(H(\overline{\theta}_s^\lambda) - H(\overline{\theta}_{\lfloor s \rfloor}^\lambda) \right) \, \mathrm{d}B_s^\lambda \right|^2 \right] \\
&= 2\lambda\beta^{-1} \int_{\lfloor t \rfloor}^t \mathbb{E}\left[|H(\overline{\theta}_s^\lambda) - H(\overline{\theta}_{\lfloor s \rfloor}^\lambda)|_\mathsf{F}^2 \right] \, \mathrm{d}s \\
&\leq \lambda^2 \left(e^{-\mathsf{a}_\mathsf{D} \min\{\kappa,1/2\}n/2} \overline{\mathsf{C}}_{\mathbf{A}2} \mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2} \right)
\end{aligned}
$$

as indicated in the calculations in (v) of Lemma A.3.

This completes the proof. □

Definition A.5. *Define $\mathfrak{M} = (\mathfrak{M}^{(i,j)})_{i,j=1,\dots,d} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ by setting, for every $i, j = 1, \dots, d$,*

$$
\mathfrak{M}^{(i,j)}(\theta, \overline{\theta}) = \langle \nabla H^{(i,j)}(\overline{\theta}), \theta - \overline{\theta} \rangle, \qquad \theta, \overline{\theta} \in \mathbb{R}^d.
$$

Lemma A.6. *Let Assumptions 1, 2, and 3 hold. Then, for any $0 < \lambda \leq \lambda_{\max}$, $n \in \mathbb{N}_0$, $, t \geq nT$, we obtain the following inequalities:*

$$\begin{aligned} &\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t \rfloor}^{t}\left(H(\overline{\theta}_s^{\lambda}) - H(\overline{\theta}_{\lfloor s \rfloor}^{\lambda}) - \mathfrak{M}(\overline{\theta}_s^{\lambda}, \overline{\theta}_{\lfloor s \rfloor}^{\lambda})\right)\,\mathrm{d}B_s^{\lambda}\right|^2\right] \\ &\leq \lambda^{2+q}\left(e^{-\mathsf{a}_{\mathsf{D}}\kappa n/2}\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2}\right), \end{aligned} \tag{91}$$

$$\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t \rfloor}^{t}\mathfrak{M}(\overline{\theta}_s^{\lambda}, \overline{\theta}_{\lfloor s \rfloor}^{\lambda})\,\mathrm{d}B_s^{\lambda}\right|^2\right] \leq \lambda^{2}\left(e^{-\mathsf{a}_{\mathsf{D}}\kappa n/2}\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2}\right), \tag{92}$$

$$\begin{aligned} &\mathbb{E}\left[2\lambda\beta^{-1}\left\langle\int_{\lfloor t \rfloor}^{t}\int_{\lfloor s \rfloor}^{s}\left(H(\overline{\zeta}_{\lfloor r \rfloor}^{\lambda,n}) - H(\overline{\theta}_{\lfloor r \rfloor}^{\lambda})\right)\,\mathrm{d}B_r^{\lambda}\,\mathrm{d}s, \int_{\lfloor t \rfloor}^{t}\mathfrak{M}(\overline{\theta}_s^{\lambda}, \overline{\theta}_{\lfloor s \rfloor}^{\lambda})\,\mathrm{d}B_s^{\lambda}\right\rangle\right] \\ &\leq \lambda^{2}\left(e^{-\mathsf{a}_{\mathsf{D}}\min\{\kappa,1/2\}n/2}10\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}] + 10\widetilde{\mathsf{C}}_{\mathbf{A}2}\right), \end{aligned} \tag{93}$$

where $\overline{\mathsf{C}}_{\mathbf{A}2}$ and $\widetilde{\mathsf{C}}_{\mathbf{A}2}$ are given in (88).

Proof. To show the inequalities hold, we follow the arguments below:

(i) To establish the first inequality (91), by using Definition A.5, we observe that, for fixed $\theta, \overline{\theta} \in \mathbb{R}^d$,

$$\begin{aligned} &|H(\theta) - H(\overline{\theta}) - \mathfrak{M}(\theta, \overline{\theta})|_{\mathsf{F}}^2 \\ &= \sum_{i,j=1}^{d}|H^{(i,j)}(\theta) - H^{(i,j)}(\overline{\theta}) - \mathfrak{M}^{(i,j)}(\theta, \overline{\theta})|^2 \\ &= \sum_{i,j=1}^{d}\left|\int_0^1\langle\nabla H^{(i,j)}(\nu\theta + (1-\nu)\overline{\theta}), \theta - \overline{\theta}\rangle\,\mathrm{d}\nu - \langle\nabla H^{(i,j)}(\overline{\theta}), \theta - \overline{\theta}\rangle\right|^2 \\ &\leq \int_0^1\sum_{i,j=1}^{d}|\nabla H^{(i,j)}(\nu\theta + (1-\nu)\overline{\theta}) - \nabla H^{(i,j)}(\overline{\theta})|^2\,\mathrm{d}\nu|\theta - \overline{\theta}|^2 \\ &\leq \int_0^1\sum_{i=1}^{d}d|\nabla^2 h^{(i)}(\nu\theta + (1-\nu)\overline{\theta}) - \nabla^2 h^{(i)}(\overline{\theta})|^2\,\mathrm{d}\nu|\theta - \overline{\theta}|^2 \\ &\leq 2^{2(\rho-2)}d^2L^2(1 + |\theta| + |\overline{\theta}|)^{2(\rho-2)}|\theta - \overline{\theta}|^{2+2q}. \end{aligned} \tag{94}$$

By using (94), we obtain that

$$\begin{aligned} &\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t \rfloor}^{t}\left(H(\overline{\theta}_s^{\lambda}) - H(\overline{\theta}_{\lfloor s \rfloor}^{\lambda}) - \mathfrak{M}(\overline{\theta}_s^{\lambda}, \overline{\theta}_{\lfloor s \rfloor}^{\lambda})\right)\,\mathrm{d}B_s^{\lambda}\right|^2\right] \\ &= 2\lambda\beta^{-1}\int_{\lfloor t \rfloor}^{t}\mathbb{E}\left[\left|H(\overline{\theta}_s^{\lambda}) - H(\overline{\theta}_{\lfloor s \rfloor}^{\lambda}) - \mathfrak{M}(\overline{\theta}_s^{\lambda}, \overline{\theta}_{\lfloor s \rfloor}^{\lambda})\right|_{\mathsf{F}}^2\right]\,\mathrm{d}s \\ &\leq \lambda\beta^{-1}2^{2\rho-3}d^2L^2\int_{\lfloor t \rfloor}^{t}\mathbb{E}\left[(1 + |\overline{\theta}_s^{\lambda}| + |\overline{\theta}_{\lfloor s \rfloor}^{\lambda}|)^{2(\rho-2)}|\overline{\theta}_s^{\lambda} - \overline{\theta}_{\lfloor s \rfloor}^{\lambda}|^{2+2q}\right]\,\mathrm{d}s \\ &\leq \lambda\beta^{-1}2^{2\rho-3}d^2L^2\int_{\lfloor t \rfloor}^{t}3^{8(\rho+1)}\left(\mathbb{E}\left[1 + |\overline{\theta}_s^{\lambda}|^{16(\rho+1)} + |\overline{\theta}_{\lfloor s \rfloor}^{\lambda}|^{16(\rho+1)}\right]\right)^{1/2} \\ &\qquad\times\left(\mathbb{E}\left[|\overline{\theta}_s^{\lambda} - \overline{\theta}_{\lfloor s \rfloor}^{\lambda}|^{4+4q}\right]\right)^{1/2}\,\mathrm{d}s \\ &\leq \lambda\beta^{-1}3^{16(\rho+1)}d^2L^2\left(2e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t \rfloor}\mathbb{E}\left[|\theta_0|^{16(\rho+1)}\right] + 2\mathsf{c}_{8(\rho+1)}\left(1 + 1/(\mathsf{a}_{\mathsf{D}}\kappa)\right) + 1\right)^{1/2} \\ &\qquad\times\lambda^{1+q}\left(e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t \rfloor}\overline{\mathsf{C}}_{\mathbf{A}0,2+2q}\mathbb{E}[|\theta_0|^{4\lceil 2+2q\rceil(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}0,2+2q}\right)^{1/2} \\ &\leq \lambda^{2+q}\left(e^{-\mathsf{a}_{\mathsf{D}}\kappa n/2}\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2}\right), \end{aligned}$$

where the third inequality holds due to Lemma 6.3 and A.2, and $\overline{\mathsf{C}}_{\mathbf{A}2}$, $\widetilde{\mathsf{C}}_{\mathbf{A}2}$ are given in (88).

(ii) To establish the second inequality (92), we recall Definition A.5 and use Remark 2.2 to obtain

$$\begin{aligned}
&\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^{t}\mathfrak{M}(\overline{\theta}_s^{\lambda},\overline{\theta}_{\lfloor s\rfloor}^{\lambda})\,\mathrm{d}B_s^{\lambda}\right|^2\right]\\
&=2\lambda\beta^{-1}\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[\sum_{i,j=1}^{d}\left|\left\langle\nabla H^{(i,j)}(\overline{\theta}_{\lfloor s\rfloor}^{\lambda}),\overline{\theta}_s^{\lambda}-\overline{\theta}_{\lfloor s\rfloor}^{\lambda}\right\rangle\right|^2\right]\mathrm{d}s\\
&\leq 2\lambda\beta^{-1}\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[d^2\mathsf{K}_0^2(1+|\overline{\theta}_{\lfloor s\rfloor}^{\lambda}|)^{2(\rho+q-2)}|\overline{\theta}_s^{\lambda}-\overline{\theta}_{\lfloor s\rfloor}^{\lambda}|^2\right]\mathrm{d}s\\
&\leq \lambda\beta^{-1}d^2\mathsf{K}_0^2\int_{\lfloor t\rfloor}^{t}2^{8(\rho+1)+1}\left(\mathbb{E}\left[1+|\overline{\theta}_{\lfloor s\rfloor}^{\lambda}|^{16(\rho+1)}\right]\right)^{1/2}\left(\mathbb{E}\left[|\overline{\theta}_s^{\lambda}-\overline{\theta}_{\lfloor s\rfloor}^{\lambda}|^4\right]\right)^{1/2}\mathrm{d}s\\
&\leq \lambda\beta^{-1}d^2\mathsf{K}_0^2 2^{8(\rho+1)+1}\left(e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t\rfloor}\mathbb{E}\left[|\theta_0|^{16(\rho+1)}\right]+\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_{\mathsf{D}}\kappa)\right)+1\right)^{1/2}\\
&\qquad\times\lambda\left(e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t\rfloor}\overline{\mathsf{C}}_{\mathbf{A}0,2}\mathbb{E}[|\theta_0|^{8(\rho+1)}]+\widetilde{\mathsf{C}}_{\mathbf{A}0,2}\right)^{1/2}\\
&\leq \lambda^2\left(e^{-\mathsf{a}_{\mathsf{D}}\kappa n/2}\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}]+\widetilde{\mathsf{C}}_{\mathbf{A}2}\right),
\end{aligned}$$

where the third inequality holds due to Lemma 6.3 and A.2, and $\overline{\mathsf{C}}_{\mathbf{A}2}$, $\widetilde{\mathsf{C}}_{\mathbf{A}2}$ are given in (88).

(iii) To obtain the third inequality (93), we use Definition A.5 and write the following

$$\begin{aligned}
&\mathbb{E}\left[2\lambda\beta^{-1}\left\langle\int_{\lfloor t\rfloor}^{t}\int_{\lfloor s\rfloor}^{s}\left(H(\overline{\zeta}_{\lfloor r\rfloor}^{\lambda,n})-H(\overline{\theta}_{\lfloor r\rfloor}^{\lambda})\right)\mathrm{d}B_r^{\lambda}\,\mathrm{d}s,\int_{\lfloor t\rfloor}^{t}\mathfrak{M}(\overline{\theta}_s^{\lambda},\overline{\theta}_{\lfloor s\rfloor}^{\lambda})\,\mathrm{d}B_s^{\lambda}\right\rangle\right]\\
&=2\lambda\beta^{-1}\mathbb{E}\left[\left\langle\int_{\lfloor t\rfloor}^{t}\int_{\lfloor s\rfloor}^{s}\left(H(\overline{\zeta}_{\lfloor r\rfloor}^{\lambda,n})-H(\overline{\theta}_{\lfloor r\rfloor}^{\lambda})\right)\mathrm{d}B_r^{\lambda}\,\mathrm{d}s,\right.\right.\\
&\qquad\left.\left.\int_{\lfloor t\rfloor}^{t}\left\langle\nabla H(\overline{\theta}_{\lfloor s\rfloor}^{\lambda}),\int_{\lfloor s\rfloor}^{s}-\lambda h_\lambda(\overline{\theta}_{\lfloor r\rfloor}^{\lambda})\,\mathrm{d}r\right\rangle\mathrm{d}B_s^{\lambda}\right\rangle\right]\\
&\quad+2\lambda\beta^{-1}\mathbb{E}\left[\left\langle\int_{\lfloor t\rfloor}^{t}\int_{\lfloor s\rfloor}^{s}\left(H(\overline{\zeta}_{\lfloor r\rfloor}^{\lambda,n})-H(\overline{\theta}_{\lfloor r\rfloor}^{\lambda})\right)\mathrm{d}B_r^{\lambda}\,\mathrm{d}s,\right.\right.\\
&\qquad\left.\left.\int_{\lfloor t\rfloor}^{t}\left\langle\nabla H(\overline{\theta}_{\lfloor s\rfloor}^{\lambda}),\int_{\lfloor s\rfloor}^{s}\int_{\lfloor r\rfloor}^{r}\lambda^2H_\lambda(\overline{\theta}_{\lfloor \nu\rfloor}^{\lambda})h_\lambda(\overline{\theta}_{\lfloor \nu\rfloor}^{\lambda})\,\mathrm{d}\nu\,\mathrm{d}r\right\rangle\mathrm{d}B_s^{\lambda}\right\rangle\right]\\
&\quad+2\lambda\beta^{-1}\mathbb{E}\left[\left\langle\int_{\lfloor t\rfloor}^{t}\int_{\lfloor s\rfloor}^{s}\left(H(\overline{\zeta}_{\lfloor r\rfloor}^{\lambda,n})-H(\overline{\theta}_{\lfloor r\rfloor}^{\lambda})\right)\mathrm{d}B_r^{\lambda}\,\mathrm{d}s,\right.\right.\\
&\qquad\left.\left.\int_{\lfloor t\rfloor}^{t}\left\langle\nabla H(\overline{\theta}_{\lfloor s\rfloor}^{\lambda}),\int_{\lfloor s\rfloor}^{s}\int_{\lfloor r\rfloor}^{r}-\lambda^2\beta^{-1}\Upsilon_\lambda(\overline{\theta}_{\lfloor \nu\rfloor}^{\lambda})\,\mathrm{d}\nu\,\mathrm{d}r\right\rangle\mathrm{d}B_s^{\lambda}\right\rangle\right]\\
&\quad+2\lambda\beta^{-1}\mathbb{E}\left[\left\langle\int_{\lfloor t\rfloor}^{t}\int_{\lfloor s\rfloor}^{s}\left(H(\overline{\zeta}_{\lfloor r\rfloor}^{\lambda,n})-H(\overline{\theta}_{\lfloor r\rfloor}^{\lambda})\right)\mathrm{d}B_r^{\lambda}\,\mathrm{d}s,\right.\right.\\
&\qquad\left.\left.\int_{\lfloor t\rfloor}^{t}\left\langle\nabla H(\overline{\theta}_{\lfloor s\rfloor}^{\lambda}),\int_{\lfloor s\rfloor}^{s}-\lambda\sqrt{2\lambda\beta^{-1}}\int_{\lfloor r\rfloor}^{r}H_\lambda(\overline{\theta}_{\lfloor \nu\rfloor}^{\lambda})\,\mathrm{d}B_\nu^{\lambda}\,\mathrm{d}r\right\rangle\mathrm{d}B_s^{\lambda}\right\rangle\right]\\
&\quad+2\lambda\beta^{-1}\mathbb{E}\left[\left\langle\int_{\lfloor t\rfloor}^{t}\int_{\lfloor s\rfloor}^{s}\left(H(\overline{\zeta}_{\lfloor r\rfloor}^{\lambda,n})-H(\overline{\theta}_{\lfloor r\rfloor}^{\lambda})\right)\mathrm{d}B_r^{\lambda}\,\mathrm{d}s,\right.\right.\\
&\qquad\left.\left.\int_{\lfloor t\rfloor}^{t}\left\langle\nabla H(\overline{\theta}_{\lfloor s\rfloor}^{\lambda}),\sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^{s}\mathrm{d}B_r^{\lambda}\right\rangle\mathrm{d}B_s^{\lambda}\right\rangle\right].
\end{aligned}\tag{95}$$

Recall that $(\mathcal{F}_t^\lambda)_{t\geq 0}$ is the completed natural filtration of $(B_t^\lambda)_{t\geq 0}$. Then, we note that, for any $i,j,k=1,\dots,d$, it holds that

$$
\begin{aligned}
&\mathbb{E}\left[\left(\int_{\lfloor t\rfloor}^t\int_{\lfloor s\rfloor}^s\left(H^{(i,j)}(\overline{\zeta}^{\lambda,n}_{\lfloor r\rfloor})-H^{(i,j)}(\overline{\theta}^{\lambda}_{\lfloor r\rfloor})\right)\,\mathrm{d}(B_r^\lambda)^{(j)}\,\mathrm{d}s\right)\right.\\
&\qquad\left.\times\left(\int_{\lfloor t\rfloor}^t\left(\partial_{\theta^{(k)}}H^{(i,j)}(\overline{\theta}^\lambda_{\lfloor s\rfloor})\int_{\lfloor s\rfloor}^s\mathrm{d}(B_r^\lambda)^{(k)}\right)\,\mathrm{d}(B_s^\lambda)^{(j)}\right)\right]\\
&=\mathbb{E}\left[\left(H^{(i,j)}(\overline{\zeta}^{\lambda,n}_{\lfloor t\rfloor})-H^{(i,j)}(\overline{\theta}^{\lambda}_{\lfloor t\rfloor})\right)\partial_{\theta^{(k)}}H^{(i,j)}(\overline{\theta}^\lambda_{\lfloor t\rfloor})\right.\\
&\qquad\left.\times\,\mathbb{E}\left[\int_{\lfloor t\rfloor}^t\int_{\lfloor s\rfloor}^s\mathrm{d}(B_r^\lambda)^{(j)}\,\mathrm{d}s\int_{\lfloor t\rfloor}^t\int_{\lfloor s\rfloor}^s\mathrm{d}(B_r^\lambda)^{(k)}\,\mathrm{d}(B_s^\lambda)^{(j)}\,\middle|\,\mathcal{F}^\lambda_{\lfloor t\rfloor}\right]\right]\\
&=\mathbb{E}\left[\left(H^{(i,j)}(\overline{\zeta}^{\lambda,n}_{\lfloor t\rfloor})-H^{(i,j)}(\overline{\theta}^{\lambda}_{\lfloor t\rfloor})\right)\partial_{\theta^{(k)}}H^{(i,j)}(\overline{\theta}^\lambda_{\lfloor t\rfloor})\right.\\
&\qquad\left.\times\,\mathbb{E}\left[\int_{\lfloor t\rfloor}^t(t-s)\,\mathrm{d}(B_s^\lambda)^{(j)}\int_{\lfloor t\rfloor}^t\int_{\lfloor s\rfloor}^s\mathrm{d}(B_r^\lambda)^{(k)}\,\mathrm{d}(B_s^\lambda)^{(j)}\,\middle|\,\mathcal{F}^\lambda_{\lfloor t\rfloor}\right]\right]\\
&=\mathbb{E}\left[\left(H^{(i,j)}(\overline{\zeta}^{\lambda,n}_{\lfloor t\rfloor})-H^{(i,j)}(\overline{\theta}^{\lambda}_{\lfloor t\rfloor})\right)\partial_{\theta^{(k)}}H^{(i,j)}(\overline{\theta}^\lambda_{\lfloor t\rfloor})\right.\\
&\qquad\left.\times\left(t\int_{\lfloor t\rfloor}^t\mathbb{E}\left[\int_{\lfloor t\rfloor}^s\mathrm{d}(B_r^\lambda)^{(k)}\,\middle|\,\mathcal{F}^\lambda_{\lfloor t\rfloor}\right]\mathrm{d}s-\int_{\lfloor t\rfloor}^t s\mathbb{E}\left[\int_{\lfloor t\rfloor}^s\mathrm{d}(B_r^\lambda)^{(k)}\,\middle|\,\mathcal{F}^\lambda_{\lfloor t\rfloor}\right]\mathrm{d}s\right)\right]\\
&=0.
\end{aligned}
$$

This implies that the last term in (95) is zero. Indeed, we have that

$$
\begin{aligned}
&2\lambda\beta^{-1}\mathbb{E}\left[\left\langle\int_{\lfloor t\rfloor}^t\int_{\lfloor s\rfloor}^s\left(H(\overline{\zeta}^{\lambda,n}_{\lfloor r\rfloor})-H(\overline{\theta}^{\lambda}_{\lfloor r\rfloor})\right)\,\mathrm{d}B_r^\lambda\,\mathrm{d}s,\right.\right.\\
&\qquad\left.\left.\int_{\lfloor t\rfloor}^t\left\langle\nabla H(\overline{\theta}^\lambda_{\lfloor s\rfloor}),\sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^s\mathrm{d}B_r^\lambda\right\rangle\,\mathrm{d}B_s^\lambda\right\rangle\right]\\
&=(2\lambda\beta^{-1})^{3/2}\mathbb{E}\left[\sum_{i=1}^d\left(\int_{\lfloor t\rfloor}^t\sum_{j=1}^d\int_{\lfloor s\rfloor}^s\left(H^{(i,j)}(\overline{\zeta}^{\lambda,n}_{\lfloor r\rfloor})-H^{(i,j)}(\overline{\theta}^{\lambda}_{\lfloor r\rfloor})\right)\,\mathrm{d}(B_r^\lambda)^{(j)}\,\mathrm{d}s\right)\right.\\
&\qquad\left.\times\left(\sum_{j=1}^d\int_{\lfloor t\rfloor}^t\left(\sum_{k=1}^d\partial_{\theta^{(k)}}H^{(i,j)}(\overline{\theta}^\lambda_{\lfloor s\rfloor})\int_{\lfloor s\rfloor}^s\mathrm{d}(B_r^\lambda)^{(k)}\right)\,\mathrm{d}(B_s^\lambda)^{(j)}\right)\right]=0.
\end{aligned}
$$

Then, by using Remark 2.2 and (95) with the result above, we obtain that

$$
\begin{aligned}
&\mathbb{E}\left[2\lambda\beta^{-1}\left\langle\int_{\lfloor t\rfloor}^t\int_{\lfloor s\rfloor}^s\left(H(\overline{\zeta}^{\lambda,n}_{\lfloor r\rfloor})-H(\overline{\theta}^{\lambda}_{\lfloor r\rfloor})\right)\,\mathrm{d}B_r^\lambda\,\mathrm{d}s,\int_{\lfloor t\rfloor}^t\mathfrak{M}(\overline{\theta}^\lambda_s,\overline{\theta}^\lambda_{\lfloor s\rfloor})\,\mathrm{d}B_s^\lambda\right\rangle\right]\\
&\leq 4\lambda^2\beta^{-2}\mathbb{E}\left[\left|\int_{\lfloor t\rfloor}^t\int_{\lfloor s\rfloor}^s\left(H(\overline{\zeta}^{\lambda,n}_{\lfloor r\rfloor})-H(\overline{\theta}^{\lambda}_{\lfloor r\rfloor})\right)\,\mathrm{d}B_r^\lambda\,\mathrm{d}s\right|^2\right]\\
&\quad+\lambda^2\mathbb{E}\left[\left|\int_{\lfloor t\rfloor}^t\left\langle\nabla H(\overline{\theta}^\lambda_{\lfloor s\rfloor}),\int_{\lfloor s\rfloor}^s-h_\lambda(\overline{\theta}^\lambda_{\lfloor r\rfloor})\,\mathrm{d}r\right\rangle\,\mathrm{d}B_s^\lambda\right|^2\right]\\
&\quad+\lambda^4\mathbb{E}\left[\left|\int_{\lfloor t\rfloor}^t\left\langle\nabla H(\overline{\theta}^\lambda_{\lfloor s\rfloor}),\int_{\lfloor s\rfloor}^s\int_{\lfloor r\rfloor}^r H_\lambda(\overline{\theta}^\lambda_{\lfloor\nu\rfloor})h_\lambda(\overline{\theta}^\lambda_{\lfloor\nu\rfloor})\,\mathrm{d}\nu\,\mathrm{d}r\right\rangle\,\mathrm{d}B_s^\lambda\right|^2\right]
\end{aligned}
$$

$$
\begin{aligned}
&+ \lambda^4\beta^{-2}\mathbb{E}\left[\left|\int_{\lfloor t\rfloor}^t \left\langle \nabla H(\overline{\theta}^\lambda_{\lfloor s\rfloor}), \int_{\lfloor s\rfloor}^s \int_{\lfloor r\rfloor}^r -\Upsilon_\lambda(\overline{\theta}^\lambda_{\lfloor \nu\rfloor})\,\mathrm{d}\nu\,\mathrm{d}r \right\rangle \mathrm{d}B^\lambda_s\right|^2\right] \\
&+ 2\lambda^3\beta^{-1}\mathbb{E}\left[\left|\int_{\lfloor t\rfloor}^t \left\langle \nabla H(\overline{\theta}^\lambda_{\lfloor s\rfloor}), \int_{\lfloor s\rfloor}^s - \int_{\lfloor r\rfloor}^r H_\lambda(\overline{\theta}^\lambda_{\lfloor \nu\rfloor})\,\mathrm{d}B^\lambda_\nu\,\mathrm{d}r \right\rangle \mathrm{d}B^\lambda_s\right|^2\right] \\
\leq\ & 8\lambda^2\beta^{-2}\int_{\lfloor t\rfloor}^t \int_{\lfloor s\rfloor}^s \mathbb{E}\left[|H(\overline{\zeta}^{\lambda,n}_{\lfloor r\rfloor})|_{\mathsf{F}}^2 + |H(\overline{\theta}^\lambda_{\lfloor r\rfloor})|_{\mathsf{F}}^2\right]\mathrm{d}r\,\mathrm{d}s \\
&+ \lambda^2\int_{\lfloor t\rfloor}^t \mathbb{E}\left[\sum_{i,j=1}^d \left|\left\langle \nabla H^{(i,j)}(\overline{\theta}^\lambda_{\lfloor s\rfloor}), h_\lambda(\overline{\theta}^\lambda_{\lfloor s\rfloor})\right\rangle\right|^2\right]\mathrm{d}s \\
&+ \lambda^4\int_{\lfloor t\rfloor}^t \mathbb{E}\left[\sum_{i,j=1}^d \left|\left\langle \nabla H^{(i,j)}(\overline{\theta}^\lambda_{\lfloor s\rfloor}), H_\lambda(\overline{\theta}^\lambda_{\lfloor s\rfloor})h_\lambda(\overline{\theta}^\lambda_{\lfloor s\rfloor})\right\rangle\right|^2\right]\mathrm{d}s \\
&+ \lambda^4\beta^{-2}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[\sum_{i,j=1}^d \left|\left\langle \nabla H^{(i,j)}(\overline{\theta}^\lambda_{\lfloor s\rfloor}), \Upsilon_\lambda(\overline{\theta}^\lambda_{\lfloor s\rfloor})\right\rangle\right|^2\right]\mathrm{d}s \\
&+ 2\lambda^3\beta^{-1}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[\sum_{i,j=1}^d \left|\left\langle \nabla H^{(i,j)}(\overline{\theta}^\lambda_{\lfloor s\rfloor}), H_\lambda(\overline{\theta}^\lambda_{\lfloor s\rfloor})\int_{\lfloor s\rfloor}^s\int_{\lfloor r\rfloor}^r \mathrm{d}B^\lambda_\nu\,\mathrm{d}r\right\rangle\right|^2\right]\mathrm{d}s \\
\leq\ & 8\lambda^2\beta^{-2}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[dK_H^2(1+|\overline{\zeta}^{\lambda,n}_{\lfloor s\rfloor}|^{\rho+q-1})^2 + dK_H^2(1+|\overline{\theta}^\lambda_{\lfloor s\rfloor}|^{\rho+q-1})^2\right]\mathrm{d}s \\
&+ \lambda^2\int_{\lfloor t\rfloor}^t \mathbb{E}\left[d^2\mathsf{K}_0^2(1+|\overline{\theta}^\lambda_{\lfloor s\rfloor}|)^{2(\rho+q-2)}K_h^2(1+|\overline{\theta}^\lambda_{\lfloor s\rfloor}|^{\rho+q})^2\right]\mathrm{d}s \\
&+ \lambda^4\int_{\lfloor t\rfloor}^t \mathbb{E}\left[d^2\mathsf{K}_0^2(1+|\overline{\theta}^\lambda_{\lfloor s\rfloor}|)^{2(\rho+q-2)}K_H^2(1+|\overline{\theta}^\lambda_{\lfloor s\rfloor}|^{\rho+q-1})^2K_h^2(1+|\overline{\theta}^\lambda_{\lfloor s\rfloor}|^{\rho+q})^2\right]\mathrm{d}s \\
&+ \lambda^4\beta^{-2}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[d^2\mathsf{K}_0^2(1+|\overline{\theta}^\lambda_{\lfloor s\rfloor}|)^{2(\rho+q-2)}\mathsf{K}_{3,d}^2(1+|\overline{\theta}^\lambda_{\lfloor s\rfloor}|)^{2(\rho+q-2)}\right]\mathrm{d}s \\
&+ 2\lambda^3\beta^{-1}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[d^2\mathsf{K}_0^2(1+|\overline{\theta}^\lambda_{\lfloor s\rfloor}|)^{2(\rho+q-2)}K_H^2(1+|\overline{\theta}^\lambda_{\lfloor s\rfloor}|^{\rho+q-1})^2\left|\int_{\lfloor s\rfloor}^s\int_{\lfloor r\rfloor}^r \mathrm{d}B^\lambda_\nu\,\mathrm{d}r\right|^2\right]\mathrm{d}s \\
\leq\ & 6\lambda^2\beta^{-2}dK_H^2 3^{16(\rho+1)}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[1+|\overline{\zeta}^{\lambda,n}_{\lfloor s\rfloor}|^{16(\rho+1)} + |\overline{\theta}^\lambda_{\lfloor s\rfloor}|^{16(\rho+1)}\right]\mathrm{d}s \\
&+ \lambda^2(K_h^2 + K_H^2K_h^2 + \beta^{-2}\mathsf{K}_{3,d}^2)d^2\mathsf{K}_0^2 2^{16(\rho+1)}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[1+|\overline{\theta}^\lambda_{\lfloor s\rfloor}|^{16(\rho+1)}\right]\mathrm{d}s \\
&+ \lambda^2\beta^{-1}d^2\mathsf{K}_0^2K_H^2 2^{8(\rho+1)+1}\int_{\lfloor t\rfloor}^t \left(\mathbb{E}\left[1+|\overline{\theta}^\lambda_{\lfloor s\rfloor}|^{16(\rho+1)}\right]\right)^{1/2}(3(d+4))\,\mathrm{d}s \\
\leq\ & 6\lambda^2\beta^{-2}dK_H^2 3^{24(\rho+1)}\Big(e^{-\lambda\mathsf{a}_{\mathsf{D}}\min\{\kappa,1/2\}\lfloor t\rfloor}\mathbb{E}[|\theta_0|^{16(\rho+1)}] + \big(\mathsf{c}_{8(\rho+1)}(1+1/(\mathsf{a}_{\mathsf{D}}\kappa))+1\big) \\
&\quad +3\mathsf{v}_{16(\rho+1)}(\mathsf{M}_V(16(\rho+1)))\Big) \\
&+ \lambda^2(K_h^2 + K_H^2K_h^2 + \beta^{-2}\mathsf{K}_{3,d}^2 + \beta^{-1}K_H^2)d^3\mathsf{K}_0^2 2^{16(\rho+1)} \\
&\quad \times\left(e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t\rfloor}\mathbb{E}\left[|\theta_0|^{16(\rho+1)}\right] + \mathsf{c}_{8(\rho+1)}(1+1/(\mathsf{a}_{\mathsf{D}}\kappa))+1\right) \\
\leq\ & \lambda^2\left(e^{-\mathsf{a}_{\mathsf{D}}\min\{\kappa,1/2\}n/2}10\overline{\mathsf{C}}_{\mathbf{A2}}\mathbb{E}[|\theta_0|^{16(\rho+1)}] + 10\widetilde{\mathsf{C}}_{\mathbf{A2}}\right),
\end{aligned}
$$

where the fifth inequality holds due to Lemma 6.3, 6.5 and A.2, and $\overline{\mathsf{C}}_{\mathbf{A2}}$, $\widetilde{\mathsf{C}}_{\mathbf{A2}}$ are given in (88). This completes the proof. □

Proof of Lemma 6.6. By using the definitions of $\overline{\theta}_t^\lambda$ in (27) and $\overline{\zeta}_t^{\lambda,n}$ in Definition 6.2, and by applying Itô's formula, we obtain, for any $n \in \mathbb{N}_0$, $t \in (nT,(n+1)T]$,

$$
\begin{aligned}
&W_2^2(\mathcal{L}(\overline{\theta}_t^\lambda), \mathcal{L}(\overline{\zeta}_t^{\lambda,n}))\\
&\leq \mathbb{E}\left[\left|\overline{\theta}_t^\lambda - \overline{\zeta}_t^{\lambda,n}\right|^2\right]\\
&= -2\lambda\mathbb{E}\left[\int_{nT}^t \left\langle \overline{\theta}_s^\lambda - \overline{\zeta}_s^{\lambda,n}, h_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda) - h(\overline{\zeta}_s^{\lambda,n})\right\rangle \mathrm{d}s\right]\\
&\quad + 2\lambda\mathbb{E}\left[\int_{nT}^t \left\langle \overline{\theta}_s^\lambda - \overline{\zeta}_s^{\lambda,n}, \lambda\int_{\lfloor s\rfloor}^s \left(H_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)h_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda) - \beta^{-1}\Upsilon_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)\right)\mathrm{d}r\right\rangle \mathrm{d}s\right]\\
&\quad - 2\lambda\mathbb{E}\left[\int_{nT}^t \left\langle \overline{\theta}_s^\lambda - \overline{\zeta}_s^{\lambda,n}, \sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^s H_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\right\rangle \mathrm{d}s\right]\\
&= -2\lambda\mathbb{E}\left[\int_{nT}^t \left\langle \overline{\theta}_s^\lambda - \overline{\zeta}_s^{\lambda,n}, h(\overline{\theta}_s^\lambda) - h(\overline{\zeta}_s^{\lambda,n})\right\rangle \mathrm{d}s\right]\\
&\quad - 2\lambda\mathbb{E}\left[\int_{nT}^t \left\langle \overline{\theta}_s^\lambda - \overline{\zeta}_s^{\lambda,n}, h(\overline{\theta}_{\lfloor s\rfloor}^\lambda) - h(\overline{\theta}_s^\lambda)\right\rangle \mathrm{d}s\right]\\
&\quad - 2\lambda\mathbb{E}\left[\int_{nT}^t \left\langle \overline{\theta}_s^\lambda - \overline{\zeta}_s^{\lambda,n}, h_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda) - h(\overline{\theta}_{\lfloor s\rfloor}^\lambda)\right\rangle \mathrm{d}s\right]\\
&\quad + 2\lambda\mathbb{E}\left[\int_{nT}^t \left\langle \overline{\theta}_s^\lambda - \overline{\zeta}_s^{\lambda,n}, \lambda\int_{\lfloor s\rfloor}^s \left(H_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)h_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda) - \beta^{-1}\Upsilon_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)\right)\mathrm{d}r\right\rangle \mathrm{d}s\right]\\
&\quad - 2\lambda\mathbb{E}\left[\int_{nT}^t \left\langle \overline{\theta}_s^\lambda - \overline{\zeta}_s^{\lambda,n}, \sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^s H_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\right\rangle \mathrm{d}s\right].
\end{aligned}
\tag{96}
$$

By applying Itô's formula to $h(\overline{\theta}_s^\lambda)$, we obtain (89). Substituting (89) into (96), applying Remark 2.5 and Young's inequality yield

$$
\begin{aligned}
&\mathbb{E}\left[\left|\overline{\theta}_t^\lambda - \overline{\zeta}_t^{\lambda,n}\right|^2\right]\\
&\leq 2\lambda \mathsf{L}_{\mathsf{OS}}\int_{nT}^t \mathbb{E}\left[|\overline{\theta}_s^\lambda - \overline{\zeta}_s^{\lambda,n}|^2\right]\mathrm{d}s\\
&\quad + 2\lambda\mathbb{E}\left[\int_{nT}^t \left\langle \overline{\theta}_s^\lambda - \overline{\zeta}_s^{\lambda,n}, -\lambda\int_{\lfloor s\rfloor}^s \left(H(\overline{\theta}_r^\lambda) - H_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)\right)h_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}r\right\rangle \mathrm{d}s\right]\\
&\quad + 2\lambda\mathbb{E}\left[\int_{nT}^t \left\langle \overline{\theta}_s^\lambda - \overline{\zeta}_s^{\lambda,n}, \lambda^2\int_{\lfloor s\rfloor}^s H(\overline{\theta}_r^\lambda)\int_{\lfloor r\rfloor}^r H_\lambda(\overline{\theta}_{\lfloor \nu\rfloor}^\lambda)h_\lambda(\overline{\theta}_{\lfloor \nu\rfloor}^\lambda)\,\mathrm{d}\nu\,\mathrm{d}r\right\rangle \mathrm{d}s\right]\\
&\quad + 2\lambda\mathbb{E}\left[\int_{nT}^t \left\langle \overline{\theta}_s^\lambda - \overline{\zeta}_s^{\lambda,n}, -\lambda^2\beta^{-1}\int_{\lfloor s\rfloor}^s H(\overline{\theta}_r^\lambda)\int_{\lfloor r\rfloor}^r \Upsilon_\lambda(\overline{\theta}_{\lfloor \nu\rfloor}^\lambda)\,\mathrm{d}\nu\,\mathrm{d}r\right\rangle \mathrm{d}s\right]\\
&\quad + 2\lambda\mathbb{E}\left[\int_{nT}^t \left\langle \overline{\theta}_s^\lambda - \overline{\zeta}_s^{\lambda,n}, -\lambda\sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^s H(\overline{\theta}_r^\lambda)\int_{\lfloor r\rfloor}^r H_\lambda(\overline{\theta}_{\lfloor \nu\rfloor}^\lambda)\,\mathrm{d}B_\nu^\lambda\,\mathrm{d}r\right\rangle \mathrm{d}s\right]\\
&\quad + 2\lambda\mathbb{E}\left[\int_{nT}^t \left\langle \overline{\theta}_s^\lambda - \overline{\zeta}_s^{\lambda,n}, \sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^s \left(H(\overline{\theta}_r^\lambda) - H_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)\right)\mathrm{d}B_r^\lambda\right\rangle \mathrm{d}s\right]\\
&\quad + 2\lambda\mathbb{E}\left[\int_{nT}^t \left\langle \overline{\theta}_s^\lambda - \overline{\zeta}_s^{\lambda,n}, \lambda\beta^{-1}\int_{\lfloor s\rfloor}^s \left(\Upsilon(\overline{\theta}_r^\lambda) - \Upsilon_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)\right)\mathrm{d}r\right\rangle \mathrm{d}s\right]\\
&\quad - 2\lambda\mathbb{E}\left[\int_{nT}^t \left\langle \overline{\theta}_s^\lambda - \overline{\zeta}_s^{\lambda,n}, h_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda) - h(\overline{\theta}_{\lfloor s\rfloor}^\lambda)\right\rangle \mathrm{d}s\right]
\end{aligned}
$$

$$
\begin{aligned}
&\leq \lambda(2\mathsf{L}_{\mathsf{OS}}+6)\int_{nT}^{t}\mathbb{E}\left[|\overline{\theta}_s^\lambda-\overline{\zeta}_s^{\lambda,n}|^2\right]\,\mathrm{d}s \\
&+\lambda\int_{nT}^{t}\mathbb{E}\left[\left|-\lambda\int_{\lfloor s\rfloor}^{s}\left(H(\overline{\theta}_r^\lambda)-H_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)\right)h_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}r\right|^2\right]\,\mathrm{d}s \\
&+\lambda\int_{nT}^{t}\mathbb{E}\left[\left|\lambda^2\int_{\lfloor s\rfloor}^{s}H(\overline{\theta}_r^\lambda)\int_{\lfloor r\rfloor}^{r}H_\lambda(\overline{\theta}_{\lfloor\nu\rfloor}^\lambda)h_\lambda(\overline{\theta}_{\lfloor\nu\rfloor}^\lambda)\,\mathrm{d}\nu\,\mathrm{d}r\right|^2\right]\,\mathrm{d}s \\
&+\lambda\int_{nT}^{t}\mathbb{E}\left[\left|-\lambda^2\beta^{-1}\int_{\lfloor s\rfloor}^{s}H(\overline{\theta}_r^\lambda)\int_{\lfloor r\rfloor}^{r}\Upsilon_\lambda(\overline{\theta}_{\lfloor\nu\rfloor}^\lambda)\,\mathrm{d}\nu\,\mathrm{d}r\right|^2\right]\,\mathrm{d}s \\
&+\lambda\int_{nT}^{t}\mathbb{E}\left[\left|-\lambda\sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^{s}H(\overline{\theta}_r^\lambda)\int_{\lfloor r\rfloor}^{r}H_\lambda(\overline{\theta}_{\lfloor\nu\rfloor}^\lambda)\,\mathrm{d}B_\nu^\lambda\,\mathrm{d}r\right|^2\right]\,\mathrm{d}s \\
&+\lambda\int_{nT}^{t}\mathbb{E}\left[\left|\lambda\beta^{-1}\int_{\lfloor s\rfloor}^{s}\left(\Upsilon(\overline{\theta}_r^\lambda)-\Upsilon_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)\right)\mathrm{d}r\right|^2\right]\,\mathrm{d}s \\
&+\lambda\int_{nT}^{t}\mathbb{E}\left[\left|h(\overline{\theta}_{\lfloor s\rfloor}^\lambda)-h_\lambda(\overline{\theta}_{\lfloor s\rfloor}^\lambda)\right|^2\right]\,\mathrm{d}s \\
&+2\lambda\mathbb{E}\left[\int_{nT}^{t}\left\langle\overline{\theta}_s^\lambda-\overline{\zeta}_s^{\lambda,n},\sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^{s}\left(H(\overline{\theta}_r^\lambda)-H_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)\right)\mathrm{d}B_r^\lambda\right\rangle\mathrm{d}s\right].
\end{aligned}
\tag{97}
$$

We note that, by using (87), for any $t\geq nT$,

$$
\begin{aligned}
\mathbb{E}\left[|h(\overline{\theta}_{\lfloor t\rfloor}^\lambda)-h_\lambda(\overline{\theta}_{\lfloor t\rfloor}^\lambda)|^2\right]&\leq\mathbb{E}\left[\lambda^3|\overline{\theta}_{\lfloor t\rfloor}^\lambda|^{6(\rho+q-1)}K_h^2(1+|\overline{\theta}_{\lfloor t\rfloor}^\lambda|^{\rho+q})^2\right] \\
&\leq\lambda^3K_h^2 2^{16(\rho+1)}\mathbb{E}\left[1+|\overline{\theta}_{\lfloor t\rfloor}^\lambda|^{16(\rho+1)}\right] \\
&\leq\lambda^3K_h^2 2^{16(\rho+1)}\left(e^{-\lambda\mathsf{a}_{\mathsf{D}}\kappa\lfloor t\rfloor}\mathbb{E}\left[|\theta_0|^{16(\rho+1)}\right]+\mathsf{c}_{8(\rho+1)}\left(1+1/(\mathsf{a}_{\mathsf{D}}\kappa)\right)+1\right) \\
&\leq\lambda^3\left(e^{-\mathsf{a}_{\mathsf{D}}\kappa n/2}\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}]+\widetilde{\mathsf{C}}_{\mathbf{A}2}\right),
\end{aligned}
\tag{98}
$$

where the third inequality holds due to Lemma 6.3, and $\overline{\mathsf{C}}_{\mathbf{A}2}$, $\widetilde{\mathsf{C}}_{\mathbf{A}2}$ are given in (88). By using Lemma A.3 and (98), (97) becomes

$$
\begin{aligned}
\mathbb{E}\left[\left|\overline{\theta}_t^\lambda-\overline{\zeta}_t^{\lambda,n}\right|^2\right]\leq&\ \lambda(2\mathsf{L}_{\mathsf{OS}}+6)\int_{nT}^{t}\mathbb{E}\left[|\overline{\theta}_s^\lambda-\overline{\zeta}_s^{\lambda,n}|^2\right]\,\mathrm{d}s \\
&+6\lambda^{2+q}\left(e^{-\mathsf{a}_{\mathsf{D}}\kappa n/2}\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}]+\widetilde{\mathsf{C}}_{\mathbf{A}2}\right) \\
&+\mathfrak{J}_1+\mathfrak{J}_2,
\end{aligned}
\tag{99}
$$

where

$$
\begin{aligned}
\mathfrak{J}_1^\lambda(t)&:=2\lambda\mathbb{E}\left[\int_{nT}^{t}\left\langle\overline{\theta}_s^\lambda-\overline{\zeta}_s^{\lambda,n},\sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^{s}\left(H(\overline{\theta}_r^\lambda)-H_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)-\mathfrak{M}(\overline{\theta}_r^\lambda,\overline{\theta}_{\lfloor r\rfloor}^\lambda)\right)\mathrm{d}B_r^\lambda\right\rangle\mathrm{d}s\right], \\
\mathfrak{J}_2^\lambda(t)&:=2\lambda\mathbb{E}\left[\int_{nT}^{t}\left\langle\overline{\theta}_s^\lambda-\overline{\zeta}_s^{\lambda,n},\sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^{s}\mathfrak{M}(\overline{\theta}_r^\lambda,\overline{\theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\right\rangle\mathrm{d}s\right]
\end{aligned}
$$

with $\mathfrak{M}$ defined in Definition A.5. By using Young's inequality and Lemma A.6, we have that

$$
\mathfrak{J}_1^\lambda(t)\leq\lambda\int_{nT}^{t}\mathbb{E}\left[|\overline{\theta}_s^\lambda-\overline{\zeta}_s^{\lambda,n}|^2\right]\,\mathrm{d}s+\lambda^{2+q}\left(e^{-\mathsf{a}_{\mathsf{D}}\kappa n/2}\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}]+\widetilde{\mathsf{C}}_{\mathbf{A}2}\right).
\tag{100}
$$

To establish an upper bound for $\mathfrak{J}_2^\lambda(t)$, we recall the definitions of $(\overline{\theta}_t^\lambda)_{t\geq 0}$ and $(\overline{\zeta}_t^{\lambda,n})_{t\geq 0}$ given in (27) and Definition 6.2, respectively, and consider the following splitting:

$$\begin{aligned}
\mathfrak{J}_2^\lambda(t) &= 2\lambda \int_{nT}^t \mathbb{E}\left[\left\langle \overline{\theta}_s^\lambda - \overline{\theta}_{\lfloor s\rfloor}^\lambda - (\overline{\zeta}_s^{\lambda,n} - \overline{\zeta}_{\lfloor s\rfloor}^{\lambda,n}), \sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^s \mathfrak{M}(\overline{\theta}_r^\lambda, \overline{\theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\right\rangle\right]\mathrm{d}s \\
&= 2\lambda \int_{nT}^t \mathbb{E}\left[\left\langle \int_{\lfloor s\rfloor}^s \left(\lambda(h(\overline{\zeta}_r^{\lambda,n}) - h_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda)) + \lambda^2\int_{\lfloor r\rfloor}^r \left(H_\lambda(\overline{\theta}_{\lfloor \nu\rfloor}^\lambda)h_\lambda(\overline{\theta}_{\lfloor \nu\rfloor}^\lambda) - \beta^{-1}\Upsilon_\lambda(\overline{\theta}_{\lfloor \nu\rfloor}^\lambda)\right)\mathrm{d}\nu\right.\right.\right. \\
&\qquad \left.\left.\left. -\lambda\sqrt{2\lambda\beta^{-1}}\int_{\lfloor r\rfloor}^r H_\lambda(\overline{\theta}_{\lfloor \nu\rfloor}^\lambda)\,\mathrm{d}B_\nu^\lambda\right)\mathrm{d}r, \sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^s \mathfrak{M}(\overline{\theta}_r^\lambda, \overline{\theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\right\rangle\right]\mathrm{d}s \\
&= 2\lambda \int_{nT}^t \mathbb{E}\left[\left\langle \int_{\lfloor s\rfloor}^s \left(\lambda(h(\overline{\theta}_r^{\lambda}) - h(\overline{\theta}_{\lfloor r\rfloor}^\lambda)) + \lambda^2\int_{\lfloor r\rfloor}^r \left(H_\lambda(\overline{\theta}_{\lfloor \nu\rfloor}^\lambda)h_\lambda(\overline{\theta}_{\lfloor \nu\rfloor}^\lambda) - \beta^{-1}\Upsilon_\lambda(\overline{\theta}_{\lfloor \nu\rfloor}^\lambda)\right)\mathrm{d}\nu\right.\right.\right. \\
&\qquad \left.\left.\left. -\lambda\sqrt{2\lambda\beta^{-1}}\int_{\lfloor r\rfloor}^r H_\lambda(\overline{\theta}_{\lfloor \nu\rfloor}^\lambda)\,\mathrm{d}B_\nu^\lambda\right)\mathrm{d}r, \sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^s \mathfrak{M}(\overline{\theta}_r^\lambda, \overline{\theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\right\rangle\right]\mathrm{d}s \\
&\quad + 2\lambda \int_{nT}^t \mathbb{E}\left[\left\langle \int_{\lfloor s\rfloor}^s \lambda(h(\overline{\theta}_{\lfloor r\rfloor}^\lambda) - h_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda))\,\mathrm{d}r, \sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^s \mathfrak{M}(\overline{\theta}_r^\lambda, \overline{\theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\right\rangle\right]\mathrm{d}s + \mathfrak{J}_{2,1}^\lambda(t),
\end{aligned}$$

where the first equality holds due to the following:

$$\begin{aligned}
&\mathbb{E}\left[\left\langle \overline{\theta}_{\lfloor s\rfloor}^\lambda + \overline{\zeta}_{\lfloor s\rfloor}^{\lambda,n}, \sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^s \mathfrak{M}(\overline{\theta}_r^\lambda, \overline{\theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\right\rangle\right] \\
&= \mathbb{E}\left[\left\langle \overline{\theta}_{\lfloor s\rfloor}^\lambda + \overline{\zeta}_{\lfloor s\rfloor}^{\lambda,n}, \mathbb{E}\left[\sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^s \mathfrak{M}(\overline{\theta}_r^\lambda, \overline{\theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\middle|\mathcal{F}_{\lfloor s\rfloor}^\lambda\right]\right\rangle\right] = 0,
\end{aligned} \tag{101}$$

and where

$$\mathfrak{J}_{2,1}^\lambda(t) = 2\lambda \int_{nT}^t \mathbb{E}\left[\left\langle \int_{\lfloor s\rfloor}^s \lambda(h(\overline{\zeta}_r^{\lambda,n}) - h(\overline{\theta}_r^\lambda))\,\mathrm{d}r, \sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^s \mathfrak{M}(\overline{\theta}_r^\lambda, \overline{\theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\right\rangle\right]\mathrm{d}s.$$

By applying Cauchy-Schwarz inequality, Corollary A.4, Lemma A.6 and (98), we further obtain that

$$\begin{aligned}
\mathfrak{J}_2^\lambda(t) &\leq 2\lambda \int_{nT}^t \left(\mathbb{E}\left[\left|\int_{\lfloor s\rfloor}^s \lambda\left(h(\overline{\theta}_r^\lambda) - h(\overline{\theta}_{\lfloor r\rfloor}^\lambda) + \lambda\int_{\lfloor r\rfloor}^r \left(H_\lambda(\overline{\theta}_{\lfloor \nu\rfloor}^\lambda)h_\lambda(\overline{\theta}_{\lfloor \nu\rfloor}^\lambda) - \beta^{-1}\Upsilon_\lambda(\overline{\theta}_{\lfloor \nu\rfloor}^\lambda)\right)\mathrm{d}\nu\right.\right.\right. \\
&\qquad \left.\left.\left. -\sqrt{2\lambda\beta^{-1}}\int_{\lfloor r\rfloor}^r H_\lambda(\overline{\theta}_{\lfloor \nu\rfloor}^\lambda)\,\mathrm{d}B_\nu^\lambda\right)\mathrm{d}r\right|^2\right]\right)^{1/2} \\
&\qquad \times \left(\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^s \mathfrak{M}(\overline{\theta}_r^\lambda, \overline{\theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\right|^2\right]\right)^{1/2}\mathrm{d}s \\
&\quad + 2\lambda \int_{nT}^t \left(\mathbb{E}\left[\left|\int_{\lfloor s\rfloor}^s \lambda(h(\overline{\theta}_{\lfloor r\rfloor}^\lambda) - h_\lambda(\overline{\theta}_{\lfloor r\rfloor}^\lambda))\,\mathrm{d}r\right|^2\right]\right)^{1/2} \\
&\qquad \times \left(\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^s \mathfrak{M}(\overline{\theta}_r^\lambda, \overline{\theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\right|^2\right]\right)^{1/2}\mathrm{d}s + \mathfrak{J}_{2,1}^\lambda(t) \\
&\leq 2\lambda \int_{nT}^t \left(\lambda^4\left(e^{-\mathsf{a}_\mathsf{D}\kappa n/2}36\overline{\mathsf{C}}_{\mathbf{A2}}\mathbb{E}[|\theta_0|^{16(\rho+1)}] + 36\widetilde{\mathsf{C}}_{\mathbf{A2}}\right)\right)^{1/2} \\
&\qquad \times \left(\lambda^2\left(e^{-\mathsf{a}_\mathsf{D}\kappa n/2}\overline{\mathsf{C}}_{\mathbf{A2}}\mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A2}}\right)\right)^{1/2}\mathrm{d}s \\
&\quad + 2\lambda \int_{nT}^t \left(\lambda^5\left(e^{-\mathsf{a}_\mathsf{D}\kappa n/2}\overline{\mathsf{C}}_{\mathbf{A2}}\mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A2}}\right)\right)^{1/2}
\end{aligned}$$

$$\times \left(\lambda^2 \left(e^{-\mathsf{a}_\mathsf{D}\kappa n/2}\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2}\right)\right)^{1/2} \mathrm{d}s + \mathfrak{J}^\lambda_{2,1}(t)$$
$$\leq 14\lambda^3 \left(e^{-\mathsf{a}_\mathsf{D}\kappa n/2}\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2}\right) + \mathfrak{J}^\lambda_{2,1}(t). \tag{102}$$

At this stage, the task reduces to upper bound $\mathfrak{J}^\lambda_{2,1}(t)$. To achieve this, we apply Cauchy-Schwarz inequality, Corollary A.4 and Lemma A.6 to obtain

$$\begin{aligned}
\mathfrak{J}^\lambda_{2,1}(t) &= 2\lambda \int_{nT}^t \mathbb{E}\Bigg[\Bigg\langle \int_{\lfloor s\rfloor}^s \lambda \Bigg(h(\overline{\zeta}^{\lambda,n}_r) - h(\overline{\zeta}^{\lambda,n}_{\lfloor r\rfloor}) - \sqrt{2\lambda\beta^{-1}}\int_{\lfloor r\rfloor}^r H(\overline{\zeta}^{\lambda,n}_{\lfloor \nu\rfloor})\,\mathrm{d}B^\lambda_\nu - \Big(h(\overline{\theta}^\lambda_r) \\
&\quad - h(\overline{\theta}^\lambda_{\lfloor r\rfloor}) - \sqrt{2\lambda\beta^{-1}}\int_{\lfloor r\rfloor}^r H(\overline{\theta}^\lambda_{\lfloor \nu\rfloor})\,\mathrm{d}B^\lambda_\nu\Big)\Bigg)\,\mathrm{d}r, \sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^s \mathfrak{M}(\overline{\theta}^\lambda_r, \overline{\theta}^\lambda_{\lfloor r\rfloor})\,\mathrm{d}B^\lambda_r\Bigg\rangle\Bigg]\,\mathrm{d}s \\
&\quad + 2\lambda \int_{nT}^t \mathbb{E}\Bigg[\Bigg\langle \int_{\lfloor s\rfloor}^s \lambda\sqrt{2\lambda\beta^{-1}}\int_{\lfloor r\rfloor}^r (H(\overline{\zeta}^{\lambda,n}_{\lfloor \nu\rfloor}) - H(\overline{\theta}^\lambda_{\lfloor \nu\rfloor}))\,\mathrm{d}B^\lambda_\nu\,\mathrm{d}r, \\
&\quad \sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^s \mathfrak{M}(\overline{\theta}^\lambda_r, \overline{\theta}^\lambda_{\lfloor r\rfloor})\,\mathrm{d}B^\lambda_r\Bigg\rangle\Bigg]\,\mathrm{d}s \\
&\leq 2\lambda \int_{nT}^t \Bigg(\mathbb{E}\Bigg[\Bigg|\int_{\lfloor s\rfloor}^s \lambda \Bigg(h(\overline{\zeta}^{\lambda,n}_r) - h(\overline{\zeta}^{\lambda,n}_{\lfloor r\rfloor}) - \sqrt{2\lambda\beta^{-1}}\int_{\lfloor r\rfloor}^r H(\overline{\zeta}^{\lambda,n}_{\lfloor \nu\rfloor})\,\mathrm{d}B^\lambda_\nu - \Big(h(\overline{\theta}^\lambda_r) \\
&\quad - h(\overline{\theta}^\lambda_{\lfloor r\rfloor}) - \sqrt{2\lambda\beta^{-1}}\int_{\lfloor r\rfloor}^r H(\overline{\theta}^\lambda_{\lfloor \nu\rfloor})\,\mathrm{d}B^\lambda_\nu\Big)\Bigg)\,\mathrm{d}r\Bigg|^2\Bigg]\Bigg)^{1/2} \\
&\quad \times \left(\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^s \mathfrak{M}(\overline{\theta}^\lambda_r, \overline{\theta}^\lambda_{\lfloor r\rfloor})\,\mathrm{d}B^\lambda_r\right|^2\right]\right)^{1/2} \mathrm{d}s \\
&\quad + 2\lambda^3 \left(e^{-\mathsf{a}_\mathsf{D}\min\{\kappa,1/2\}n/2}10\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}] + 10\widetilde{\mathsf{C}}_{\mathbf{A}2}\right) \\
&\leq 2\lambda \int_{nT}^t \left(\lambda^4 \left(e^{-\mathsf{a}_\mathsf{D}\min\{\kappa,1/2\}n/2}72\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}] + 72\widetilde{\mathsf{C}}_{\mathbf{A}2}\right)\right)^{1/2} \\
&\quad \times \left(\lambda^2 \left(e^{-\mathsf{a}_\mathsf{D}\kappa n/2}\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2}\right)\right)^{1/2} \mathrm{d}s \\
&\quad + 20\lambda^3 \left(e^{-\mathsf{a}_\mathsf{D}\min\{\kappa,1/2\}n/2}\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2}\right) \\
&\leq 44\lambda^3 \left(e^{-\mathsf{a}_\mathsf{D}\min\{\kappa,1/2\}n/2}\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2}\right).
\end{aligned} \tag{103}$$

Substituting (103) into (102) yields

$$\mathfrak{J}^\lambda_2(t) \leq 58\lambda^3 \left(e^{-\mathsf{a}_\mathsf{D}\min\{\kappa,1/2\}n/2}\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2}\right). \tag{104}$$

By applying (100) and (104) to (99), we obtain that

$$\begin{aligned}
\mathbb{E}\left[\left|\overline{\theta}^\lambda_t - \overline{\zeta}^{\lambda,n}_t\right|^2\right] &\leq \lambda(2\mathsf{L}_\mathsf{OS} + 7)\int_{nT}^t \mathbb{E}\left[|\overline{\theta}^\lambda_s - \overline{\zeta}^{\lambda,n}_s|^2\right]\mathrm{d}s \\
&\quad + 65\lambda^{2+q}\left(e^{-\mathsf{a}_\mathsf{D}\min\{\kappa,1/2\}n/2}\overline{\mathsf{C}}_{\mathbf{A}2}\mathbb{E}[|\theta_0|^{16(\rho+1)}] + \widetilde{\mathsf{C}}_{\mathbf{A}2}\right),
\end{aligned}$$

which, by applying Grönwall's lemma, yields

$$W_2^2(\mathcal{L}(\overline{\theta}^\lambda_t), \mathcal{L}(\overline{\zeta}^{\lambda,n}_t)) \leq \mathbb{E}\left[\left|\overline{\theta}^\lambda_t - \overline{\zeta}^{\lambda,n}_t\right|^2\right] \leq \lambda^{2+q}\left(e^{-\mathsf{a}_\mathsf{D}\min\{\kappa,1/2\}n/2}\mathsf{C}_0\mathbb{E}[|\theta_0|^{16(\rho+1)}] + \mathsf{C}_1\right),$$

where κ is given in (73) and

$$\mathsf{C}_0 := 65e^{2\mathsf{L}_\mathsf{OS}+7}\overline{\mathsf{C}}_{\mathbf{A}2}, \quad \mathsf{C}_1 := 65e^{2\mathsf{L}_\mathsf{OS}+7}\widetilde{\mathsf{C}}_{\mathbf{A}2} \tag{105}$$

with $\overline{\mathsf{C}}_{\mathbf{A}2}$, $\widetilde{\mathsf{C}}_{\mathbf{A}2}$ given in (88). □

APPENDIX B. ANALYTIC EXPRESSION OF CONSTANTS FOR aHOLA

CONSTANT		FULL EXPRESSION
Remark 2.2	K_0	$2^{1-q}\max\{L, \lvert\nabla^2 h^{(1)}(0)\rvert,\dots,\lvert\nabla^2 h^{(d)}(0)\rvert\}$
	K_1	$\max\{K_0, \lvert\nabla h^{(1)}(0)\rvert,\dots,\lvert\nabla h^{(d)}(0)\rvert\}$
	K_2	$\max\{K_1, \lvert h^{(1)}(0)\rvert,\dots,\lvert h^{(d)}(0)\rvert\}$
	$K_{3,d}$	$\max\{d^{3/2}L, \Upsilon(0)\}$
Remark 2.4	a_D	$a/2$
	b_D	$(a/2+b)R_D^{\overline{r}+2}+\lvert h(0)\rvert^2/(2a)$
	R_D	$\max\{(4b/a)^{1/(r-\overline{r})}, 2^{1/r}\}$
	$\overline{b}_D$	a_D+b_D
Remark 2.5	L_{OS}	$\sqrt{d}K_1(1+2R_{OS})^{\rho+q-1}$
	R_{OS}	$(b/a)^{1/(r-\overline{r})}$
Theorem 2.6	C_0	$\min\{\dot{c}, a_D\kappa, a_D/2\}/4$
	C_1	$e^{\min\{\dot{c},a_D\kappa,a_D/2\}/4}\left[\mathsf{C}_0^{1/2}+\mathsf{C}_2+\hat{c}\left(3+\int_{\mathbb{R}^d}V_2(\theta)\pi_\beta(\mathrm{d}\theta)\right)\right]$
	C_2	$\mathsf{C}_1^{1/2}+\mathsf{C}_3$
Corollary 2.7	C_3	$\min\{\dot{c}, a_D\kappa, a_D/2\}/8$
	C_4	$e^{\min\{\dot{c},a_D\kappa,a_D/2\}/8}\left[\mathsf{C}_0^{1/2}+\mathsf{C}_4+\sqrt{2\hat{c}}\left(3+\int_{\mathbb{R}^d}V_2(\theta)\pi_\beta(\mathrm{d}\theta)\right)^{1/2}\right]$
	C_5	$\mathsf{C}_1^{1/2}+\mathsf{C}_5$
	C_4	$\sqrt{\hat{c}}\left(1+\frac{8}{\min\{\dot{c},a_D\kappa,a_D/2\}}\right)e^{\min\{\dot{c},a_D\kappa,a_D/2\}/8}\left(\sqrt{2}\mathsf{C}_0^{1/2}+2^{4\rho+5}\right)$
	C_5	$4(\sqrt{\hat{c}}/\dot{c})e^{\dot{c}/4}(\sqrt{2}\mathsf{C}_1^{1/2}+3\sqrt{2}+2^{4\rho+5}(c_{8(\rho+1)}(1+1/(a_D\kappa))+1)^{1/2}+\sqrt{6}v_{16(\rho+1)}^{1/2}(M_V(16(\rho+1))))$
Lemma 6.3	κ	$M_1^{\rho+q-1}/\left(1+M_1^{3(\rho+q-1)}\right)^{1/3}$ (with $M_1>0$)
	c_0	$a_D\kappa M_1^2+c_1+2^{2(\rho+q)}\beta^{-1}d(1+K_1)^2$
	c_1	$2b_D+4K_hK_H+\beta^{-1}2^{\rho+q-1}K_{3,d}+2K_h^2+3K_H^2K_h^2+6K_h^2K_H+2^{2(\rho+q)-6}\beta^{-2}K_{3,d}^2+\beta^{-1}K_hK_{3,d}2^{\rho+q}+\beta^{-1}K_HK_hK_{3,d}2^{\rho+q-1}$
	c_p	$(1+2/(a_D\kappa))^{p-1}(a_D\kappa M_1^2+c_1)^p+c_\Xi(p)+2^{2(p+\rho+q)-3}p(2p-1)\beta^{-1}d(1+K_1)^2((1+2/(a_D\kappa))^{p-2}(a_D\kappa M_1^2+c_1)^{p-1}+M_2(p)^{2p-2})$
	$c_\Xi(p)$	$2^{2p-4}(\beta^{-1}d)^p(2p(2p-1))^{p+1}3^{p-1}\left(1+2^{p(\rho+q-1)}d^{p/2}K_1^p\right)^2$
	$M_2(p)$	$(2^{2(p+\rho+q)-1}p(2p-1)\beta^{-1}d(1+K_1)^2/(a_D\kappa))^{1/2}$
Lemma 6.4	$c_{V,1}(p)$	$a_Dp/4$
	$c_{V,2}(p)$	$(3/4)a_Dpv_p(M_V(p))$
	$M_V(p)$	$(1/3+4\overline{b}_D/(3a_D)+4d/(3a_D\beta)+4(p-2)/(3a_D\beta))^{1/2}$
Lemma 6.6	C_0	$65e^{2L_{OS}+7}\widetilde{C}_{\mathbf{A}2}$
	C_1	$65e^{2L_{OS}+7}\widetilde{C}_{\mathbf{A}2}$
	$\overline{C}_{\mathbf{A}2}$	$3^{24(\rho+1)}d^3(1+K_0)^2(1+K_h)^2(1+K_H)^4(1+\beta^{-1})^2(1+L)^2(1+K_{3,d})^2(2+\max\{\overline{C}_{\mathbf{A}0,2},\overline{C}_{\mathbf{A}0,2q},\overline{C}_{\mathbf{A}1,2},\overline{C}_{\mathbf{A}0,2+2q}\})$
	$\widetilde{C}_{\mathbf{A}2}$	$3^{24(\rho+1)}d^3(1+K_0)^2(1+K_h)^2(1+K_H)^4(1+\beta^{-1})^2(1+L)^2(1+K_{3,d})^2(2c_{8(\rho+1)}(1+1/(a_D\kappa))+2\max\{\widetilde{C}_{\mathbf{A}0,2},\widetilde{C}_{\mathbf{A}0,2q},\widetilde{C}_{\mathbf{A}1,2},\widetilde{C}_{\mathbf{A}0,2+2q}\}+1+v_{16(\rho+1)}(M_V(16(\rho+1))))$
	$\overline{C}_{\mathbf{A}0,p}$	$5^{2p}2^{4\lceil p\rceil(\rho+1)}(K_h^{2p}+\beta^{-2p}K_{3,d}^{2p}+K_H^{2p}K_h^{2p}+(\beta^{-1}(4p+2)(d+4p))^p(1+K_H^{2p}))$
	$\widetilde{C}_{\mathbf{A}0,p}$	$\overline{C}_{\mathbf{A}0,p}\left(c_{2\lceil p\rceil(\rho+1)}(1+1/(a_D\kappa))+1\right)$
	$\overline{C}_{\mathbf{A}1,p}$	$2^{2p+\lceil p\rceil(\rho+1)-1}K_h^{2p}$
	$\widetilde{C}_{\mathbf{A}1,p}$	$\overline{C}_{\mathbf{A}1,p}\left(c_{\lceil p\rceil(\rho+1)}(1+1/(a_D\kappa))+1\right)+2^{2p}K_h^{2p}3v_{2\lceil p\rceil(\rho+1)}(M_V(2\lceil p\rceil(\rho+1)))+2^{2p}(2\beta^{-1}(p+1)(d+2p))^p$
Proposition 6.7	$\dot{c}$	$\min\{\bar{\phi}, c_{V,1}(2), 4c_{V,2}(2)\epsilon c_{V,1}(2)\}/2$
	$\bar{\phi}$	$\left(\sqrt{8\pi/(\beta L_{OS})}\dot{c}_0\exp\left(\left(\dot{c}_0\sqrt{\beta L_{OS}/8}+\sqrt{8/(\beta L_{OS})}\right)^2\right)\right)^{-1}$
	ϵ	$1\wedge(4c_{V,2}(2)\sqrt{2\beta\pi/L_{OS}}\int_0^{\dot{c}_1}\exp\left(\left(s\sqrt{\beta L_{OS}/8}+\sqrt{8/(\beta L_{OS})}\right)^2\right)\mathrm{d}s)^{-1}$
	$\dot{c}_0$	$2(4c_{V,2}(2)(1+c_{V,1}(2))/c_{V,1}(2)-1)^{1/2}$
	$\dot{c}_1$	$2(2c_{V,2}(2)/c_{V,1}(2)-1)^{1/2}$
	$\hat{c}$	$2(1+\dot{c}_0)\exp(\beta L_{OS}\dot{c}_0^2/8+2\dot{c}_0)/\epsilon$
Lemma 6.8	C_2	$\hat{c}\left(1+\frac{4}{\min\{\dot{c},a_D\kappa,a_D/2\}}\right)e^{\min\{\dot{c},a_D\kappa,a_D/2\}/4}\left(\mathsf{C}_0+2^{8\rho+11}\right)$
	C_3	$2(\hat{c}/\dot{c})e^{\dot{c}/2}(\mathsf{C}_1+15+2^{8\rho+11}(c_{8(\rho+1)}(1+1/(a_D\kappa))+1)+18v_{16(\rho+1)}(M_V(16(\rho+1))))$

APPENDIX C. ANALYTIC EXPRESSION OF CONSTANTS FOR aHOLLA

CONSTANT		FULL EXPRESSION
Remark 3.2	$\overline{\mathsf{K}}_0$	$\max\{\overline{L}_1, \lvert\nabla^2 h^{(1)}(0)\rvert, \ldots, \lvert\nabla^2 h^{(d)}(0)\rvert\}$
	$\overline{\mathsf{K}}_1$	$\max\{\overline{L}_3, \lvert h(0)\rvert\}$
Theorem 3.3	$C_{\mathsf{Lin},0}$	$\min\{\dot{\mathsf{c}}_{\mathsf{Lin}}, \overline{a}/2\}/4$
	$C_{\mathsf{Lin},1}$	$e^{\min\{\dot{\mathsf{c}}_{\mathsf{Lin}}, \overline{a}/2\}/4} \left[\mathsf{C}_{\mathsf{Lin},0}^{1/2} + \mathsf{C}_{\mathsf{Lin},2} + \hat{\mathsf{c}}_{\mathsf{Lin}} \left(3 + \int_{\mathbb{R}^d} V_2(\theta)\pi_\beta(\mathrm{d}\theta)\right)\right]$
	$C_{\mathsf{Lin},2}$	$\mathsf{C}_{\mathsf{Lin},1}^{1/2} + \mathsf{C}_{\mathsf{Lin},3}$
Corollary 3.4	$C_{\mathsf{Lin},3}$	$\min\{\dot{\mathsf{c}}_{\mathsf{Lin}}, \overline{a}/2\}/8$
	$C_{\mathsf{Lin},4}$	$e^{\min\{\dot{\mathsf{c}}_{\mathsf{Lin}}, \overline{a}/2\}/8} \left[\mathsf{C}_{\mathsf{Lin},0}^{1/2} + \mathsf{C}_{\mathsf{Lin},4} + \sqrt{2\hat{\mathsf{c}}_{\mathsf{Lin}}} \left(3 + \int_{\mathbb{R}^d} V_2(\theta)\pi_\beta(\mathrm{d}\theta)\right)^{1/2}\right]$
	$C_{\mathsf{Lin},5}$	$\mathsf{C}_{\mathsf{Lin},1}^{1/2} + \mathsf{C}_{\mathsf{Lin},5}$
	$\mathsf{C}_{\mathsf{Lin},4}$	$\sqrt{\hat{\mathsf{c}}_{\mathsf{Lin}}} \left(1 + \frac{8}{\min\{\dot{\mathsf{c}}_{\mathsf{Lin}}, \overline{a}/2\}}\right) e^{\min\{\dot{\mathsf{c}}_{\mathsf{Lin}}, \overline{a}/2\}/8} \left(\mathsf{C}_{\mathsf{Lin},0}^{1/2} + 3\right)$
	$\mathsf{C}_{\mathsf{Lin},5}$	$4(\sqrt{\hat{\mathsf{c}}_{\mathsf{Lin}}}/\dot{\mathsf{c}}_{\mathsf{Lin}}) e^{\dot{\mathsf{c}}_{\mathsf{Lin}}/4} \left(\mathsf{C}_{\mathsf{Lin},1}^{1/2} + 1 + 3(\overline{\mathsf{c}}_2(1 + 1/\overline{a}) + 1)^{1/2} + \sqrt{3}\mathrm{v}_4^{1/2}(\overline{\mathsf{M}}_V(4))\right)$
Lemma D.3	$\overline{\mathsf{c}}_0$	$\beta^{-1} d \left(\overline{L}_2 + \overline{L}_2\overline{\mathsf{K}}_1 + \overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2\right) \overline{\mathsf{M}}_1 + \overline{\mathsf{c}}_1 + 2\beta^{-1} d(1 + \overline{L}_3)^2$
	$\overline{\mathsf{c}}_1$	$2\overline{b} + 2\overline{L}_3\overline{\mathsf{K}}_1 + 2\overline{\mathsf{K}}_1^2 + \overline{L}_3^2\overline{\mathsf{K}}_1^2/2 + \beta^{-2}d^2\overline{L}_2^2/4 + 2\overline{L}_3\overline{\mathsf{K}}_1^2 + \beta^{-1}d(\overline{L}_2\overline{\mathsf{K}}_1 + \overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2)$
	$\overline{\mathsf{M}}_1$	$2(\overline{a}\beta)^{-1} d(\overline{L}_2 + \overline{L}_2\overline{\mathsf{K}}_1 + \overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2)$
	$\overline{\mathsf{c}}_p$	$\overline{\mathsf{c}}_{\Xi}(p) + (1 + 2/\overline{a})^{p-1} \left(\overline{\mathsf{c}}_1 + \beta^{-1} d \left(\overline{L}_2 + \overline{L}_2\overline{\mathsf{K}}_1 + \overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2\right) \overline{\mathsf{M}}_1\right)^p + 2^{2p-2}p(2p - 1)\beta^{-1}d(1 + \overline{L}_3)^2((1 + 2/\overline{a})^{p-2} \left(\overline{\mathsf{c}}_1 + \beta^{-1} d \left(\overline{L}_2 + \overline{L}_2\overline{\mathsf{K}}_1 + \overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2\right) \overline{\mathsf{M}}_1\right)^{p-1} + \overline{\mathsf{M}}_2(p)^{2p-2})$
	$\overline{\mathsf{c}}_{\Xi}(p)$	$2^{2p-3}p(2p - 1)(2\beta^{-1}dp(2p - 1)(1 + \overline{L}_3)^2)^p$
	$\overline{\mathsf{M}}_2(p)$	$(2^{2p}p(2p - 1)\beta^{-1}d(1 + \overline{L}_3)^2/\overline{a})^{1/2}$
Lemma D.4	$\overline{\mathsf{c}}_{V,1}(p)$	$\overline{a}p/4$
	$\overline{\mathsf{c}}_{V,2}(p)$	$(3/4)\overline{a}p\mathrm{v}_p(\overline{\mathsf{M}}_V(p))$
	$\overline{\mathsf{M}}_V(p)$	$(1/3 + 4\overline{b}/(3\overline{a}) + 4d/(3\overline{a}\beta) + 4(p - 2)/(3\overline{a}\beta))^{1/2}$
Lemma D.6	$\mathsf{C}_{\mathsf{Lin},0}$	$52e^{2\overline{L}_3+6}\overline{\mathsf{C}}_{\mathbf{s}2}$
	$\mathsf{C}_{\mathsf{Lin},1}$	$52e^{2\overline{L}_3+6}\widetilde{\mathsf{C}}_{\mathbf{s}2}$
	$\overline{\mathsf{C}}_{\mathbf{s}2}$	$2^7 d^4(1 + \beta^{-1})^2(1 + \overline{L}_3)^4(1 + \overline{L}_2)^2(1 + \overline{L}_1)^2(1 + \overline{\mathsf{K}}_1)^2(1 + \overline{\mathsf{K}}_0)^2(1 + \max\{\overline{\mathsf{C}}_{\mathbf{s}0,2}, \overline{\mathsf{C}}_{\mathbf{s}0,2q}, \overline{\mathsf{C}}_{\mathbf{s}0,1+q}, \overline{\mathsf{C}}_{\mathbf{s}1,2}\})$
	$\widetilde{\mathsf{C}}_{\mathbf{s}2}$	$2^7 d^4(1 + \beta^{-1})^2(1 + \overline{L}_3)^4(1 + \overline{L}_2)^2(1 + \overline{L}_1)^2(1 + \overline{\mathsf{K}}_1)^2(1 + \overline{\mathsf{K}}_0)^2(\overline{\mathsf{c}}_2(1 + 1/\overline{a}) + \max\{\widetilde{\mathsf{C}}_{\mathbf{s}0,2}, \widetilde{\mathsf{C}}_{\mathbf{s}0,2q}, \widetilde{\mathsf{C}}_{\mathbf{s}0,1+q}, \widetilde{\mathsf{C}}_{\mathbf{s}1,2}\} + 1)$
	$\overline{\mathsf{C}}_{\mathbf{s}0,p}$	$10^{2p}(\overline{\mathsf{K}}_1^{2p} + (\overline{L}_3\overline{\mathsf{K}}_1)^{2p})$
	$\widetilde{\mathsf{C}}_{\mathbf{s}0,p}$	$\overline{\mathsf{C}}_{\mathbf{s}0,p}(\overline{\mathsf{c}}_{\lceil p\rceil}(1 + 1/\overline{a}) + 1) + 5^{2p}((\beta^{-1}d\overline{L}_2)^{2p} + (2\beta^{-1}(p + 1)(d + 2p))^p(1 + \overline{L}_3^{2p}))$
	$\overline{\mathsf{C}}_{\mathbf{s}1,p}$	$2^{4\lceil p\rceil}(1 + \overline{\mathsf{K}}_1)^{2p}$
	$\widetilde{\mathsf{C}}_{\mathbf{s}1,p}$	$\overline{\mathsf{C}}_{\mathbf{s}1,p} \left(\overline{\mathsf{c}}_{\lceil p\rceil}(1 + 1/\overline{a}) + 1\right) + 2^{3\lceil p\rceil}(1 + \overline{\mathsf{K}}_1)^{2p} 3\mathrm{v}_{2\lceil p\rceil}(\overline{\mathsf{M}}_V(2\lceil p\rceil)) + 2^{2p}(2\beta^{-1}(p + 1)(d + 2p))^p$
Proposition D.7	$\dot{\mathsf{c}}_{\mathsf{Lin}}$	$\min\{\bar{\phi}_{\mathsf{Lin}}, \overline{\mathsf{c}}_{V,1}(2), 4\overline{\mathsf{c}}_{V,2}(2)\overline{\epsilon}\overline{\mathsf{c}}_{V,1}(2)\}/2$
	$\bar{\phi}_{\mathsf{Lin}}$	$\left(\sqrt{8\pi/(\beta\overline{L}_3)}\dot{\mathsf{c}}_{\mathsf{Lin},0} \exp\left(\left(\dot{\mathsf{c}}_{\mathsf{Lin},0}\sqrt{\beta\overline{L}_3/8} + \sqrt{8/(\beta\overline{L}_3)}\right)^2\right)\right)^{-1}$
	$\overline{\epsilon}$	$1 \wedge \left(4\overline{\mathsf{c}}_{V,2}(2)\sqrt{2\beta\pi/\overline{L}_3} \int_0^{\dot{\mathsf{c}}_{\mathsf{Lin},1}} \exp\left(\left(s\sqrt{\beta\overline{L}_3/8} + \sqrt{8/(\beta\overline{L}_3)}\right)^2\right) \mathrm{d}s\right)^{-1}$
	$\dot{\mathsf{c}}_{\mathsf{Lin},0}$	$2(4\overline{\mathsf{c}}_{V,2}(2)(1 + \overline{\mathsf{c}}_{V,1}(2))/\overline{\mathsf{c}}_{V,1}(2) - 1)^{1/2}$
	$\dot{\mathsf{c}}_{\mathsf{Lin},1}$	$2(2\overline{\mathsf{c}}_{V,2}(2)/\overline{\mathsf{c}}_{V,1}(2) - 1)^{1/2}$
	$\hat{\mathsf{c}}_{\mathsf{Lin}}$	$2(1 + \dot{\mathsf{c}}_{\mathsf{Lin},0}) \exp(\beta\overline{L}_3\dot{\mathsf{c}}_{\mathsf{Lin},0}^2/8 + 2\dot{\mathsf{c}}_{\mathsf{Lin},0})/\overline{\epsilon}$
Lemma D.8	$\mathsf{C}_{\mathsf{Lin},2}$	$\hat{\mathsf{c}}_{\mathsf{Lin}} \left(1 + \frac{4}{\min\{\dot{\mathsf{c}}_{\mathsf{Lin}}, \overline{a}/2\}}\right) e^{\min\{\dot{\mathsf{c}}_{\mathsf{Lin}}, \overline{a}/2\}/4} (\mathsf{C}_{\mathsf{Lin},0} + 9)$
	$\mathsf{C}_{\mathsf{Lin},3}$	$2(\hat{\mathsf{c}}_{\mathsf{Lin}}/\dot{\mathsf{c}}_{\mathsf{Lin}}) e^{\dot{\mathsf{c}}_{\mathsf{Lin}}/2} \left(\mathsf{C}_{\mathsf{Lin},1} + 9 + 9\overline{\mathsf{c}}_2(1 + 1/\overline{a}) + 9\mathrm{v}_4(\overline{\mathsf{M}}_V(4))\right)$

APPENDIX D. PROOF OF MAIN RESULTS FOR aHOLLA

We provide the proofs for Theorem 3.3 and Corollary 3.4, which follow the same lines as those for aHOLA. We first introduce auxiliary processes which we use throughout the convergence analysis in Appendix D.1. Then, we provide moment estimates for the newly introduced processes in Appendix D.2, which are followed by the detailed proofs for the main results in Appendix D.3. We postpone the proofs of the results in Appendices D.2 and D.3 to Appendix D.4.

D.1. Auxiliary processes.

Fix $\beta > 0$. Consider the Langevin SDE $(Z_t)_{t\geq 0}$ given by

$$Z_0 := \theta_0, \quad \mathrm{d}Z_t = -h\left(Z_t\right)\mathrm{d}t + \sqrt{2\beta^{-1}}\mathrm{d}B_t, \tag{106}$$

where $(B_t)_{t\geq 0}$ is a d-dimensional Brownian motion on $(\Omega, \mathcal{F}, P)$ with its completed natural filtration denoted by $(\mathcal{F}_t)_{t\geq 0}$. Moreover, we assume that $(\mathcal{F}_t)_{t\geq 0}$ is independent of $\sigma(\theta_0)$. Under Assumption 5, it is a well-known result that the Langevin SDE (106) admits a unique solution, which is adapted to $\mathcal{F}_t \vee \sigma(\theta_0)$, $t \geq 0$. Its $2p$-th moment estimate with $p \in \mathbb{N}$ is provided in Lemma D.9, which can be used to further deduce the $2p$-th moment estimate of $\pi_\beta(\theta) \propto e^{-\beta U(\theta)}$.

For each $\lambda > 0$, recall $B_t^\lambda := B_{\lambda t}/\sqrt{\lambda}, t \geq 0$. Denote by $(\mathcal{F}_t^\lambda)_{t\geq 0}$ with $\mathcal{F}_t^\lambda := \mathcal{F}_{\lambda t}$, $t \geq 0$, its completed natural filtration, which is independent of $\sigma(\theta_0)$. Moreover, we denote by $Z_t^\lambda := Z_{\lambda t}$, $t \geq 0$, the time-changed version of Langevin SDE (106), which is given by

$$Z_0^\lambda := \theta_0, \quad \mathrm{d}Z_t^\lambda = -\lambda h(Z_t^\lambda)\,\mathrm{d}t + \sqrt{2\lambda\beta^{-1}}\,\mathrm{d}B_t^\lambda. \tag{107}$$

Furthermore, we denote by $(\widetilde{\Theta}_t^\lambda)_{t\geq 0}$ the continuous-time interpolation of aHOLLA (15)-(17) given by

$$\widetilde{\Theta}_0^\lambda := \theta_0, \quad \mathrm{d}\widetilde{\Theta}_t^\lambda = \lambda\phi_{\mathsf{Lin}}^\lambda(\widetilde{\Theta}_{\lfloor t\rfloor}^\lambda)\,\mathrm{d}t + \sqrt{2\lambda\beta^{-1}}\psi_{\mathsf{Lin}}^\lambda(\widetilde{\Theta}_{\lfloor t\rfloor}^\lambda)\,\mathrm{d}B_t^\lambda, \tag{108}$$

where $\phi_{\mathsf{Lin}}^\lambda$ and $\psi_{\mathsf{Lin}}^\lambda$ are defined in (16) and (17), respectively.

Remark D.1. *Similarly, denote by $(\overline{\Theta}_t^\lambda)_{t\geq 0}$ the continuous-time interpolation of the order 1.5 scheme (18) given by*

$$\begin{aligned}\mathrm{d}\overline{\Theta}_t^\lambda = &-\lambda h(\overline{\Theta}_{\lfloor t\rfloor}^\lambda)\,\mathrm{d}t + \lambda^2\int_{\lfloor t\rfloor}^t \left(H(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)h(\overline{\Theta}_{\lfloor s\rfloor}^\lambda) - \beta^{-1}\Upsilon(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\right)\mathrm{d}s\,\mathrm{d}t \\ &- \lambda\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t H(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\,\mathrm{d}B_s^\lambda\,\mathrm{d}t + \sqrt{2\lambda\beta^{-1}}\,\mathrm{d}B_t^\lambda\end{aligned} \tag{109}$$

with $\overline{\Theta}_0^\lambda := \theta_0$. We note that $\mathcal{L}(\widetilde{\Theta}_n^\lambda) = \mathcal{L}(\Theta_n^{\mathsf{aHOLLA}}) = \mathcal{L}(\Theta_n^\lambda) = \mathcal{L}(\overline{\Theta}_n^\lambda)$, for each $n \in \mathbb{N}_0$.

Finally, for any $s \geq 0$, consider a continuous-time process $(\zeta_t^{s,v,\lambda})_{t\geq s}$ defined by

$$\zeta_s^{s,v,\lambda} := v \in \mathbb{R}^d, \quad \mathrm{d}\zeta_t^{s,v,\lambda} = -\lambda h(\zeta_t^{s,v,\lambda})\,\mathrm{d}t + \sqrt{2\lambda\beta^{-1}}\,\mathrm{d}B_t^\lambda. \tag{110}$$

Definition D.2. *Fix $\lambda > 0$. Define $T \equiv T(\lambda) := \lfloor 1/\lambda\rfloor$. Then, for any $n \in \mathbb{N}_0$ and $t \geq nT$, define*

$$\widetilde{\zeta}_t^{\lambda,n} := \zeta_t^{nT,\overline{\Theta}_{nT}^\lambda,\lambda}.$$

D.2. Moment estimates.

Recall the following Lyapunov functions: for each $p \in [2, \infty) \cap \mathbb{N}$, define $V_p(\theta) := (1 + |\theta|^2)^{p/2}$, for all $\theta \in \mathbb{R}^d$, and moreover, define $\mathrm{v}_p(w) := (1 + w^2)^{p/2}$, for all $w \geq 0$. We observe that V_p is twice continuously differentiable and satisfies:

$$\sup_{\theta\in\mathbb{R}^d} |\nabla V_p(\theta)|/V_p(\theta) < \infty, \quad \lim_{|\theta|\to\infty} \nabla V_p(\theta)/V_p(\theta) = 0. \tag{111}$$

Furthermore, we denote by $\mathcal{P}_{V_p}(\mathbb{R}^d)$ the set of probability measures $\mu \in \mathcal{P}(\mathbb{R}^d)$ which satisfies $\int_{\mathbb{R}^d} V_p(\theta)\,\mu(\mathrm{d}\theta) < \infty$.

Next, we establish moment estimates for $(\widetilde{\Theta}_t^\lambda)_{t\geq 0}$ given in (108). The results with explicit constants are provided below. We note that for any $p \in [2, \infty) \cap \mathbb{N}$ and $t \geq 0$, we have that $\mathbb{E}[|\widetilde{\Theta}_t^\lambda|^{2p}] = \mathbb{E}[|\overline{\Theta}_t^\lambda|^{2p}]$.

Lemma D.3. *Let Assumptions 4, 5, and 6 hold. Then, we obtain the following estimates:*

(i) *For any* $0 < \lambda \leq \overline{\lambda}_{\max}$, $n \in \mathbb{N}_0$, *and* $t \in (n, n+1]$,

$$\mathbb{E}\left[|\widetilde{\Theta}_t^\lambda|^2\right] \leq (1-\lambda(t-n)\overline{a})(1-\lambda\overline{a})^n\,\mathbb{E}\left[|\theta_0|^2\right] + \overline{\mathsf{c}}_0\,(1+1/\overline{a}),$$

where the constant $\overline{\mathsf{c}}_0$ *is given explicitly in* (137). *In particular, the above inequality implies* $\sup_{t\geq 0}\mathbb{E}\left[|\widetilde{\Theta}_t^\lambda|^2\right] \leq \mathbb{E}\left[|\theta_0|^2\right] + \overline{\mathsf{c}}_0\,(1+1/\overline{a}) < \infty$.

(ii) *For any* $p \in [2,\infty) \cap \mathbb{N}$, $0 < \lambda \leq \overline{\lambda}_{\max}$, $n \in \mathbb{N}_0$, *and* $t \in (n, n+1]$,

$$\mathbb{E}\left[|\widetilde{\Theta}_t^\lambda|^{2p}\right] \leq (1-\lambda(t-n)\overline{a})(1-\lambda\overline{a})^n\,\mathbb{E}\left[|\theta_0|^{2p}\right] + \overline{\mathsf{c}}_p\,(1+1/\overline{a}),$$

where the constant $\overline{\mathsf{c}}_p$ *is given explicitly in* (145). *In particular, the above estimate implies* $\sup_{t\geq 0}\mathbb{E}\left[|\widetilde{\Theta}_t^\lambda|^{2p}\right] \leq \mathbb{E}\left[|\theta_0|^{2p}\right] + \overline{\mathsf{c}}_p\,(1+1/\overline{a}) < \infty$.

Proof. See Appendix D.4. □

We provide below a drift condition for V_p (defined in the beginning of Appendix D.2).

Lemma D.4. *Let Assumptions 4, 5, and 6 hold. Then, for any* $p \in [2,\infty) \cap \mathbb{N}$, $\theta \in \mathbb{R}^d$, *we obtain*

$$\Delta V_p(\theta)/\beta - \langle \nabla V_p(\theta), h(\theta)\rangle \leq -\overline{\mathsf{c}}_{V,1}(p)V_p(\theta) + \overline{\mathsf{c}}_{V,2}(p),$$

where $\overline{\mathsf{c}}_{V,1}(p) := \overline{a}p/4$, $\overline{\mathsf{c}}_{V,2}(p) := (3/4)\overline{a}p\mathrm{v}_p(\overline{\mathsf{M}}_V(p))$ *with* $\overline{\mathsf{M}}_V(p) := (1/3 + 4\overline{b}/(3\overline{a}) + 4d/(3\overline{a}\beta) + 4(p-2)/(3\overline{a}\beta))^{1/2}$.

Proof. See [4, Lemma 3.5]. □

By applying Lemma D.3 and D.4, we obtain moment estimates for $(\widetilde{\zeta}_t^{\lambda,n})_{t\geq nT}$ defined in Definition D.2.

Lemma D.5. *Let Assumptions 4, 5, and 6 hold. Then, for any* $p \in \mathbb{N}$, $0 < \lambda \leq \overline{\lambda}_{\max}$, $n \in \mathbb{N}_0$, *and* $t \geq nT$, *we obtain*

$$\mathbb{E}[V_{2p}(\widetilde{\zeta}_t^{\lambda,n})] \leq 2^{p-1}e^{-\lambda\overline{a}t/2}\mathbb{E}[|\theta_0|^{2p}] + 2^{p-1}\left(\overline{\mathsf{c}}_p(1+1/\overline{a})+1\right) + 3\mathrm{v}_{2p}(\overline{\mathsf{M}}_V(2p)),$$

where $\overline{\mathsf{c}}_p$ *is given in* (145) *(see also Lemma D.3) and* $\overline{\mathsf{M}}_V(2p)$ *is given in Lemma D.4.*

Proof. See [22, Corollary 4.6]. □

D.3. **Proof of main results.** In this section, we present key results used to obtain Theorem 3.3. To this end, we split $W_1(\mathcal{L}(\overline{\Theta}_t^\lambda), \pi_\beta)$, for any $t \in (nT, (n+1)T]$ and $n \in \mathbb{N}_0$, by using the law of $\widetilde{\zeta}_t^{\lambda,n}$ defined in Definition D.2 and the law of Z_t^λ given in (107) as follows:

$$W_1(\mathcal{L}(\overline{\Theta}_t^\lambda), \pi_\beta) \leq W_1(\mathcal{L}(\overline{\Theta}_t^\lambda), \mathcal{L}(\widetilde{\zeta}_t^{\lambda,n})) + W_1(\mathcal{L}(\widetilde{\zeta}_t^{\lambda,n}), \mathcal{L}(Z_t^\lambda)) + W_1(\mathcal{L}(Z_t^\lambda), \pi_\beta). \tag{112}$$

In the following lemma, we provide a non-asymptotic estimate for $W_2(\mathcal{L}(\overline{\Theta}_t^\lambda), \mathcal{L}(\widetilde{\zeta}_t^{\lambda,n}))$, which can be used to upper bound the first term on the RHS of (112).

Lemma D.6. *Let Assumptions 4, 5, and 6 hold. Then, for any* $0 < \lambda \leq \overline{\lambda}_{\max}$, $n \in \mathbb{N}_0$, *and* $t \in (nT, (n+1)T]$, *we obtain*

$$W_2(\mathcal{L}(\overline{\Theta}_t^\lambda), \mathcal{L}(\widetilde{\zeta}_t^{\lambda,n})) \leq \lambda^{1+q/2}\left(e^{-\overline{a}n/4}\mathsf{C}_{\mathsf{Lin},0}\mathbb{E}[|\theta_0|^4] + \mathsf{C}_{\mathsf{Lin},1}\right)^{1/2},$$

where $\mathsf{C}_{\mathsf{Lin},0}, \mathsf{C}_{\mathsf{Lin},1}$ *are given explicitly in* (167).

Proof. See Appendix D.4. □

For the last two terms on the RHS of (112), we observe that they can be viewed as Wasserstein-1 distances between distributions of Langevin processes starting from different initial points. Therefore, to obtain their upper bounds, we introduce a semi-metric which allows us to establish a contraction result for the Langevin SDE (106) under our assumptions.

We consider the following semi-metric: for any $p \in [2,\infty) \cap \mathbb{N}$, $\mu, \nu \in \mathcal{P}_{V_p}(\mathbb{R}^d)$, let

$$w_{1,p}(\mu,\nu) := \inf_{\zeta\in\mathcal{C}(\mu,\nu)} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} [1 \wedge |\theta-\theta'|](1+V_p(\theta)+V_p(\theta'))\zeta(\mathrm{d}\theta,\mathrm{d}\theta'). \tag{113}$$

Then, we provide a result which states the contraction property of the Langevin SDE (106) in $w_{1,2}$.

Proposition D.7. *Let Assumptions 4, 5, and 6 hold. Moreover, let $\theta_0' \in L^2$, and let $(Z_t')_{t\geq 0}$ be the solution of SDE (106) whose starting point $Z_0' := \theta_0'$ is assumed to be independent of $\mathcal{F}_\infty := \sigma(\bigcup_{t\geq 0}\mathcal{F}_t)$. Then, we obtain*

$$w_{1,2}(\mathcal{L}(Z_t), \mathcal{L}(Z_t')) \leq \hat{\mathsf{c}}_{\mathsf{Lin}} e^{-\dot{\mathsf{c}}_{\mathsf{Lin}} t} w_{1,2}(\mathcal{L}(\theta_0), \mathcal{L}(\theta_0')), \tag{114}$$

where the explicit expressions for $\dot{\mathsf{c}}_{\mathsf{Lin}}, \hat{\mathsf{c}}_{\mathsf{Lin}}$ are given below.

The contraction constant $\dot{\mathsf{c}}_{\mathsf{Lin}}$ is given by:

$$\dot{\mathsf{c}}_{\mathsf{Lin}} := \min\{\bar{\phi}_{\mathsf{Lin}}, \bar{\mathsf{c}}_{V,1}(2), 4\bar{\mathsf{c}}_{V,2}(2)\bar{\epsilon}\bar{\mathsf{c}}_{V,1}(2)\}/2,$$

where $\bar{\mathsf{c}}_{V,1}(2) := \bar{a}/2$, $\bar{\mathsf{c}}_{V,2}(2) := 3\bar{a}\mathrm{v}_2(\overline{\mathsf{M}}_V(2))/2$ with $\overline{\mathsf{M}}_V(2) := (1/3 + 4\bar{b}/(3\bar{a}) + 4d/(3\bar{a}\beta))^{1/2}$, the constant $\bar{\phi}_{\mathsf{Lin}}$ is given by

$$\bar{\phi}_{\mathsf{Lin}} := \left(\sqrt{8\pi/(\beta\overline{L}_3)}\,\dot{\mathsf{c}}_{\mathsf{Lin},0} \exp\left(\left(\dot{\mathsf{c}}_{\mathsf{Lin},0}\sqrt{\beta\overline{L}_3/8} + \sqrt{8/(\beta\overline{L}_3)}\right)^2\right)\right)^{-1},$$

and $\bar{\epsilon} > 0$ is chosen such that

$$\bar{\epsilon} \leq 1 \wedge \left(4\bar{\mathsf{c}}_{V,2}(2)\sqrt{2\beta\pi/\overline{L}_3}\int_0^{\dot{\mathsf{c}}_{\mathsf{Lin},1}} \exp\left(\left(s\sqrt{\beta\overline{L}_3/8} + \sqrt{8/(\beta\overline{L}_3)}\right)^2\right) \mathrm{d}s\right)^{-1}$$

with $\dot{\mathsf{c}}_{\mathsf{Lin},0} := 2(4\bar{\mathsf{c}}_{V,2}(2)(1+\bar{\mathsf{c}}_{V,1}(2))/\bar{\mathsf{c}}_{V,1}(2) - 1)^{1/2}$ and $\dot{\mathsf{c}}_{\mathsf{Lin},1} := 2(2\bar{\mathsf{c}}_{V,2}(2)/\bar{\mathsf{c}}_{V,1}(2) - 1)^{1/2}$.

Moreover, the constant $\hat{\mathsf{c}}_{\mathsf{Lin}}$ is given by:

$$\hat{\mathsf{c}}_{\mathsf{Lin}} := 2(1+\dot{\mathsf{c}}_{\mathsf{Lin},0}) \exp(\beta\overline{L}_3\dot{\mathsf{c}}_{\mathsf{Lin},0}^2/8 + 2\dot{\mathsf{c}}_{\mathsf{Lin},0})/\bar{\epsilon}.$$

Proof. We note that [11, Assumption 2.1] holds with $\kappa = \overline{L}_3$ due to Assumption 5, [11, Assumption 2.2] holds with $V = V_2$ due to Remark D.4, and [11, Assumptions 2.4 and 2.5] hold due to (111). Therefore, we can obtain (114) following the same arguments as in the proof of [4, Proposition 3.14] based on [11, Theorem 2.2, Corollary 2.3]. In addition, $\dot{\mathsf{c}}_{\mathsf{Lin}}, \hat{\mathsf{c}}_{\mathsf{Lin}}$ can be obtained following the arguments in the proof of [18, Proposition 4.6]. □

By using the above result and $W_1 \leq w_{1,2}$ (see [18, Lemma A.3]), we can establish a non-asymptotic error bound for the second term on the RHS of (112). The explicit statement is given below.

Lemma D.8. *Let Assumptions 4, 5, and 6 hold. Then, for any $0 < \lambda \leq \overline{\lambda}_{\max}$, $n \in \mathbb{N}_0$, and $t \in (nT, (n+1)T]$, we obtain*

$$W_1(\mathcal{L}(\widetilde{\zeta}_t^{\lambda,n}), \mathcal{L}(Z_t^\lambda)) \leq \lambda^{1+q/2}\left(e^{-\min\{\dot{\mathsf{c}}_{\mathsf{Lin}},\bar{a}/2\}n/4}\mathsf{C}_{\mathsf{Lin},2}\mathbb{E}[|\theta_0|^4] + \mathsf{C}_{\mathsf{Lin},3}\right),$$

where

$$\begin{aligned}\mathsf{C}_{\mathsf{Lin},2} &:= \hat{\mathsf{c}}_{\mathsf{Lin}}\left(1 + \frac{4}{\min\{\dot{\mathsf{c}}_{\mathsf{Lin}}, \bar{a}/2\}}\right)e^{\min\{\dot{\mathsf{c}}_{\mathsf{Lin}},\bar{a}/2\}/4}\left(\mathsf{C}_{\mathsf{Lin},0} + 9\right),\\ \mathsf{C}_{\mathsf{Lin},3} &:= 2(\hat{\mathsf{c}}_{\mathsf{Lin}}/\dot{\mathsf{c}}_{\mathsf{Lin}})e^{\dot{\mathsf{c}}_{\mathsf{Lin}}/2}\left(\mathsf{C}_{\mathsf{Lin},1} + 9 + 9\bar{\mathsf{c}}_2(1+1/\bar{a}) + 9\mathrm{v}_4(\overline{\mathsf{M}}_V(4))\right)\end{aligned} \tag{115}$$

with $\dot{\mathsf{c}}_{\mathsf{Lin}}, \hat{\mathsf{c}}_{\mathsf{Lin}}$ given in Proposition D.7, $\mathsf{C}_{\mathsf{Lin},0}, \mathsf{C}_{\mathsf{Lin},1}$ given in (167) (see also Lemma D.6), $\bar{\mathsf{c}}_2$ given in Lemma D.3, and $\overline{\mathsf{M}}_V(4)$ given in Lemma D.5.

Proof. See [18, Lemma 4.7]. □

To obtain an upper bound for the last term on the RHS of (112), we observe that π_β is the invariant measure of the Langevin SDE (107). Thus, by applying Proposition D.7, we have that

$$W_1(\mathcal{L}(Z_t^\lambda), \pi_\beta) \leq \hat{\mathsf{c}}_{\mathsf{Lin}}e^{-\dot{\mathsf{c}}_{\mathsf{Lin}}\lambda t}w_{1,2}(\mathcal{L}(\theta_0), \pi_\beta) \leq \hat{\mathsf{c}}_{\mathsf{Lin}}e^{-\dot{\mathsf{c}}_{\mathsf{Lin}}\lambda t}\left[1 + \mathbb{E}[V_2(\theta_0)] + \int_{\mathbb{R}^d} V_2(\theta)\pi_\beta(\mathrm{d}\theta)\right]. \tag{116}$$

By using Lemma D.6, D.8 and (116), we can obtain an upper bound for each $W_2(\mathcal{L}(\Theta_n^{\mathsf{aHOLLA}}), \pi_\beta)$, $n \in \mathbb{N}_0$, as stated in Theorem 3.3.

Proof of Theorem 3.3. Substituting the results in Lemma D.6, D.8 and (116) into (112), for any $0 < \lambda \leq \overline{\lambda}_{\max}$, $n \in \mathbb{N}_0$, and $t \in (nT, (n+1)T]$, we have that

$$W_1(\mathcal{L}(\overline{\Theta}_t^\lambda), \pi_\beta) \leq \lambda^{1+q/2}\left(e^{-\bar{a}n/4}\mathsf{C}_{\mathsf{Lin},0}\mathbb{E}[|\theta_0|^4] + \mathsf{C}_{\mathsf{Lin},1}\right)^{1/2}$$

$$
\begin{aligned}
&+ \lambda^{1+q/2} \left(e^{-\min\{\dot{\mathsf{c}}_{\mathsf{Lin}}, \bar{a}/2\} n/4} \mathsf{C}_{\mathsf{Lin},2} \mathbb{E}[|\theta_0|^4] + \mathsf{C}_{\mathsf{Lin},3} \right) \\
&+ \hat{\mathsf{c}}_{\mathsf{Lin}} e^{-\dot{\mathsf{c}}_{\mathsf{Lin}} \lambda t} \left[1 + \mathbb{E}[V_2(\theta_0)] + \int_{\mathbb{R}^d} V_2(\theta) \pi_\beta(\mathrm{d}\theta) \right] \\
\leq\, & C_{\mathsf{Lin},1} e^{-C_{\mathsf{Lin},0}(n+1)} (\mathbb{E}[|\theta_0|^4] + 1) + C_{\mathsf{Lin},2} \lambda^{1+q/2},
\end{aligned}
$$

where

$$
\begin{aligned}
C_{\mathsf{Lin},0} &:= \min\{\dot{\mathsf{c}}_{\mathsf{Lin}}, \bar{a}/2\}/4, \\
C_{\mathsf{Lin},1} &:= e^{\min\{\dot{\mathsf{c}}_{\mathsf{Lin}}, \bar{a}/2\}/4} \left[\mathsf{C}_{\mathsf{Lin},0}^{1/2} + \mathsf{C}_{\mathsf{Lin},2} + \hat{\mathsf{c}}_{\mathsf{Lin}} \left(3 + \int_{\mathbb{R}^d} V_2(\theta) \pi_\beta(\mathrm{d}\theta) \right) \right], \\
C_{\mathsf{Lin},2} &:= \mathsf{C}_{\mathsf{Lin},1}^{1/2} + \mathsf{C}_{\mathsf{Lin},3}
\end{aligned}
\tag{117}
$$

with $\dot{\mathsf{c}}_{\mathsf{Lin}}, \hat{\mathsf{c}}_{\mathsf{Lin}}$ given in Proposition D.7, $\mathsf{C}_{\mathsf{Lin},0}, \mathsf{C}_{\mathsf{Lin},1}$ given in (167) (see also Lemma D.6), $\mathsf{C}_{\mathsf{Lin},2}, \mathsf{C}_{\mathsf{Lin},3}$ given in (115) (see also Lemma D.8). The above result implies that, for each $n \in \mathbb{N}_0$,

$$
W_1(\mathcal{L}(\overline{\Theta}_{nT}^\lambda), \pi_\beta) \leq C_{\mathsf{Lin},1} e^{-C_{\mathsf{Lin},0} n} (\mathbb{E}[|\theta_0|^4] + 1) + C_{\mathsf{Lin},2} \lambda^{1+q/2},
$$

which further yields, by setting nT to n on the LHS and n to n/T on the RHS, that

$$
\begin{aligned}
W_1(\mathcal{L}(\overline{\Theta}_n^\lambda), \pi_\beta) = W_1(\mathcal{L}(\Theta_n^\lambda), \pi_\beta) &= W_1(\mathcal{L}(\Theta_n^{\mathsf{aHOLLA}}), \pi_\beta) \\
&\leq C_{\mathsf{Lin},1} e^{-C_{\mathsf{Lin},0} \lambda n} (\mathbb{E}[|\theta_0|^4] + 1) + C_{\mathsf{Lin},2} \lambda^{1+q/2},
\end{aligned}
$$

where the inequality holds due to $n\lambda \leq n/T$. This completes the proof. □

We can obtain the upper bound for $W_2(\mathcal{L}(\Theta_n^{\mathsf{aHOLLA}}), \pi_\beta)$, $n \in \mathbb{N}_0$, as stated in Corollary 3.4, by applying similar arguments as in the proof of Theorem 3.3.

Proof of Corollary 3.4. To establish a non-asymptotic error bound for $W_2(\mathcal{L}(\Theta_n^{\mathsf{aHOLLA}}), \pi_\beta)$, we consider the following splitting: for any $0 < \lambda \leq \overline{\lambda}_{\max}$, $n \in \mathbb{N}_0$, and $t \in (nT, (n+1)T]$,

$$
W_2(\mathcal{L}(\overline{\Theta}_t^\lambda), \pi_\beta) \leq W_2(\mathcal{L}(\overline{\Theta}_t^\lambda), \mathcal{L}(\widetilde{\zeta}_t^{\lambda,n})) + W_2(\mathcal{L}(\widetilde{\zeta}_t^{\lambda,n}), \mathcal{L}(Z_t^\lambda)) + W_2(\mathcal{L}(Z_t^\lambda), \pi_\beta). \tag{118}
$$

An upper bound for the first term on the RHS of (118) is provided in Lemma D.6. To establish an estimate for the second term on the RHS of (118), we use $W_2 \leq \sqrt{2w_{1,2}}$ (see [18, Lemma A.3] for the proof) and follow the same arguments as that in the proof of [18, Lemma 4.7]. Consequently, for any $0 < \lambda \leq \overline{\lambda}_{\max}$, $n \in \mathbb{N}_0$, and $t \in (nT, (n+1)T]$, we obtain that,

$$
W_2(\mathcal{L}(\widetilde{\zeta}_t^{\lambda,n}), \mathcal{L}(Z_t^\lambda)) \leq \lambda^{1/2+q/4} \left(e^{-\min\{\dot{\mathsf{c}}_{\mathsf{Lin}}, \bar{a}/2\} n/8} \mathsf{C}_{\mathsf{Lin},4} \mathbb{E}^{1/2}[|\theta_0|^4] + \mathsf{C}_{\mathsf{Lin},5} \right), \tag{119}
$$

where

$$
\begin{aligned}
\mathsf{C}_{\mathsf{Lin},4} &:= \sqrt{\hat{\mathsf{c}}_{\mathsf{Lin}}} \left(1 + \frac{8}{\min\{\dot{\mathsf{c}}_{\mathsf{Lin}}, \bar{a}/2\}} \right) e^{\min\{\dot{\mathsf{c}}_{\mathsf{Lin}}, \bar{a}/2\}/8} \left(\mathsf{C}_{\mathsf{Lin},0}^{1/2} + 3 \right), \\
\mathsf{C}_{\mathsf{Lin},5} &:= 4(\sqrt{\hat{\mathsf{c}}_{\mathsf{Lin}}}/\dot{\mathsf{c}}_{\mathsf{Lin}}) e^{\dot{\mathsf{c}}_{\mathsf{Lin}}/4} \left(\mathsf{C}_{\mathsf{Lin},1}^{1/2} + 1 + 3(\bar{\mathsf{c}}_2(1 + 1/\bar{a}) + 1)^{1/2} + \sqrt{3} \mathsf{v}_4^{1/2}(\overline{\mathsf{M}}_V(4)) \right)
\end{aligned}
\tag{120}
$$

with $\dot{\mathsf{c}}_{\mathsf{Lin}}, \hat{\mathsf{c}}_{\mathsf{Lin}}$ given in Proposition D.7, $\mathsf{C}_{\mathsf{Lin},0}, \mathsf{C}_{\mathsf{Lin},1}$ given in (167) (see also Lemma D.6), $\bar{\mathsf{c}}_2$ given in Lemma D.3, and $\overline{\mathsf{M}}_V(4)$ given in Lemma D.5. An upper bound for the last term on the RHS of (118) can be obtained by using $W_2 \leq \sqrt{2w_{1,2}}$ and Proposition D.7:

$$
\begin{aligned}
W_2(\mathcal{L}(Z_t^\lambda), \pi_\beta) &\leq \sqrt{2\hat{\mathsf{c}}_{\mathsf{Lin}}} e^{-\dot{\mathsf{c}}_{\mathsf{Lin}} \lambda t/2} w_{1,2}^{1/2}(\mathcal{L}(\theta_0), \pi_\beta) \\
&\leq \sqrt{2\hat{\mathsf{c}}_{\mathsf{Lin}}} e^{-\dot{\mathsf{c}}_{\mathsf{Lin}} \lambda t/2} \left[1 + \mathbb{E}[V_2(\theta_0)] + \int_{\mathbb{R}^d} V_2(\theta) \pi_\beta(\mathrm{d}\theta) \right]^{1/2}.
\end{aligned}
\tag{121}
$$

Applying the results in Lemma D.6, (119), (121) to (118) yields, for any $0 < \lambda \leq \overline{\lambda}_{\max}$, $n \in \mathbb{N}_0$, and $t \in (nT, (n+1)T]$, that

$$
\begin{aligned}
W_2(\mathcal{L}(\overline{\Theta}_t^\lambda), \pi_\beta) \leq\, & \lambda^{1+q/2} \left(e^{-\bar{a}n/4} \mathsf{C}_{\mathsf{Lin},0} \mathbb{E}[|\theta_0|^4] + \mathsf{C}_{\mathsf{Lin},1} \right)^{1/2} \\
&+ \lambda^{1/2+q/4} \left(e^{-\min\{\dot{\mathsf{c}}_{\mathsf{Lin}}, \bar{a}/2\} n/8} \mathsf{C}_{\mathsf{Lin},4} \mathbb{E}^{1/2}[|\theta_0|^4] + \mathsf{C}_{\mathsf{Lin},5} \right)
\end{aligned}
$$

$$+\sqrt{2\hat{\mathsf{c}}_{\mathsf{Lin}}}e^{-\dot{\mathsf{c}}_{\mathsf{Lin}}\lambda t/2}\left[1+\mathbb{E}[V_2(\theta_0)]+\int_{\mathbb{R}^d}V_2(\theta)\pi_\beta(\mathrm{d}\theta)\right]^{1/2}$$
$$\le C_{\mathsf{Lin},4}e^{-C_{\mathsf{Lin},3}(n+1)}(\mathbb{E}[|\theta_0|^4]+1)^{1/2}+C_{\mathsf{Lin},5}\lambda^{1/2+q/4},$$

where

$$C_{\mathsf{Lin},3}:=\min\{\dot{\mathsf{c}}_{\mathsf{Lin}},\overline{a}/2\}/8,$$
$$C_{\mathsf{Lin},4}:=e^{\min\{\dot{\mathsf{c}}_{\mathsf{Lin}},\overline{a}/2\}/8}\left[\mathsf{C}_{\mathsf{Lin},0}^{1/2}+\mathsf{C}_{\mathsf{Lin},4}+\sqrt{2\hat{\mathsf{c}}_{\mathsf{Lin}}}\left(3+\int_{\mathbb{R}^d}V_2(\theta)\pi_\beta(\mathrm{d}\theta)\right)^{1/2}\right], \tag{122}$$
$$C_{\mathsf{Lin},5}:=\mathsf{C}_{\mathsf{Lin},1}^{1/2}+\mathsf{C}_{\mathsf{Lin},5}$$

with $\dot{\mathsf{c}}_{\mathsf{Lin}},\hat{\mathsf{c}}_{\mathsf{Lin}}$ given in Proposition D.7, $\mathsf{C}_{\mathsf{Lin},0},\mathsf{C}_{\mathsf{Lin},1}$ given in (167) (see also Lemma D.6), $\mathsf{C}_{\mathsf{Lin},4},\mathsf{C}_{\mathsf{Lin},5}$ given in (120). This further implies that, for each $n\in\mathbb{N}_0$,

$$W_2(\mathcal{L}(\overline{\Theta}^\lambda_n),\pi_\beta)=W_2(\mathcal{L}(\Theta^\lambda_n),\pi_\beta)=W_2(\mathcal{L}(\Theta^{\mathsf{aHOLLA}}_n),\pi_\beta)$$
$$\le C_{\mathsf{Lin},4}e^{-C_{\mathsf{Lin},3}\lambda n}(\mathbb{E}[|\theta_0|^4]+1)^{1/2}+C_{\mathsf{Lin},5}\lambda^{1/2+q/4},$$

which completes the proof. □

D.4. **Proof of auxiliary results in Appendices D.2 and D.3.**

Lemma D.9. *Let Assumption 6 hold. Then, for any $p\in\mathbb{N},t\ge 0$, we obtain that*

$$\mathbb{E}[|Z_t^\lambda|^{2p}]\le e^{-\lambda p\overline{a}t}\mathbb{E}[|\theta_0|^{2p}]+2(\overline{b}+\beta^{-1}(d+2(p-1)))\overline{\mathsf{M}}_0^{2p-2}/\overline{a}<\infty,$$

where $\overline{\mathsf{M}}_0:=(2(\overline{b}+\beta^{-1}(d+2(p-1)))/\overline{a})^{1/2}$.

Proof. See [18, Lemma A.1]. □

Proof of Lemma D.3-(i). For any $0<\lambda\le\overline{\lambda}_{\max}\le 1$ with $\overline{\lambda}_{\max}$ given in (19), $t\in(n,n+1]$, $n\in\mathbb{N}_0$, we define

$$\overline{\Delta}^\lambda_{n,t}:=\widetilde{\Theta}^\lambda_n+\lambda\phi^\lambda_{\mathsf{Lin}}(\widetilde{\Theta}^\lambda_n)(t-n),\quad \overline{\Xi}^\lambda_{n,t}:=\sqrt{2\lambda\beta^{-1}}\psi^\lambda_{\mathsf{Lin}}(\widetilde{\Theta}^\lambda_n)(B^\lambda_t-B^\lambda_n), \tag{123}$$

where for all $\theta\in\mathbb{R}^d$,

$$\phi^\lambda_{\mathsf{Lin}}(\theta):=-h(\theta)+(\lambda/2)\left(H(\theta)h(\theta)-\beta^{-1}\Upsilon(\theta)\right), \tag{124}$$

and

$$\psi^\lambda_{\mathsf{Lin}}(\theta):=\sqrt{\mathsf{I}_d-\lambda H(\theta)+(\lambda^2/3)(H(\theta))^2}. \tag{125}$$

Then, by using (108), (123) – (125), and by noticing $\mathbb{E}\left[\langle\overline{\Delta}^\lambda_{n,t},\overline{\Xi}^\lambda_{n,t}\rangle\middle|\widetilde{\Theta}^\lambda_n\right]=0$, we have that

$$\mathbb{E}\left[|\widetilde{\Theta}^\lambda_t|^2\middle|\widetilde{\Theta}^\lambda_n\right]=|\overline{\Delta}^\lambda_{n,t}|^2+\mathbb{E}\left[|\overline{\Xi}^\lambda_{n,t}|^2\middle|\widetilde{\Theta}^\lambda_n\right]. \tag{126}$$

The second term on the RHS of (126) can be further upper bounded as follows:

$$\begin{aligned}\mathbb{E}\left[|\overline{\Xi}^\lambda_{n,t}|^2\middle|\widetilde{\Theta}^\lambda_n\right]&=2\lambda\beta^{-1}\mathbb{E}\left[\left\langle\psi^\lambda_{\mathsf{Lin}}(\widetilde{\Theta}^\lambda_n)(B^\lambda_t-B^\lambda_n),\psi^\lambda_{\mathsf{Lin}}(\widetilde{\Theta}^\lambda_n)(B^\lambda_t-B^\lambda_n)\right\rangle\middle|\widetilde{\Theta}^\lambda_n\right]\\&=2\lambda\beta^{-1}\mathbb{E}\left[\left\langle B^\lambda_t-B^\lambda_n,\left(\mathsf{I}_d-\lambda H(\widetilde{\Theta}^\lambda_n)+(\lambda^2/3)(H(\widetilde{\Theta}^\lambda_n))^2\right)(B^\lambda_t-B^\lambda_n)\right\rangle\middle|\widetilde{\Theta}^\lambda_n\right]\\&\le 2\lambda\beta^{-1}\left(d(t-n)+\lambda\overline{L}_3d(t-n)+(\lambda^2/3)\overline{L}_3^2d(t-n)\right)\\&\le 2\lambda\beta^{-1}(t-n)d(1+\overline{L}_3)^2,\end{aligned} \tag{127}$$

where the first inequality holds due to Remark 3.2. Next, to upper bound the first term on the RHS of (126), we use (123) to obtain

$$|\overline{\Delta}^\lambda_{n,t}|^2=|\widetilde{\Theta}^\lambda_n|^2+2\lambda(t-n)\left\langle\widetilde{\Theta}^\lambda_n,\phi^\lambda_{\mathsf{Lin}}(\widetilde{\Theta}^\lambda_n)\right\rangle+\lambda^2(t-n)^2|\phi^\lambda_{\mathsf{Lin}}(\widetilde{\Theta}^\lambda_n)|^2. \tag{128}$$

By using (124), the second term on the RHS of (128) can be estimated as follows:

$$\begin{aligned}\left\langle\widetilde{\Theta}^\lambda_n,\phi^\lambda_{\mathsf{Lin}}(\widetilde{\Theta}^\lambda_n)\right\rangle&=-\left\langle\widetilde{\Theta}^\lambda_n,h(\widetilde{\Theta}^\lambda_n)\right\rangle+(\lambda/2)\left\langle\widetilde{\Theta}^\lambda_n,H(\widetilde{\Theta}^\lambda_n)h(\widetilde{\Theta}^\lambda_n)\right\rangle-(\lambda/2)\beta^{-1}\left\langle\widetilde{\Theta}^\lambda_n,\Upsilon(\widetilde{\Theta}^\lambda_n)\right\rangle\\&\le-\overline{a}|\widetilde{\Theta}^\lambda_n|^2+\overline{b}+\lambda\overline{L}_3\overline{\mathsf{K}}_1(1+|\widetilde{\Theta}^\lambda_n|^2)+(\lambda/2)\beta^{-1}d\overline{L}_2|\widetilde{\Theta}^\lambda_n|,\end{aligned} \tag{129}$$

where the last inequality holds due to Assumption 6, Remark 3.2 and the following calculations: for all $\theta \in \mathbb{R}^d$,

$$|\theta||H(\theta)h(\theta)| \leq \overline{L}_3|\theta||h(\theta)| = \overline{L}_3\overline{\mathsf{K}}_1(|\theta|+|\theta|^2) \leq 2\overline{L}_3\overline{\mathsf{K}}_1(1+|\theta|^2).$$

To provide an upper bound for the third term on the RHS of (128), we write by straightforward calculations that

$$\begin{aligned}|\phi^{\lambda}_{\mathsf{Lin}}(\widetilde{\Theta}^{\lambda}_n)|^2 &= |-h(\widetilde{\Theta}^{\lambda}_n) + (\lambda/2)H(\widetilde{\Theta}^{\lambda}_n)h(\widetilde{\Theta}^{\lambda}_n) - (\lambda/2)\beta^{-1}\Upsilon(\widetilde{\Theta}^{\lambda}_n)|^2 \\ &= |h(\widetilde{\Theta}^{\lambda}_n)|^2 + (\lambda^2/4)|H(\widetilde{\Theta}^{\lambda}_n)h(\widetilde{\Theta}^{\lambda}_n)|^2 + (\lambda^2/4)\beta^{-2}|\Upsilon(\widetilde{\Theta}^{\lambda}_n)|^2 \\ &\quad - \lambda\left\langle h(\widetilde{\Theta}^{\lambda}_n), H(\widetilde{\Theta}^{\lambda}_n)h(\widetilde{\Theta}^{\lambda}_n)\right\rangle + \lambda\beta^{-1}\left\langle h(\widetilde{\Theta}^{\lambda}_n), \Upsilon(\widetilde{\Theta}^{\lambda}_n)\right\rangle \\ &\quad - (\lambda^2/2)\beta^{-1}\left\langle H(\widetilde{\Theta}^{\lambda}_n)h(\widetilde{\Theta}^{\lambda}_n), \Upsilon(\widetilde{\Theta}^{\lambda}_n)\right\rangle. \end{aligned} \tag{130}$$

We then provide upper bounds for each of the terms on the RHS of (130). By using Remark 3.2, we obtain that

$$\begin{aligned} |h(\widetilde{\Theta}^{\lambda}_n)|^2 &\leq 2\overline{\mathsf{K}}_1^2(1+|\widetilde{\Theta}^{\lambda}_n|^2), \\ (\lambda^2/4)|H(\widetilde{\Theta}^{\lambda}_n)h(\widetilde{\Theta}^{\lambda}_n)|^2 &\leq (\lambda^2/2)\overline{L}_3^2\overline{\mathsf{K}}_1^2(1+|\widetilde{\Theta}^{\lambda}_n|^2), \\ (\lambda^2/4)\beta^{-2}|\Upsilon(\widetilde{\Theta}^{\lambda}_n)|^2 &\leq (\lambda^2/4)\beta^{-2}d^2\overline{L}_2^2, \\ -\lambda\left\langle h(\widetilde{\Theta}^{\lambda}_n), H(\widetilde{\Theta}^{\lambda}_n)h(\widetilde{\Theta}^{\lambda}_n)\right\rangle &\leq 2\lambda\overline{L}_3\overline{\mathsf{K}}_1^2(1+|\widetilde{\Theta}^{\lambda}_n|^2), \\ \lambda\beta^{-1}\left\langle h(\widetilde{\Theta}^{\lambda}_n), \Upsilon(\widetilde{\Theta}^{\lambda}_n)\right\rangle &\leq \lambda\beta^{-1}d\overline{L}_2\overline{\mathsf{K}}_1(1+|\widetilde{\Theta}^{\lambda}_n|), \\ -(\lambda^2/2)\beta^{-1}\left\langle H(\widetilde{\Theta}^{\lambda}_n)h(\widetilde{\Theta}^{\lambda}_n), \Upsilon(\widetilde{\Theta}^{\lambda}_n)\right\rangle &\leq (\lambda^2/2)\beta^{-1}d\overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1(1+|\widetilde{\Theta}^{\lambda}_n|). \end{aligned} \tag{131}$$

Substituting (131) into (130) yields

$$\begin{aligned}|\phi^{\lambda}_{\mathsf{Lin}}(\widetilde{\Theta}^{\lambda}_n)|^2 &\leq 2\overline{\mathsf{K}}_1^2(1+|\widetilde{\Theta}^{\lambda}_n|^2) + (\lambda^2/2)\overline{L}_3^2\overline{\mathsf{K}}_1^2(1+|\widetilde{\Theta}^{\lambda}_n|^2) + (\lambda^2/4)\beta^{-2}d^2\overline{L}_2^2 \\ &\quad + 2\lambda\overline{L}_3\overline{\mathsf{K}}_1^2(1+|\widetilde{\Theta}^{\lambda}_n|^2) + \lambda\beta^{-1}d\overline{L}_2\overline{\mathsf{K}}_1(1+|\widetilde{\Theta}^{\lambda}_n|) + (\lambda^2/2)\beta^{-1}d\overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1(1+|\widetilde{\Theta}^{\lambda}_n|). \end{aligned} \tag{132}$$

Combining the results in (129) and (132), we obtain the following upper bound for (128):

$$\begin{aligned}|\overline{\Delta}^{\lambda}_{n,t}|^2 &\leq |\widetilde{\Theta}^{\lambda}_n|^2 - 2\lambda(t-n)\overline{a}|\widetilde{\Theta}^{\lambda}_n|^2 + \lambda(t-n)\left(2\lambda\overline{L}_3\overline{\mathsf{K}}_1 + 2\lambda\overline{\mathsf{K}}_1^2 + \lambda^3\overline{L}_3^2\overline{\mathsf{K}}_1^2/2 + 2\lambda^2\overline{L}_3\overline{\mathsf{K}}_1^2\right)|\widetilde{\Theta}^{\lambda}_n|^2 \\ &\quad + \lambda(t-n)\left(\beta^{-1}d\overline{L}_2 + \beta^{-1}d\overline{L}_2\overline{\mathsf{K}}_1 + \beta^{-1}d\overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2\right)|\widetilde{\Theta}^{\lambda}_n| + \lambda(t-n)\Big(2\overline{b} + 2\overline{L}_3\overline{\mathsf{K}}_1 \\ &\quad + 2\overline{\mathsf{K}}_1^2 + \overline{L}_3^2\overline{\mathsf{K}}_1^2/2 + \beta^{-2}d^2\overline{L}_2^2/4 + 2\overline{L}_3\overline{\mathsf{K}}_1^2 + \beta^{-1}d\overline{L}_2\overline{\mathsf{K}}_1 + \beta^{-1}d\overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2\Big) \\ &= (1-\lambda(t-n)\overline{a})|\widetilde{\Theta}^{\lambda}_n|^2 - \lambda(t-n)\overline{\mathfrak{I}}^{\lambda}_1(\widetilde{\Theta}^{\lambda}_n) - \lambda(t-n)\overline{\mathfrak{I}}^{\lambda}_2(\widetilde{\Theta}^{\lambda}_n) + \lambda(t-n)\overline{\mathsf{c}}_1, \end{aligned} \tag{133}$$

where, for all $\theta \in \mathbb{R}^d$,

$$\begin{aligned}\overline{\mathfrak{I}}^{\lambda}_1(\theta) &:= (\overline{a}/2)|\theta|^2 - \left(2\lambda\overline{L}_3\overline{\mathsf{K}}_1 + 2\lambda\overline{\mathsf{K}}_1^2 + \lambda^3\overline{L}_3^2\overline{\mathsf{K}}_1^2/2 + 2\lambda^2\overline{L}_3\overline{\mathsf{K}}_1^2\right)|\theta|^2 \\ \overline{\mathfrak{I}}^{\lambda}_2(\theta) &:= (\overline{a}/2)|\theta|^2 - \left(\beta^{-1}d\overline{L}_2 + \beta^{-1}d\overline{L}_2\overline{\mathsf{K}}_1 + \beta^{-1}d\overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2\right)|\theta|,\end{aligned}$$

and where $\overline{\mathsf{c}}_1 := 2\overline{b} + 2\overline{L}_3\overline{\mathsf{K}}_1 + 2\overline{\mathsf{K}}_1^2 + \overline{L}_3^2\overline{\mathsf{K}}_1^2/2 + \beta^{-2}d^2\overline{L}_2^2/4 + 2\overline{L}_3\overline{\mathsf{K}}_1^2 + \beta^{-1}d\overline{L}_2\overline{\mathsf{K}}_1 + \beta^{-1}d\overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2$. We note that, for all $\theta \in \mathbb{R}^d$, $0 < \lambda \leq \overline{\lambda}_{\max} \leq \min\{\overline{a}/(16\overline{L}_3\overline{\mathsf{K}}_1), \overline{a}/(16\overline{\mathsf{K}}_1^2), \overline{a}^{1/3}/(4\overline{L}_3^2\overline{\mathsf{K}}_1^2)^{1/3}, \overline{a}^{1/2}/(16\overline{L}_3\overline{\mathsf{K}}_1^2)^{1/3}\}$,

$$\overline{\mathfrak{I}}^{\lambda}_1(\theta) = \left((\overline{a}/8 - 2\lambda\overline{L}_3\overline{\mathsf{K}}_1) + (\overline{a}/8 - 2\lambda\overline{\mathsf{K}}_1^2) + (\overline{a}/8 - \lambda^3\overline{L}_3^2\overline{\mathsf{K}}_1^2/2) + (\overline{a}/8 - 2\lambda^2\overline{L}_3\overline{\mathsf{K}}_1^2)\right)|\theta|^2 \geq 0. \tag{134}$$

Substituting (134) into (133) yields

$$|\overline{\Delta}^{\lambda}_{n,t}|^2 \leq (1-\lambda(t-n)\overline{a})|\widetilde{\Theta}^{\lambda}_n|^2 - \lambda(t-n)\overline{\mathfrak{I}}^{\lambda}_2(\widetilde{\Theta}^{\lambda}_n) + \lambda(t-n)\overline{\mathsf{c}}_1.$$

Denote by $\overline{\mathsf{M}}_1 := 2(\overline{a}\beta)^{-1}d(\overline{L}_2 + \overline{L}_2\overline{\mathsf{K}}_1 + \overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2)$ and $\overline{\mathsf{S}}_{n,\overline{\mathsf{M}}_1} := \{\omega \in \Omega : |\widetilde{\Theta}_n^\lambda(\omega)| > \overline{\mathsf{M}}_1\}$. By observing the fact that

$$\overline{\mathfrak{I}}_2^\lambda(\theta) > 0 \qquad \Longleftrightarrow \qquad |\theta| > \overline{\mathsf{M}},$$

we obtain the following:

$$|\overline{\Delta}_{n,t}^\lambda|^2 \mathbb{1}_{\overline{\mathsf{S}}_{n,\overline{\mathsf{M}}_1}} \leq (1 - \lambda(t-n)\overline{a})|\widetilde{\Theta}_n^\lambda|^2 \mathbb{1}_{\overline{\mathsf{S}}_{n,\overline{\mathsf{M}}_1}} + \lambda(t-n)\overline{\mathsf{c}}_1 \mathbb{1}_{\overline{\mathsf{S}}_{n,\overline{\mathsf{M}}_1}}.$$

Similarly, we have that

$$\begin{aligned}|\overline{\Delta}_{n,t}^\lambda|^2 \mathbb{1}_{\overline{\mathsf{S}}^{\mathsf{c}}_{n,\overline{\mathsf{M}}_1}} &\leq (1 - \lambda(t-n)\overline{a})|\widetilde{\Theta}_n^\lambda|^2 \mathbb{1}_{\overline{\mathsf{S}}^{\mathsf{c}}_{n,\overline{\mathsf{M}}_1}} + \lambda(t-n)\overline{\mathsf{c}}_1 \mathbb{1}_{\overline{\mathsf{S}}^{\mathsf{c}}_{n,\overline{\mathsf{M}}_1}} \\ &\quad + \lambda(t-n)\left(\beta^{-1}d\overline{L}_2 + \beta^{-1}d\overline{L}_2\overline{\mathsf{K}}_1 + \beta^{-1}d\overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2\right)\overline{\mathsf{M}}_1 \mathbb{1}_{\overline{\mathsf{S}}^{\mathsf{c}}_{n,\overline{\mathsf{M}}_1}}.\end{aligned}$$

Combining the two cases yields

$$|\overline{\Delta}_{n,t}^\lambda|^2 \leq (1 - \lambda(t-n)\overline{a})|\widetilde{\Theta}_n^\lambda|^2 + \lambda(t-n)\overline{\mathsf{c}}_1 + \lambda(t-n)\beta^{-1}d\left(\overline{L}_2 + \overline{L}_2\overline{\mathsf{K}}_1 + \overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2\right)\overline{\mathsf{M}}_1. \tag{135}$$

Finally, by substituting (127) and (135) into (126), we obtain

$$\mathbb{E}\left[|\widetilde{\Theta}_t^\lambda|^2 \middle| \widetilde{\Theta}_n^\lambda\right] \leq (1 - \lambda(t-n)\overline{a})|\widetilde{\Theta}_n^\lambda|^2 + \lambda(t-n)\overline{\mathsf{c}}_0, \tag{136}$$

where

$$\begin{aligned}\overline{\mathsf{c}}_0 &:= \beta^{-1}d\left(\overline{L}_2 + \overline{L}_2\overline{\mathsf{K}}_1 + \overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2\right)\overline{\mathsf{M}}_1 + \overline{\mathsf{c}}_1 + 2\beta^{-1}d(1+\overline{L}_3)^2, \\ \overline{\mathsf{c}}_1 &:= 2\overline{b} + 2\overline{L}_3\overline{\mathsf{K}}_1 + 2\overline{\mathsf{K}}_1^2 + \overline{L}_3^2\overline{\mathsf{K}}_1^2/2 + \beta^{-2}d^2\overline{L}_2^2/4 + 2\overline{L}_3\overline{\mathsf{K}}_1^2 + \beta^{-1}d(\overline{L}_2\overline{\mathsf{K}}_1 + \overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2), \\ \overline{\mathsf{M}}_1 &:= 2(\overline{a}\beta)^{-1}d(\overline{L}_2 + \overline{L}_2\overline{\mathsf{K}}_1 + \overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2).\end{aligned} \tag{137}$$

We observe that, for $0 < \lambda \leq \overline{\lambda}_{\max} \leq 1/\overline{a}$,

$$1 > 1 - \lambda(t-n)\overline{a} > 1 - \lambda\overline{a} \geq 0,$$

then, by induction, (136) implies, for $t \in (n, n+1]$, $n \in \mathbb{N}_0$, $0 < \lambda \leq \overline{\lambda}_{\max} \leq 1$, that,

$$\begin{aligned}\mathbb{E}\left[|\widetilde{\Theta}_t^\lambda|^2\right] &\leq (1 - \lambda(t-n)\overline{a})\mathbb{E}\left[|\widetilde{\Theta}_n^\lambda|^2\right] + \lambda(t-n)\overline{\mathsf{c}}_0 \\ &\leq (1 - \lambda(t-n)\overline{a})(1 - \lambda\overline{a})\,\mathbb{E}\left[|\widetilde{\Theta}_{n-1}^\lambda|^2\right] + \overline{\mathsf{c}}_0 + \lambda\overline{\mathsf{c}}_0 \\ &\leq (1 - \lambda(t-n)\overline{a})(1 - \lambda\overline{a})^2\,\mathbb{E}\left[|\widetilde{\Theta}_{n-2}^\lambda|^2\right] + \overline{\mathsf{c}}_0 + \lambda\overline{\mathsf{c}}_0\left(1 + (1 - \lambda\overline{a})\right) \\ &\leq \dots \\ &\leq (1 - \lambda(t-n)\overline{a})(1 - \lambda\overline{a})^n\,\mathbb{E}\left[|\theta_0|^2\right] + \overline{\mathsf{c}}_0\left(1 + 1/\overline{a}\right),\end{aligned}$$

which completes the proof. □

Proof of Lemma D.3-(ii). For any $p \in [2,\infty) \cap \mathbb{N}$, $0 < \lambda \leq \overline{\lambda}_{\max} \leq 1$ with $\overline{\lambda}_{\max}$ given in (19), $t \in (n, n+1]$, $n \in \mathbb{N}_0$, by using the same arguments as in the proof of [18, Lemma 4.2-(ii)] up to the inequality before [18, Eq. (134)] and by using (108) with (123), we obtain that

$$\begin{aligned}\mathbb{E}\left[|\widetilde{\Theta}_t^\lambda|^{2p} \middle| \widetilde{\Theta}_n^\lambda\right] &\leq |\overline{\Delta}_{n,t}^\lambda|^{2p} + 2^{2p-3}p(2p-1)|\overline{\Delta}_{n,t}^\lambda|^{2p-2}\mathbb{E}\left[|\overline{\Xi}_{n,t}^\lambda|^2 \middle| \widetilde{\Theta}_n^\lambda\right] \\ &\quad + 2^{2p-3}p(2p-1)\mathbb{E}\left[|\overline{\Xi}_{n,t}^\lambda|^{2p} \middle| \widetilde{\Theta}_n^\lambda\right].\end{aligned} \tag{138}$$

Then, by using (123) and (125), we can obtain an upper estimate for the last term in (138) as follows:

$$\begin{aligned}\mathbb{E}\left[|\overline{\Xi}_{n,t}^\lambda|^{2p} \middle| \widetilde{\Theta}_n^\lambda\right] &= (2\lambda\beta^{-1})^p\mathbb{E}\left[\left\langle (B_t^\lambda - B_n^\lambda), \left(\psi_{\mathsf{Lin}}^\lambda(\widetilde{\Theta}_n^\lambda)\right)^2 (B_t^\lambda - B_n^\lambda)\right\rangle^p \middle| \widetilde{\Theta}_n^\lambda\right] \\ &\leq (2\lambda\beta^{-1})^p\mathbb{E}\left[\left(|B_t^\lambda - B_n^\lambda|^2 + \lambda\overline{L}_3|B_t^\lambda - B_n^\lambda|^2 + (\lambda^2/3)\overline{L}_3^2|B_t^\lambda - B_n^\lambda|^2\right)^p \middle| \widetilde{\Theta}_n^\lambda\right] \\ &\leq (2\lambda\beta^{-1})^p(1+\overline{L}_3)^{2p}\mathbb{E}\left[|B_t^\lambda - B_n^\lambda|^{2p}\right] \\ &\leq (2\lambda\beta^{-1}dp(2p-1)(t-n))^p(1+\overline{L}_3)^{2p} \\ &\leq \lambda(t-n)(2\beta^{-1}dp(2p-1)(1+\overline{L}_3)^2)^p.\end{aligned} \tag{139}$$

Substituting (127) and (139) into (138) yields

$$\mathbb{E}\left[|\widetilde{\Theta}_t^\lambda|^{2p}\middle|\widetilde{\Theta}_n^\lambda\right] \leq |\overline{\Delta}_{n,t}^\lambda|^{2p} + \lambda(t-n)2^{2p-2}p(2p-1)\beta^{-1}d(1+\overline{L}_3)^2|\overline{\Delta}_{n,t}^\lambda|^{2p-2} + \lambda(t-n)\bar{\mathsf{c}}_\Xi(p), \tag{140}$$

where $\bar{\mathsf{c}}_\Xi(p) := 2^{2p-3}p(2p-1)(2\beta^{-1}dp(2p-1)(1+\overline{L}_3)^2)^p$. Next, we apply (135) to obtain

$$\begin{aligned}
|\overline{\Delta}_{n,t}^\lambda|^{2p} &\leq \left((1-\lambda(t-n)\bar{a})|\widetilde{\Theta}_n^\lambda|^2 + \lambda(t-n)\bar{\mathsf{c}}_1 + \lambda(t-n)\beta^{-1}d\left(\overline{L}_2 + \overline{L}_2\overline{\mathsf{K}}_1 + \overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2\right)\overline{\mathsf{M}}_1\right)^p \\
&\leq (1+\lambda(t-n)\bar{a}/2)^{p-1}\left(1-\lambda(t-n)\bar{a}\right)^p|\widetilde{\Theta}_n^\lambda|^{2p} \\
&\quad + (1+2/(\lambda(t-n)\bar{a}))^{p-1}\left(\lambda(t-n)\right)^p(\bar{\mathsf{c}}_1 + \beta^{-1}d\left(\overline{L}_2 + \overline{L}_2\overline{\mathsf{K}}_1 + \overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2\right)\overline{\mathsf{M}}_1)^p \\
&\leq (1-\lambda(t-n)\bar{a}/2)^{p-1}\left(1-\lambda(t-n)\bar{a}\right)|\widetilde{\Theta}_n^\lambda|^{2p} \\
&\quad + \lambda(t-n)\left(1+2/\bar{a}\right)^{p-1}(\bar{\mathsf{c}}_1 + \beta^{-1}d\left(\overline{L}_2 + \overline{L}_2\overline{\mathsf{K}}_1 + \overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2\right)\overline{\mathsf{M}}_1)^p \\
&= \bar{\mathsf{c}}_{n,t}^\lambda(p)|\widetilde{\Theta}_n^\lambda|^{2p} + \widetilde{\mathsf{c}}_{n,t}^\lambda(p),
\end{aligned} \tag{141}$$

where the second inequality holds due to $(u+v)^p \leq (1+\varepsilon)^{p-1}u^p + (1+\varepsilon^{-1})^{p-1}v^p$, $u, v \geq 0$, $\varepsilon > 0$ with $\varepsilon = \lambda(t-n)\bar{a}/2$, and where

$$\begin{aligned}
\bar{\mathsf{c}}_{n,t}^\lambda(p) &:= (1-\lambda(t-n)\bar{a}/2)^{p-1}\left(1-\lambda(t-n)\bar{a}\right), \\
\widetilde{\mathsf{c}}_{n,t}^\lambda(p) &:= \lambda(t-n)\left(1+2/\bar{a}\right)^{p-1}(\bar{\mathsf{c}}_1 + \beta^{-1}d\left(\overline{L}_2 + \overline{L}_2\overline{\mathsf{K}}_1 + \overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2\right)\overline{\mathsf{M}}_1)^p.
\end{aligned}$$

In addition, we observe that by (141),

$$|\overline{\Delta}_{n,t}^\lambda|^{2p-2} \leq \bar{\mathsf{c}}_{n,t}^\lambda(p-1)|\widetilde{\Theta}_n^\lambda|^{2p-2} + \widetilde{\mathsf{c}}_{n,t}^\lambda(p-1), \tag{142}$$

and, in particular, when $p = 2$, (142) yields $|\overline{\Delta}_{n,t}^\lambda|^2 \leq \bar{\mathsf{c}}_{n,t}^\lambda(1)|\widetilde{\Theta}_n^\lambda|^2 + \widetilde{\mathsf{c}}_{n,t}^\lambda(1)$ which is exactly the upper bound (135). By substituting (141) and (142) into (140), we have that

$$\begin{aligned}
\mathbb{E}\left[|\widetilde{\Theta}_t^\lambda|^{2p}\middle|\widetilde{\Theta}_n^\lambda\right] &\leq \bar{\mathsf{c}}_{n,t}^\lambda(p)|\widetilde{\Theta}_n^\lambda|^{2p} + \widetilde{\mathsf{c}}_{n,t}^\lambda(p) + \lambda(t-n)\bar{\mathsf{c}}_\Xi(p) \\
&\quad + \lambda(t-n)2^{2p-2}p(2p-1)\beta^{-1}d(1+\overline{L}_3)^2\left(\bar{\mathsf{c}}_{n,t}^\lambda(p-1)|\widetilde{\Theta}_n^\lambda|^{2p-2} + \widetilde{\mathsf{c}}_{n,t}^\lambda(p-1)\right).
\end{aligned} \tag{143}$$

Denote by $\overline{\mathsf{M}}_2(p) := (2^{2p}p(2p-1)\beta^{-1}d(1+\overline{L}_3)^2/\bar{a})^{1/2}$. For all $|\theta| > \overline{\mathsf{M}}_2(p)$, we have that

$$(\lambda(t-n)\bar{a}/4)|\theta|^{2p} > (\lambda(t-n)2^{2p}p(2p-1)\beta^{-1}d(1+\overline{L}_3)^2/4)|\theta|^{2p-2}.$$

Denote by $\overline{\mathsf{S}}_{n,\overline{\mathsf{M}}_2(p)} := \{\omega \in \Omega : |\widetilde{\Theta}_n^\lambda(\omega)| > \overline{\mathsf{M}}_2(p)\}$. By using the above inequality, (143) can be further bounded as follows:

$$\begin{aligned}
&\mathbb{E}\left[|\widetilde{\Theta}_t^\lambda|^{2p}\mathbb{1}_{\overline{\mathsf{S}}_{n,\overline{\mathsf{M}}_2(p)}}\middle|\widetilde{\Theta}_n^\lambda\right] \\
&\leq (1-\lambda(t-n)\bar{a}/4)\,\bar{\mathsf{c}}_{n,t}^\lambda(p-1)|\widetilde{\Theta}_n^\lambda|^{2p}\mathbb{1}_{\overline{\mathsf{S}}_{n,\overline{\mathsf{M}}_2(p)}} + \left(\widetilde{\mathsf{c}}_{n,t}^\lambda(p) + \lambda(t-n)\bar{\mathsf{c}}_\Xi(p)\right)\mathbb{1}_{\overline{\mathsf{S}}_{n,\overline{\mathsf{M}}_2(p)}} \\
&\quad + \lambda(t-n)2^{2p-2}p(2p-1)\beta^{-1}d(1+\overline{L}_3)^2\widetilde{\mathsf{c}}_{n,t}^\lambda(p-1)\mathbb{1}_{\overline{\mathsf{S}}_{n,\overline{\mathsf{M}}_2(p)}} \\
&\quad - (\lambda(t-n)\bar{a}/4)\,\bar{\mathsf{c}}_{n,t}^\lambda(p-1)|\widetilde{\Theta}_n^\lambda|^{2p}\mathbb{1}_{\overline{\mathsf{S}}_{n,\overline{\mathsf{M}}_2(p)}} \\
&\quad + (\lambda(t-n)2^{2p}p(2p-1)\beta^{-1}d(1+\overline{L}_3)^2/4)\bar{\mathsf{c}}_{n,t}^\lambda(p-1)|\widetilde{\Theta}_n^\lambda|^{2p-2}\mathbb{1}_{\overline{\mathsf{S}}_{n,\overline{\mathsf{M}}_2(p)}} \\
&\leq (1-\lambda(t-n)\bar{a})\,|\widetilde{\Theta}_n^\lambda|^{2p}\mathbb{1}_{\overline{\mathsf{S}}_{n,\overline{\mathsf{M}}_2(p)}} + \lambda(t-n)\left[\bar{\mathsf{c}}_\Xi(p) + (1+2/\bar{a})^{p-1}\right. \\
&\quad \times(\bar{\mathsf{c}}_1 + \beta^{-1}d\left(\overline{L}_2 + \overline{L}_2\overline{\mathsf{K}}_1 + \overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2\right)\overline{\mathsf{M}}_1)^p + 2^{2p-2}p(2p-1)\beta^{-1}d(1+\overline{L}_3)^2 \\
&\qquad \times\left.\left((1+2/\bar{a})^{p-2}(\bar{\mathsf{c}}_1 + \beta^{-1}d\left(\overline{L}_2 + \overline{L}_2\overline{\mathsf{K}}_1 + \overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2\right)\overline{\mathsf{M}}_1)^{p-1} + \overline{\mathsf{M}}_2(p)^{2p-2}\right)\right]\mathbb{1}_{\overline{\mathsf{S}}_{n,\overline{\mathsf{M}}_2(p)}} \\
&\leq (1-\lambda(t-n)\bar{a})\,|\widetilde{\Theta}_n^\lambda|^{2p}\mathbb{1}_{\overline{\mathsf{S}}_{n,\overline{\mathsf{M}}_2(p)}} + \lambda(t-n)\bar{\mathsf{c}}_p\mathbb{1}_{\overline{\mathsf{S}}_{n,\overline{\mathsf{M}}_2(p)}},
\end{aligned} \tag{144}$$

where

$$\begin{aligned}\bar{\mathsf{c}}_p &:= \bar{\mathsf{c}}_{\Xi}(p) + (1+2/\bar{a})^{p-1}\left(\bar{\mathsf{c}}_1 + \beta^{-1}d\left(\overline{L}_2 + \overline{L}_2\overline{\mathsf{K}}_1 + \overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2\right)\overline{\mathsf{M}}_1\right)^p \\ &\quad + 2^{2p-2}p(2p-1)\beta^{-1}d(1+\overline{L}_3)^2 \\ &\qquad \times \left((1+2/\bar{a})^{p-2}\left(\bar{\mathsf{c}}_1 + \beta^{-1}d\left(\overline{L}_2 + \overline{L}_2\overline{\mathsf{K}}_1 + \overline{L}_2\overline{L}_3\overline{\mathsf{K}}_1/2\right)\overline{\mathsf{M}}_1\right)^{p-1} + \overline{\mathsf{M}}_2(p)^{2p-2}\right), \quad (145) \\ \bar{\mathsf{c}}_{\Xi}(p) &:= 2^{2p-3}p(2p-1)(2\beta^{-1}dp(2p-1)(1+\overline{L}_3)^2)^p, \\ \overline{\mathsf{M}}_2(p) &:= (2^{2p}p(2p-1)\beta^{-1}d(1+\overline{L}_3)^2/\bar{a})^{1/2}\end{aligned}$$

with $\bar{\mathsf{c}}_1$ given in (137). Similarly, by using (143), we have that

$$\mathbb{E}\left[|\widetilde{\Theta}_t^\lambda|^{2p}\mathbb{1}_{\overline{\mathsf{S}}^{\mathsf{c}}_{n,\overline{\mathsf{M}}_2(p)}}\,\middle|\,\widetilde{\Theta}_n^\lambda\right] \leq (1-\lambda(t-n)\bar{a})\,|\widetilde{\Theta}_n^\lambda|^{2p}\mathbb{1}_{\overline{\mathsf{S}}^{\mathsf{c}}_{n,\overline{\mathsf{M}}_2(p)}} + \lambda(t-n)\bar{\mathsf{c}}_p\mathbb{1}_{\overline{\mathsf{S}}^{\mathsf{c}}_{n,\overline{\mathsf{M}}_2(p)}}. \quad (146)$$

Combing (144) and (146) yields the desired result. □

Lemma D.10. *Let Assumptions 4, 5, and 6 hold. Then, for any* $0 < \lambda \leq \overline{\lambda}_{\max}$, $p > 0, t \geq 0$, *we obtain*

$$\mathbb{E}\left[|\overline{\Theta}_t^\lambda - \overline{\Theta}_{\lfloor t\rfloor}^\lambda|^{2p}\right] \leq \lambda^p\left(e^{-\lambda\bar{a}\lfloor t\rfloor}\overline{\mathsf{C}}_{\mathbf{S}0,p}\mathbb{E}[|\theta_0|^{2\lceil p\rceil}] + \widetilde{\mathsf{C}}_{\mathbf{S}0,p}\right), \quad (147)$$

$$\mathbb{E}\left[|\widetilde{\zeta}_t^{\lambda,n} - \widetilde{\zeta}_{\lfloor t\rfloor}^{\lambda,n}|^{2p}\right] \leq \lambda^p\left(e^{-\lambda\bar{a}\lfloor t\rfloor/2}\overline{\mathsf{C}}_{\mathbf{S}1,p}\mathbb{E}[|\theta_0|^{2\lceil p\rceil}] + \widetilde{\mathsf{C}}_{\mathbf{S}1,p}\right), \quad (148)$$

where $\overline{\mathsf{C}}_{\mathbf{S}0,p}$ *and* $\widetilde{\mathsf{C}}_{\mathbf{S}0,p}$ *are given in* (149), *and* $\overline{\mathsf{C}}_{\mathbf{S}1,p}$ *and* $\widetilde{\mathsf{C}}_{\mathbf{S}1,p}$ *are given in* (150).

Proof. To show that (147) holds, we use the definition of $(\overline{\Theta}_t^\lambda)_{t\geq 0}$ given in (109) and obtain that, for any $t \geq 0$,

$$\begin{aligned}&\mathbb{E}\left[|\overline{\Theta}_t^\lambda - \overline{\Theta}_{\lfloor t\rfloor}^\lambda|^{2p}\right] \\ &= \mathbb{E}\left[\left|-\lambda\int_{\lfloor t\rfloor}^t h(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\,\mathrm{d}s + \lambda^2\int_{\lfloor t\rfloor}^t\int_{\lfloor s\rfloor}^s\left(H(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)h(\overline{\Theta}_{\lfloor r\rfloor}^\lambda) - \beta^{-1}\Upsilon(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)\right)\mathrm{d}r\,\mathrm{d}s\right.\right. \\ &\qquad\left.\left. -\lambda\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t\int_{\lfloor s\rfloor}^s H(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\,\mathrm{d}s + \sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t\mathrm{d}B_s^\lambda\right|^{2p}\right] \\ &\leq 5^{2p}\left(\lambda^{2p}\mathbb{E}\left[|h(\overline{\Theta}_{\lfloor t\rfloor}^\lambda)|^{2p}\right] + \lambda^{4p}\mathbb{E}\left[|H(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)h(\overline{\Theta}_{\lfloor t\rfloor}^\lambda)|^{2p}\right]\right. \\ &\qquad + \lambda^{4p}\beta^{-2p}\mathbb{E}\left[|\Upsilon(\overline{\Theta}_{\lfloor t\rfloor}^\lambda)|^{2p}\right] + \lambda^p(2\beta^{-1})^p\mathbb{E}\left[\left|\int_{\lfloor t\rfloor}^t\mathrm{d}B_s^\lambda\right|^{2p}\right] \\ &\qquad\left. + \lambda^{3p}(2\beta^{-1})^p\mathbb{E}\left[\left|H(\overline{\Theta}_{\lfloor t\rfloor}^\lambda)\int_{\lfloor t\rfloor}^t\int_{\lfloor s\rfloor}^s\mathrm{d}B_r^\lambda\,\mathrm{d}s\right|^{2p}\right]\right) \\ &\leq 5^{2p}\lambda^p\left(\overline{\mathsf{K}}_1^{2p}\mathbb{E}\left[(1+|\overline{\Theta}_{\lfloor t\rfloor}^\lambda|)^{2p}\right] + (\overline{L}_3\overline{\mathsf{K}}_1)^{2p}\mathbb{E}\left[(1+|\overline{\Theta}_{\lfloor t\rfloor}^\lambda|)^{2p}\right] + (\beta^{-1}d\overline{L}_2)^{2p}\right. \\ &\qquad\left. + (2\beta^{-1}(p+1)(d+2p))^p + (2\beta^{-1}\overline{L}_3^2(p+1)(d+2p))^p\right) \\ &\leq \lambda^p 10^{2p}(\overline{\mathsf{K}}_1^{2p} + (\overline{L}_3\overline{\mathsf{K}}_1)^{2p})\left(\mathbb{E}\left[|\overline{\Theta}_{\lfloor t\rfloor}^\lambda|^{2\lceil p\rceil}\right] + 1\right) \\ &\quad + \lambda^p 5^{2p}((\beta^{-1}d\overline{L}_2)^{2p} + (2\beta^{-1}(p+1)(d+2p))^p + (2\beta^{-1}\overline{L}_3^2(p+1)(d+2p))^p) \\ &\leq \lambda^p\left(e^{-\lambda\bar{a}\lfloor t\rfloor}\overline{\mathsf{C}}_{\mathbf{S}0,p}\mathbb{E}[|\theta_0|^{2\lceil p\rceil}] + \widetilde{\mathsf{C}}_{\mathbf{S}0,p}\right),\end{aligned}$$

where the first inequality holds due to $(\sum_{l=1}^v u_l)^w \leq v^w\sum_{l=1}^v u_l^w$, $v \in \mathbb{N}$, $u_l \geq 0$, $w > 0$, the second inequality holds due to Remark 3.2, Cauchy-Schwarz inequality, and the following inequality:

$$\mathbb{E}\left[\left|\int_{\lfloor t\rfloor}^t\mathrm{d}B_s^\lambda\right|^{2p}\right] \leq \max\{d^p, (p(d+2p-2))^p\} \leq ((p+1)(d+2p))^p,$$

the fourth inequality holds due to Lemma D.3, and where

$$\overline{\mathsf{C}}_{\mathbf{S}0,p} := 10^{2p}(\overline{\mathsf{K}}_1^{2p} + (\overline{L}_3\overline{\mathsf{K}}_1)^{2p}),$$
$$\widetilde{\mathsf{C}}_{\mathbf{S}0,p} := \overline{\mathsf{C}}_{\mathbf{S}0,p}\left(\bar{\mathsf{c}}_{\lceil p\rceil}(1+1/\bar{a})+1\right) + 5^{2p}\left((\beta^{-1}d\overline{L}_2)^{2p} + (2\beta^{-1}(p+1)(d+2p))^p(1+\overline{L}_3^{2p})\right). \tag{149}$$

The inequality (148) can be obtained by using similar arguments. More precisely, by using Definition D.2 with (110), we obtain that, for any $t \geq 0$,

$$\begin{aligned}
\mathbb{E}\left[|\widetilde{\zeta}_t^{\lambda,n} - \widetilde{\zeta}_{\lfloor t\rfloor}^{\lambda,n}|^{2p}\right] &= \mathbb{E}\left[\left|-\lambda\int_{\lfloor t\rfloor}^t h(\widetilde{\zeta}_s^{\lambda,n})\,\mathrm{d}s + \sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t \mathrm{d}B_s^\lambda\right|^{2p}\right]\\
&\leq 2^{2p}\left(\lambda^{2p}\mathbb{E}\left[\int_{\lfloor t\rfloor}^t |h(\widetilde{\zeta}_s^{\lambda,n})|^{2p}\,\mathrm{d}s\right] + \lambda^p(2\beta^{-1})^p\mathbb{E}\left[\left|\int_{\lfloor t\rfloor}^t \mathrm{d}B_s^\lambda\right|^{2p}\right]\right)\\
&\leq 2^{2p}\lambda^p\left(\overline{\mathsf{K}}_1^{2p}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[(1+|\widetilde{\zeta}_s^{\lambda,n}|)^{2\lceil p\rceil}\right]\mathrm{d}s + (2\beta^{-1}(p+1)(d+2p))^p\right)\\
&\leq \lambda^p\left(2^{3\lceil p\rceil}\overline{\mathsf{K}}_1^{2p}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[V_{2\lceil p\rceil}(\widetilde{\zeta}_s^{\lambda,n})\right]\mathrm{d}s + 2^{2p}(2\beta^{-1}(p+1)(d+2p))^p\right)\\
&\leq \lambda^p\left(e^{-\lambda\bar{a}\lfloor t\rfloor/2}\overline{\mathsf{C}}_{\mathbf{S}1,p}\mathbb{E}[|\theta_0|^{2\lceil p\rceil}] + \widetilde{\mathsf{C}}_{\mathbf{S}1,p}\right),
\end{aligned}$$

where the second inequality holds due to Remark 3.2 and the last inequality holds due to Lemma D.5 and where

$$\begin{aligned}
\overline{\mathsf{C}}_{\mathbf{S}1,p} &:= 2^{4\lceil p\rceil}(1+\overline{\mathsf{K}}_1)^{2p},\\
\widetilde{\mathsf{C}}_{\mathbf{S}1,p} &:= \overline{\mathsf{C}}_{\mathbf{S}1,p}\left(\bar{\mathsf{c}}_{\lceil p\rceil}(1+1/\bar{a})+1\right) + 2^{3\lceil p\rceil}(1+\overline{\mathsf{K}}_1)^{2p}3\mathrm{v}_{2\lceil p\rceil}(\overline{\mathsf{M}}_V(2\lceil p\rceil))\\
&\quad + 2^{2p}(2\beta^{-1}(p+1)(d+2p))^p.
\end{aligned} \tag{150}$$

This completes the proof. □

Lemma D.11. *Let Assumptions 4, 5, and 6 hold. Then, for any $0 < \lambda \leq \lambda_{\max}$, $n \in \mathbb{N}_0$, , $t \geq nT$, we obtain the following inequalities:*

$$\mathbb{E}\left[\left|-\lambda\int_{\lfloor t\rfloor}^t \left(H(\overline{\Theta}_s^\lambda) - H(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\right)h(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\,\mathrm{d}s\right|^2\right] \leq \lambda^3\left(e^{-\bar{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2}\right),$$

$$\mathbb{E}\left[\left|\lambda^2\int_{\lfloor t\rfloor}^t H(\overline{\Theta}_s^\lambda)\int_{\lfloor s\rfloor}^s H(\overline{\Theta}_{\lfloor r\rfloor}^\lambda))h(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}r\,\mathrm{d}s\right|^2\right] \leq \lambda^3\left(e^{-\bar{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2}\right),$$

$$\mathbb{E}\left[\left|-\lambda^2\beta^{-1}\int_{\lfloor t\rfloor}^t H(\overline{\Theta}_s^\lambda)\int_{\lfloor s\rfloor}^s \Upsilon(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}r\,\mathrm{d}s\right|^2\right] \leq \lambda^3\left(e^{-\bar{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2}\right),$$

$$\mathbb{E}\left[\left|-\lambda\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t H(\overline{\Theta}_s^\lambda)\int_{\lfloor s\rfloor}^s H(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\,\mathrm{d}s\right|^2\right] \leq \lambda^3\left(e^{-\bar{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2}\right),$$

$$\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t \left(H(\overline{\Theta}_s^\lambda) - H(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\right)\mathrm{d}B_s^\lambda\right|^2\right] \leq \lambda^2\left(e^{-\bar{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2}\right),$$

$$\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t \left(H(\widetilde{\zeta}_s^{\lambda,n}) - H(\widetilde{\zeta}_{\lfloor s\rfloor}^{\lambda,n})\right)\mathrm{d}B_s^\lambda\right|^2\right] \leq \lambda^2\left(e^{-\bar{a}n/4}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2}\right),$$

$$\mathbb{E}\left[\left|\lambda\beta^{-1}\int_{\lfloor t\rfloor}^t \left(\Upsilon(\overline{\Theta}_s^\lambda) - \Upsilon(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\right)\mathrm{d}s\right|^2\right] \leq \lambda^{2+q}\left(e^{-\bar{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2}\right),$$

$$\mathbb{E}\left[\left|-\lambda\int_{\lfloor t\rfloor}^{t} H(\overline{\Theta}_s^\lambda)h(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\,\mathrm{d}s\right|^2+\left|\lambda\beta^{-1}\int_{\lfloor t\rfloor}^{t}\Upsilon(\overline{\Theta}_s^\lambda)\mathrm{d}s\right|^2\right]\le\lambda^2\left(e^{-\bar{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4]+\widetilde{\mathsf{C}}_{\mathbf{S}2}\right),$$

$$\mathbb{E}\left[\left|-\lambda\int_{\lfloor t\rfloor}^{t} H(\widetilde{\zeta}_s^{\lambda,n})h(\widetilde{\zeta}_s^{\lambda,n})\,\mathrm{d}s\right|^2+\left|\lambda\beta^{-1}\int_{\lfloor t\rfloor}^{t}\Upsilon(\widetilde{\zeta}_s^{\lambda,n})\mathrm{d}s\right|^2\right]\le\lambda^2\left(e^{-\bar{a}n/4}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4]+\widetilde{\mathsf{C}}_{\mathbf{S}2}\right),$$

where $\overline{\mathsf{C}}_{\mathbf{S}2}$ *and* $\widetilde{\mathsf{C}}_{\mathbf{S}2}$ *are given in* (151).

Proof. The inequalities can be obtained by using the following arguments:

(i) To show that the first inequality holds, by using Assumption 5, Remark 3.2, we have that

$$\begin{aligned}
&\mathbb{E}\left[\left|-\lambda\int_{\lfloor t\rfloor}^{t}\left(H(\overline{\Theta}_s^\lambda)-H(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\right)h(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\,\mathrm{d}s\right|^2\right]\\
&\le\lambda^2\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[|H(\overline{\Theta}_s^\lambda)-H(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)|^2|h(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)|^2\right]\,\mathrm{d}s\\
&\le\lambda^2\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[\overline{L}_2^2|\overline{\Theta}_s^\lambda-\overline{\Theta}_{\lfloor s\rfloor}^\lambda|^2\overline{\mathsf{K}}_1^2(1+|\overline{\Theta}_{\lfloor s\rfloor}^\lambda|)^2\right]\,\mathrm{d}s\\
&\le 2^{3/2}\lambda^2(\overline{L}_2\overline{\mathsf{K}}_1)^2\int_{\lfloor t\rfloor}^{t}\left(\mathbb{E}\left[1+|\overline{\Theta}_{\lfloor s\rfloor}^\lambda|^4\right]\right)^{1/2}\left(\mathbb{E}\left[|\overline{\Theta}_s^\lambda-\overline{\Theta}_{\lfloor s\rfloor}^\lambda|^4\right]\right)^{1/2}\,\mathrm{d}s\\
&\le 2^{3/2}\lambda^2(\overline{L}_2\overline{\mathsf{K}}_1)^2\left(e^{-\lambda\bar{a}\lfloor t\rfloor}\mathbb{E}\left[|\theta_0|^4\right]+\bar{\mathsf{c}}_2(1+1/\bar{a})+1\right)^{1/2}\\
&\quad\times\lambda\left(e^{-\lambda\bar{a}\lfloor t\rfloor}\overline{\mathsf{C}}_{\mathbf{S}0,2}\mathbb{E}[|\theta_0|^4]+\widetilde{\mathsf{C}}_{\mathbf{S}0,2}\right)^{1/2}\\
&\le\lambda^3 3(\overline{L}_2\overline{\mathsf{K}}_1)^2\left(e^{-\lambda\bar{a}\lfloor t\rfloor}(1+\overline{\mathsf{C}}_{\mathbf{S}0,2})\mathbb{E}\left[|\theta_0|^4\right]+\bar{\mathsf{c}}_2(1+1/\bar{a})+\widetilde{\mathsf{C}}_{\mathbf{S}0,2}+1\right)\\
&\le\lambda^3\left(e^{-\bar{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}\left[|\theta_0|^4\right]+\widetilde{\mathsf{C}}_{\mathbf{S}2}\right),
\end{aligned}$$

where the fourth inequality holds by applying Lemma D.3 and D.10, the last inequality holds due to $\lambda\lfloor t\rfloor\ge\lambda nT\ge n/2$, and where

$$\begin{aligned}
\overline{\mathsf{C}}_{\mathbf{S}2}&:=2^7d^4(1+\beta^{-1})^2(1+\overline{L}_3)^4(1+\overline{L}_2)^2(1+\overline{L}_1)^2(1+\overline{\mathsf{K}}_1)^2(1+\overline{\mathsf{K}}_0)^2\\
&\quad\times(1+\max\{\overline{\mathsf{C}}_{\mathbf{S}0,2},\overline{\mathsf{C}}_{\mathbf{S}0,2q},\overline{\mathsf{C}}_{\mathbf{S}0,1+q},\overline{\mathsf{C}}_{\mathbf{S}1,2}\}),\\
\widetilde{\mathsf{C}}_{\mathbf{S}2}&:=2^7d^4(1+\beta^{-1})^2(1+\overline{L}_3)^4(1+\overline{L}_2)^2(1+\overline{L}_1)^2(1+\overline{\mathsf{K}}_1)^2(1+\overline{\mathsf{K}}_0)^2\\
&\quad\times(\bar{\mathsf{c}}_2(1+1/\bar{a})+\max\{\widetilde{\mathsf{C}}_{\mathbf{S}0,2},\widetilde{\mathsf{C}}_{\mathbf{S}0,2q},\widetilde{\mathsf{C}}_{\mathbf{S}0,1+q},\widetilde{\mathsf{C}}_{\mathbf{S}1,2}\}+1)
\end{aligned}\tag{151}$$

with $\overline{\mathsf{C}}_{\mathbf{S}0,p}$, $\widetilde{\mathsf{C}}_{\mathbf{S}0,p}$, $\overline{\mathsf{C}}_{\mathbf{S}1,p}$, $\widetilde{\mathsf{C}}_{\mathbf{S}1,p}$, $p>0$, given in (149) and (150), and $\bar{\mathsf{c}}_p$, $\overline{\mathsf{M}}_V(p)$, $p\in[2,\infty)\cap\mathbb{N}$ given in (145) (see also Lemma D.3) and Lemma D.4.

(ii) To establish the second inequality, we apply Remark 3.2 and Lemma D.3 to obtain

$$\begin{aligned}
&\mathbb{E}\left[\left|\lambda^2\int_{\lfloor t\rfloor}^{t}H(\overline{\Theta}_s^\lambda)\int_{\lfloor s\rfloor}^{s}H(\overline{\Theta}_{\lfloor r\rfloor}^\lambda))h(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}r\,\mathrm{d}s\right|^2\right]\\
&\le\lambda^4\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[|H(\overline{\Theta}_s^\lambda)H(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)h(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)|^2\right]\,\mathrm{d}s\\
&\le\lambda^4\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[\overline{L}_3^4\overline{\mathsf{K}}_1^2(1+|\overline{\Theta}_{\lfloor s\rfloor}^\lambda|)^2\right]\,\mathrm{d}s\\
&\le 2\lambda^4\overline{L}_3^4\overline{\mathsf{K}}_1^2\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[1+|\overline{\Theta}_{\lfloor s\rfloor}^\lambda|^2\right]\,\mathrm{d}s\\
&\le 4\lambda^4\overline{L}_3^4\overline{\mathsf{K}}_1^2\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[1+|\overline{\Theta}_{\lfloor s\rfloor}^\lambda|^4\right]\,\mathrm{d}s
\end{aligned}$$

$$
\begin{aligned}
&\leq \lambda^3 4\overline{L}_3^4 \overline{\mathsf{K}}_1^2 \left(e^{-\lambda \overline{a}\lfloor t\rfloor}\mathbb{E}\left[|\theta_0|^4\right] + \overline{\mathsf{c}}_2(1+1/\overline{a})+1\right)\\
&\leq \lambda^3 \left(e^{-\overline{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}\left[|\theta_0|^4\right] + \widetilde{\mathsf{C}}_{\mathbf{S}2}\right),
\end{aligned}
$$

where $\overline{\mathsf{C}}_{\mathbf{S}2}$ and $\widetilde{\mathsf{C}}_{\mathbf{S}2}$ are given in (151).

(iii) To obtain the third inequality, we apply Remark 3.2 and write

$$
\begin{aligned}
&\mathbb{E}\left[\left|-\lambda^2\beta^{-1}\int_{\lfloor t\rfloor}^t H(\overline{\Theta}_s^\lambda)\int_{\lfloor s\rfloor}^s \Upsilon(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}r\,\mathrm{d}s\right|^2\right]\\
&\leq \lambda^4\beta^{-2}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[|H(\overline{\Theta}_s^\lambda)\Upsilon(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)|^2\right]\mathrm{d}s\\
&\leq \lambda^4\beta^{-2}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[\overline{L}_3^2|\Upsilon(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)|^2\right]\mathrm{d}s\\
&\leq \lambda^4(\beta^{-1}\overline{L}_3 d\overline{L}_2)^2\\
&\leq \lambda^3\left(e^{-\overline{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2}\right),
\end{aligned}
$$

where $\overline{\mathsf{C}}_{\mathbf{S}2}$ and $\widetilde{\mathsf{C}}_{\mathbf{S}2}$ are given in (151).

(iv) To obtain the fourth inequality, we use Cauchy-Schwarz inequality and Lemma D.3:

$$
\begin{aligned}
&\mathbb{E}\left[\left|-\lambda\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t H(\overline{\Theta}_s^\lambda)\int_{\lfloor s\rfloor}^s H(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\,\mathrm{d}s\right|^2\right]\\
&\leq 2\lambda^3\beta^{-1}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[\overline{L}_3^2\left|H(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\int_{\lfloor s\rfloor}^s \mathrm{d}B_r^\lambda\right|^2\right]\mathrm{d}s\\
&\leq 2\lambda^3\beta^{-1}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[\overline{L}_3^4\left|\int_{\lfloor s\rfloor}^s \mathrm{d}B_r^\lambda\right|^2\right]\mathrm{d}s\\
&\leq 2\lambda^3\beta^{-1}\overline{L}_3^4 d\\
&\leq \lambda^3\left(e^{-\overline{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2}\right),
\end{aligned}
$$

where $\overline{\mathsf{C}}_{\mathbf{S}2}$ and $\widetilde{\mathsf{C}}_{\mathbf{S}2}$ are given in (151).

(v) To obtain the fifth inequality, we apply Assumption 5 and Cauchy-Schwarz inequality:

$$
\begin{aligned}
&\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t \left(H(\overline{\Theta}_s^\lambda) - H(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\right)\mathrm{d}B_s^\lambda\right|^2\right]\\
&= 2\lambda\beta^{-1}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[|H(\overline{\Theta}_s^\lambda) - H(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)|_{\mathsf{F}}^2\right]\mathrm{d}s\\
&\leq 2\lambda\beta^{-1}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[d\overline{L}_2^2|\overline{\Theta}_s^\lambda - \overline{\Theta}_{\lfloor s\rfloor}^\lambda|^2\right]\mathrm{d}s\\
&\leq 2\lambda\beta^{-1}d\overline{L}_2^2\int_{\lfloor t\rfloor}^t \left(\mathbb{E}\left[|\overline{\Theta}_s^\lambda - \overline{\Theta}_{\lfloor s\rfloor}^\lambda|^4\right]\right)^{1/2}\mathrm{d}s\\
&\leq 2\lambda^2\beta^{-1}d\overline{L}_2^2\left(e^{-\lambda\overline{a}\lfloor t\rfloor}\overline{\mathsf{C}}_{\mathbf{S}0,2}\mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}0,2}\right)^{1/2}\\
&\leq \lambda^2\left(e^{-\overline{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2}\right),
\end{aligned}
$$

where the third inequality holds due to Lemma D.10, and where $\overline{\mathsf{C}}_{\mathbf{S}2}$ and $\widetilde{\mathsf{C}}_{\mathbf{S}2}$ are given in (151).

(vi) To obtain the sixth inequality, we use the same arguments as in (v):

$$
\begin{aligned}
&\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^{t}\left(H(\widetilde{\zeta}_s^{\lambda,n})-H(\widetilde{\zeta}_{\lfloor s\rfloor}^{\lambda,n})\right)\,\mathrm{d}B_s^{\lambda}\right|^2\right]\\
&=2\lambda\beta^{-1}\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[|H(\widetilde{\zeta}_s^{\lambda,n})-H(\widetilde{\zeta}_{\lfloor s\rfloor}^{\lambda,n})|_{\mathsf{F}}^2\right]\,\mathrm{d}s\\
&\leq 2\lambda\beta^{-1}\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[d\overline{L}_2^2|\widetilde{\zeta}_s^{\lambda,n}-\widetilde{\zeta}_{\lfloor s\rfloor}^{\lambda,n}|^2\right]\,\mathrm{d}s\\
&\leq 2\lambda\beta^{-1}d\overline{L}_2^2\int_{\lfloor t\rfloor}^{t}\left(\mathbb{E}\left[|\widetilde{\zeta}_s^{\lambda,n}-\widetilde{\zeta}_{\lfloor s\rfloor}^{\lambda,n}|^4\right]\right)^{1/2}\,\mathrm{d}s\\
&\leq 2\lambda^2\beta^{-1}d\overline{L}_2^2\left(e^{-\lambda\bar{a}\lfloor t\rfloor/2}\overline{\mathsf{C}}_{\mathbf{S}1,2}\mathbb{E}[|\theta_0|^4]+\widetilde{\mathsf{C}}_{\mathbf{S}1,2}\right)^{1/2}\\
&\leq \lambda^2\left(e^{-\bar{a}n/4}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4]+\widetilde{\mathsf{C}}_{\mathbf{S}2}\right),
\end{aligned}
$$

where the third inequality holds due to Lemma D.10, and where $\overline{\mathsf{C}}_{\mathbf{S}2}$ and $\widetilde{\mathsf{C}}_{\mathbf{S}2}$ are given in (151).

(vii) To establish the seventh inequality, we apply Remark 3.2 and Cauchy-Schwarz inequality to obtain

$$
\begin{aligned}
&\mathbb{E}\left[\left|\lambda\beta^{-1}\int_{\lfloor t\rfloor}^{t}\left(\Upsilon(\overline{\Theta}_s^{\lambda})-\Upsilon(\overline{\Theta}_{\lfloor s\rfloor}^{\lambda})\right)\,\mathrm{d}s\right|^2\right]\\
&\leq \lambda^2\beta^{-2}\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[|\Upsilon(\overline{\Theta}_s^{\lambda})-\Upsilon(\overline{\Theta}_{\lfloor s\rfloor}^{\lambda})|^2\right]\,\mathrm{d}s\\
&\leq \lambda^2\beta^{-2}\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[d^3\overline{L}_1^2|\overline{\Theta}_s^{\lambda}-\overline{\Theta}_{\lfloor s\rfloor}^{\lambda}|^{2q}\right]\,\mathrm{d}s\\
&\leq \lambda^2\beta^{-2}d^3\overline{L}_1^2\int_{\lfloor t\rfloor}^{t}\left(\mathbb{E}\left[|\overline{\Theta}_s^{\lambda}-\overline{\Theta}_{\lfloor s\rfloor}^{\lambda}|^{4q}\right]\right)^{1/2}\,\mathrm{d}s\\
&\leq \lambda^{2+q}\beta^{-2}d^3\overline{L}_1^2\left(e^{-\lambda\bar{a}\lfloor t\rfloor}\overline{\mathsf{C}}_{\mathbf{S}0,2q}\mathbb{E}[|\theta_0|^4]+\widetilde{\mathsf{C}}_{\mathbf{S}0,2q}\right)^{1/2}\\
&\leq \lambda^{2+q}\left(e^{-\bar{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4]+\widetilde{\mathsf{C}}_{\mathbf{S}2}\right),
\end{aligned}
$$

where the fourth inequality holds due to Lemma D.10, and $\overline{\mathsf{C}}_{\mathbf{S}2}$, $\widetilde{\mathsf{C}}_{\mathbf{S}2}$ are given in (151).

(viii) To obtain the eighth inequality, we apply Remark 3.2 and write

$$
\begin{aligned}
&\mathbb{E}\left[\left|-\lambda\int_{\lfloor t\rfloor}^{t}H(\overline{\Theta}_s^{\lambda})h(\overline{\Theta}_{\lfloor s\rfloor}^{\lambda})\,\mathrm{d}s\right|^2+\left|\lambda\beta^{-1}\int_{\lfloor t\rfloor}^{t}\Upsilon(\overline{\Theta}_s^{\lambda})\,\mathrm{d}s\right|^2\right]\\
&\leq \lambda^2\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[\overline{L}_3^2|h(\overline{\Theta}_{\lfloor s\rfloor}^{\lambda})|^2\right]\,\mathrm{d}s+\lambda^2\beta^{-2}d^2\overline{L}_2^2\\
&\leq \lambda^2\overline{L}_3^2\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[\overline{\mathsf{K}}_1^2(1+|\overline{\Theta}_{\lfloor s\rfloor}^{\lambda}|)^2\right]\,\mathrm{d}s+\lambda^2\beta^{-2}d^2\overline{L}_2^2\\
&\leq \lambda^2\left(4\overline{L}_3^2\overline{\mathsf{K}}_1^2\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[(1+|\overline{\Theta}_{\lfloor s\rfloor}^{\lambda}|^4)\right]\,\mathrm{d}s+\beta^{-2}d^2\overline{L}_2^2\right)\\
&\leq \lambda^2\left(4\overline{L}_3^2\overline{\mathsf{K}}_1^2\left(e^{-\lambda\bar{a}\lfloor t\rfloor}\mathbb{E}\left[|\theta_0|^4\right]+\bar{\mathsf{c}}_2\left(1+1/\bar{a}\right)+1\right)+\beta^{-2}d^2\overline{L}_2^2\right)\\
&\leq \lambda^2 4(1+\overline{L}_3)^2(1+\overline{\mathsf{K}}_1)^2(1+\beta^{-1})^2d^2(1+\overline{L}_2)^2\left(e^{-\lambda\bar{a}\lfloor t\rfloor}\mathbb{E}\left[|\theta_0|^4\right]+\bar{\mathsf{c}}_2\left(1+1/\bar{a}\right)+1\right)\\
&\leq \lambda^2\left(e^{-\bar{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4]+\widetilde{\mathsf{C}}_{\mathbf{S}2}\right),
\end{aligned}
$$

where the fourth inequality holds due to Lemma D.3, and where $\overline{\mathsf{C}}_{\mathbf{S}2}$, $\widetilde{\mathsf{C}}_{\mathbf{S}2}$ are given in (151).

(ix) To establish the last inequality, we follow the arguments in (viii):

$$
\begin{aligned}
&\mathbb{E}\left[\left|-\lambda \int_{\lfloor t\rfloor}^{t} H(\widetilde{\zeta}_s^{\lambda,n})h(\widetilde{\zeta}_s^{\lambda,n})\,\mathrm{d}s\right|^2+\left|\lambda\beta^{-1}\int_{\lfloor t\rfloor}^{t}\Upsilon(\widetilde{\zeta}_s^{\lambda,n})\mathrm{d}s\right|^2\right]\\
&\leq \lambda^2\left(2\overline{L}_3^2\overline{\mathsf{K}}_1^2\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[(1+|\widetilde{\zeta}_s^{\lambda,n}|^2)\right]\,\mathrm{d}s+\beta^{-2}d^2\overline{L}_2^2\right)\\
&\leq \lambda^2\left(2\overline{L}_3^2\overline{\mathsf{K}}_1^2\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[V_4(\widetilde{\zeta}_s^{\lambda,n})\right]\,\mathrm{d}s+\beta^{-2}d^2\overline{L}_2^2\right)\\
&\leq \lambda^2\left(2\overline{L}_3^2\overline{\mathsf{K}}_1^2\left(2e^{-\lambda\overline{a}\lfloor t\rfloor/2}\mathbb{E}[|\theta_0|^4]+2\left(\overline{\mathsf{c}}_2(1+1/\overline{a})+1\right)+3\mathrm{v}_4(\overline{\mathsf{M}}_V(4))\right)+\beta^{-2}d^2\overline{L}_2^2\right)\\
&\leq \lambda^2 4(1+\overline{L}_3)^2(1+\overline{\mathsf{K}}_1)^2(1+\beta^{-1})^2d^2(1+\overline{L}_2)^2\\
&\quad\times\left(e^{-\lambda\overline{a}\lfloor t\rfloor/2}\mathbb{E}\left[|\theta_0|^4\right]+\overline{\mathsf{c}}_2\left(1+1/\overline{a}\right)+3\mathrm{v}_4(\overline{\mathsf{M}}_V(4))/2+2\right)\\
&\leq \lambda^2\left(e^{-\overline{a}n/4}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4]+\widetilde{\mathsf{C}}_{\mathbf{S}2}\right),
\end{aligned}
$$

where the second inequality holds due to Lemma D.5, and where $\overline{\mathsf{C}}_{\mathbf{S}2}$ and $\widetilde{\mathsf{C}}_{\mathbf{S}2}$ are given in (151). This completes the proof. ☐

Corollary D.12. *Let Assumptions 4, 5, and 6 hold. Then, for any $0<\lambda\leq\overline{\lambda}_{\max}$, $n\in\mathbb{N}_0$, , $t\geq nT$, we obtain the following inequalities:*

$$
\begin{aligned}
&\mathbb{E}\left[\left|h(\overline{\Theta}_t^{\lambda})-h(\overline{\Theta}_{\lfloor t\rfloor}^{\lambda})+\lambda\int_{\lfloor t\rfloor}^{t}\left(H(\overline{\Theta}_{\lfloor s\rfloor}^{\lambda})h(\overline{\Theta}_{\lfloor s\rfloor}^{\lambda})-\beta^{-1}\Upsilon(\overline{\Theta}_{\lfloor s\rfloor}^{\lambda})\right)\,\mathrm{d}s\right.\right.\\
&\qquad\left.\left.-\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^{t}H(\overline{\Theta}_{\lfloor s\rfloor}^{\lambda})\,\mathrm{d}B_s^{\lambda}\right|^2\right]\leq\lambda^2\left(e^{-\overline{a}n/2}36\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4]+36\widetilde{\mathsf{C}}_{\mathbf{S}2}\right),\\
&\mathbb{E}\left[\left|h(\widetilde{\zeta}_t^{\lambda,n})-h(\widetilde{\zeta}_{\lfloor t\rfloor}^{\lambda,n})-\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^{t}H(\widetilde{\zeta}_{\lfloor s\rfloor}^{\lambda,n})\,\mathrm{d}B_s^{\lambda}-\left(h(\overline{\Theta}_t^{\lambda})-h(\overline{\Theta}_{\lfloor t\rfloor}^{\lambda})\right.\right.\right.\\
&\qquad\left.\left.\left.-\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^{t}H(\overline{\Theta}_{\lfloor s\rfloor}^{\lambda})\,\mathrm{d}B_s^{\lambda}\right)\right|^2\right]\leq\lambda^2\left(e^{-\overline{a}n/4}72\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4]+72\widetilde{\mathsf{C}}_{\mathbf{S}2}\right),
\end{aligned}
$$

where $\overline{\mathsf{C}}_{\mathbf{S}2}$ and $\widetilde{\mathsf{C}}_{\mathbf{S}2}$ are given in (151).

Proof. For any $t\geq nT$, by applying Itô's formula to $h(\overline{\Theta}_t^{\lambda})$, we obtain, almost surely

$$
\begin{aligned}
h(\overline{\Theta}_t^{\lambda})-h(\overline{\Theta}_{\lfloor t\rfloor}^{\lambda})=&-\lambda\int_{\lfloor t\rfloor}^{t}H(\overline{\Theta}_s^{\lambda})h(\overline{\Theta}_{\lfloor s\rfloor}^{\lambda})\,\mathrm{d}s+\lambda^2\int_{\lfloor t\rfloor}^{t}H(\overline{\Theta}_s^{\lambda})\int_{\lfloor s\rfloor}^{s}H(\overline{\Theta}_{\lfloor r\rfloor}^{\lambda})h(\overline{\Theta}_{\lfloor r\rfloor}^{\lambda})\,\mathrm{d}r\,\mathrm{d}s\\
&-\lambda^2\beta^{-1}\int_{\lfloor t\rfloor}^{t}H(\overline{\Theta}_s^{\lambda})\int_{\lfloor s\rfloor}^{s}\Upsilon(\overline{\Theta}_{\lfloor r\rfloor}^{\lambda})\,\mathrm{d}r\,\mathrm{d}s\\
&-\lambda\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^{t}H(\overline{\Theta}_s^{\lambda})\int_{\lfloor s\rfloor}^{s}H(\overline{\Theta}_{\lfloor r\rfloor}^{\lambda})\,\mathrm{d}B_r^{\lambda}\,\mathrm{d}s\\
&+\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^{t}H(\overline{\Theta}_s^{\lambda})\,\mathrm{d}B_s^{\lambda}+\lambda\beta^{-1}\int_{\lfloor t\rfloor}^{t}\Upsilon(\overline{\Theta}_s^{\lambda})\mathrm{d}s.
\end{aligned}
\tag{152}
$$

Similarly, applying Itô's formula to $h(\widetilde{\zeta}_t^{\lambda,n})$ yields, almost surely

$$
\begin{aligned}
h(\widetilde{\zeta}_t^{\lambda,n})-h(\widetilde{\zeta}_{\lfloor t\rfloor}^{\lambda,n})=&-\lambda\int_{\lfloor t\rfloor}^{t}H(\widetilde{\zeta}_s^{\lambda,n})h(\widetilde{\zeta}_s^{\lambda,n})\,\mathrm{d}s+\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^{t}H(\widetilde{\zeta}_s^{\lambda,n})\,\mathrm{d}B_s^{\lambda}\\
&+\lambda\beta^{-1}\int_{\lfloor t\rfloor}^{t}\Upsilon(\widetilde{\zeta}_s^{\lambda,n})\mathrm{d}s.
\end{aligned}
\tag{153}
$$

(i) To obtain the first inequality, by using (152) with $(\sum_{l=1}^{v} u_l)^2 \leq v\sum_{l=1}^{v} u_l^2$, $v \in \mathbb{N}$, $u_l \geq 0$, we obtain that

$$\begin{aligned}
&\mathbb{E}\left[\left|h(\overline{\Theta}_t^\lambda) - h(\overline{\Theta}_{\lfloor t\rfloor}^\lambda) + \lambda \int_{\lfloor t\rfloor}^t \left(H(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)h(\overline{\Theta}_{\lfloor s\rfloor}^\lambda) - \beta^{-1}\Upsilon(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\right)\,\mathrm{d}s\right.\right.\\
&\qquad\left.\left. -\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t H(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\,\mathrm{d}B_s^\lambda\right|^2\right]\\
&\leq 6\mathbb{E}\left[\left|-\lambda\int_{\lfloor t\rfloor}^t \left(H(\overline{\Theta}_s^\lambda) - H(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\right)h(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\,\mathrm{d}s\right|^2\right]\\
&\quad + 6\mathbb{E}\left[\left|\lambda^2\int_{\lfloor t\rfloor}^t H(\overline{\Theta}_s^\lambda)\int_{\lfloor s\rfloor}^s H(\overline{\Theta}_{\lfloor r\rfloor}^\lambda))h(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}r\,\mathrm{d}s\right|^2\right]\\
&\quad + 6\mathbb{E}\left[\left|-\lambda^2\beta^{-1}\int_{\lfloor t\rfloor}^t H(\overline{\Theta}_s^\lambda)\int_{\lfloor s\rfloor}^s \Upsilon(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}r\,\mathrm{d}s\right|^2\right]\\
&\quad + 6\mathbb{E}\left[\left|-\lambda\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t H(\overline{\Theta}_s^\lambda)\int_{\lfloor s\rfloor}^s H(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\,\mathrm{d}s\right|^2\right]\\
&\quad + 6\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t \left(H(\overline{\Theta}_s^\lambda) - H(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\right)\,\mathrm{d}B_s^\lambda\right|^2\right]\\
&\quad + 6\mathbb{E}\left[\left|\lambda\beta^{-1}\int_{\lfloor t\rfloor}^t \left(\Upsilon(\overline{\Theta}_s^\lambda) - \Upsilon(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\right)\,\mathrm{d}s\right|^2\right]\\
&\leq \lambda^2\left(e^{-\bar{a}n/2}36\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4] + 36\widetilde{\mathsf{C}}_{\mathbf{S}2}\right),
\end{aligned}$$

where the last inequality holds due to Lemma D.11.

(ii) To establish the second inequality, we use (152) and (153) to obtain

$$\begin{aligned}
&\mathbb{E}\left[\left|h(\widetilde{\zeta}_t^{\lambda,n}) - h(\widetilde{\zeta}_{\lfloor t\rfloor}^{\lambda,n}) - \sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t H(\widetilde{\zeta}_{\lfloor s\rfloor}^{\lambda,n})\,\mathrm{d}B_s^\lambda\right.\right.\\
&\qquad\left.\left. - \left(h(\overline{\Theta}_t^\lambda) - h(\overline{\Theta}_{\lfloor t\rfloor}^\lambda) - \sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t H(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\,\mathrm{d}B_s^\lambda\right)\right|^2\right]\\
&\leq 2\mathbb{E}\left[\left|h(\widetilde{\zeta}_t^{\lambda,n}) - h(\widetilde{\zeta}_{\lfloor t\rfloor}^{\lambda,n}) - \sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t H(\widetilde{\zeta}_{\lfloor s\rfloor}^{\lambda,n})\,\mathrm{d}B_s^\lambda\right|^2\right]\\
&\quad + 2\mathbb{E}\left[\left|h(\overline{\Theta}_t^\lambda) - h(\overline{\Theta}_{\lfloor t\rfloor}^\lambda) - \sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t H(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\,\mathrm{d}B_s^\lambda\right|^2\right]\\
&\leq 6\mathbb{E}\left[\left|-\lambda\int_{\lfloor t\rfloor}^t H(\widetilde{\zeta}_s^{\lambda,n})h(\widetilde{\zeta}_s^{\lambda,n})\,\mathrm{d}s\right|^2\right] + 6\mathbb{E}\left[\left|\lambda\beta^{-1}\int_{\lfloor t\rfloor}^t \Upsilon(\widetilde{\zeta}_s^{\lambda,n})\mathrm{d}s\right|^2\right]\\
&\quad + 6\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t \left(H(\widetilde{\zeta}_s^{\lambda,n}) - H(\widetilde{\zeta}_{\lfloor s\rfloor}^{\lambda,n})\right)\,\mathrm{d}B_s^\lambda\right|^2\right]\\
&\quad + 12\mathbb{E}\left[\left|-\lambda\int_{\lfloor t\rfloor}^t H(\overline{\Theta}_s^\lambda)h(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\,\mathrm{d}s\right|^2\right] + 12\mathbb{E}\left[\left|\lambda\beta^{-1}\int_{\lfloor t\rfloor}^t \Upsilon(\overline{\Theta}_s^\lambda)\mathrm{d}s\right|^2\right]
\end{aligned}$$

$$
\begin{aligned}
&+12\mathbb{E}\left[\left|\lambda^2\int_{\lfloor t\rfloor}^t H(\overline{\Theta}_s^\lambda)\int_{\lfloor s\rfloor}^s H(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)h(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}r\,\mathrm{d}s\right|^2\right]\\
&+12\mathbb{E}\left[\left|-\lambda^2\beta^{-1}\int_{\lfloor t\rfloor}^t H(\overline{\Theta}_s^\lambda)\int_{\lfloor s\rfloor}^s \Upsilon(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}r\,\mathrm{d}s\right|^2\right]\\
&+12\mathbb{E}\left[\left|-\lambda\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t H(\overline{\Theta}_s^\lambda)\int_{\lfloor s\rfloor}^s H(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\,\mathrm{d}s\right|^2\right]\\
&+12\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t \left(H(\overline{\Theta}_s^\lambda)-H(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\right)\,\mathrm{d}B_s^\lambda\right|^2\right]\\
&\leq 12\lambda^2\left(e^{-\bar{a}n/4}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4]+\widetilde{\mathsf{C}}_{\mathbf{S}2}\right)+60\lambda^2\left(e^{-\bar{a}n/4}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4]+\widetilde{\mathsf{C}}_{\mathbf{S}2}\right)\\
&\leq \lambda^2\left(e^{-\bar{a}n/4}72\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4]+72\widetilde{\mathsf{C}}_{\mathbf{S}2}\right),
\end{aligned}
$$

where the second last inequality holds by using Lemma D.11.

This completes the proof. ☐

Definition D.13. *Define* $\mathfrak{M}=(\mathfrak{M}^{(i,j)})_{i,j=1,\dots,d}:\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}^{d\times d}$ *by setting, for every* $i,j=1,\dots,d$*,*

$$
\mathfrak{M}^{(i,j)}(\theta,\overline{\theta})=\langle\nabla H^{(i,j)}(\overline{\theta}),\theta-\overline{\theta}\rangle,\qquad \theta,\overline{\theta}\in\mathbb{R}^d.
$$

Lemma D.14. *Let Assumptions 4, 5, and 6 hold. Then, for any* $0<\lambda\leq\overline{\lambda}_{\max}$*,* $n\in\mathbb{N}_0$*, ,* $t\geq nT$*, we obtain the following inequalities:*

$$
\begin{aligned}
&\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t\left(H(\overline{\Theta}_s^\lambda)-H(\overline{\Theta}_{\lfloor s\rfloor}^\lambda)-\mathfrak{M}(\overline{\Theta}_s^\lambda,\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\right)\,\mathrm{d}B_s^\lambda\right|^2\right]\\
&\leq\lambda^{2+q}\left(e^{-\bar{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4]+\widetilde{\mathsf{C}}_{\mathbf{S}2}\right),
\end{aligned}\tag{154}
$$

$$
\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t\mathfrak{M}(\overline{\Theta}_s^\lambda,\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\,\mathrm{d}B_s^\lambda\right|^2\right]\leq\lambda^2\left(e^{-\bar{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4]+\widetilde{\mathsf{C}}_{\mathbf{S}2}\right),\tag{155}
$$

$$
\begin{aligned}
&\mathbb{E}\left[2\lambda\beta^{-1}\left\langle\int_{\lfloor t\rfloor}^t\int_{\lfloor s\rfloor}^s\left(H(\widetilde{\zeta}_{\lfloor r\rfloor}^{\lambda,n})-H(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)\right)\,\mathrm{d}B_r^\lambda\,\mathrm{d}s,\int_{\lfloor t\rfloor}^t\mathfrak{M}(\overline{\Theta}_s^\lambda,\overline{\Theta}_{\lfloor s\rfloor}^\lambda)\,\mathrm{d}B_s^\lambda\right\rangle\right]\\
&\leq\lambda^2\left(e^{-\bar{a}n/4}5\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4]+5\widetilde{\mathsf{C}}_{\mathbf{S}2}\right),
\end{aligned}\tag{156}
$$

where $\overline{\mathsf{C}}_{\mathbf{S}2}$ *and* $\widetilde{\mathsf{C}}_{\mathbf{S}2}$ *are given in (151).*

Proof. To show the inequalities hold, we follow the arguments below:

(i) To establish the first inequality (154), by using Definition D.13, we observe that, for fixed $\theta,\overline{\theta}\in\mathbb{R}^d$,

$$
\begin{aligned}
&|H(\theta)-H(\overline{\theta})-\mathfrak{M}(\theta,\overline{\theta})|_{\mathsf{F}}^2\\
&=\sum_{i,j=1}^d|H^{(i,j)}(\theta)-H^{(i,j)}(\overline{\theta})-\mathfrak{M}^{(i,j)}(\theta,\overline{\theta})|^2\\
&=\sum_{i,j=1}^d\left|\int_0^1\langle\nabla H^{(i,j)}(\nu\theta+(1-\nu)\overline{\theta}),\theta-\overline{\theta}\rangle\,\mathrm{d}\nu-\langle\nabla H^{(i,j)}(\overline{\theta}),\theta-\overline{\theta}\rangle\right|^2\\
&\leq\int_0^1\sum_{i,j=1}^d|\nabla H^{(i,j)}(\nu\theta+(1-\nu)\overline{\theta})-\nabla H^{(i,j)}(\overline{\theta})|^2\,\mathrm{d}\nu|\theta-\overline{\theta}|^2
\end{aligned}
$$

$$
\begin{aligned}
&\leq \int_0^1 \sum_{i=1}^d d|\nabla^2 h^{(i)}(\nu\theta+(1-\nu)\overline{\theta}) - \nabla^2 h^{(i)}(\overline{\theta})|^2 \,\mathrm{d}\nu |\theta-\overline{\theta}|^2 \\
&\leq d^2 \overline{L}_1^2 |\theta-\overline{\theta}|^{2+2q},
\end{aligned} \tag{157}
$$

where the last inequality holds due to Assumption 5. By using (157), we obtain that

$$
\begin{aligned}
&\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t \left(H(\overline{\Theta}_s^\lambda) - H(\overline{\Theta}_{\lfloor s\rfloor}^\lambda) - \mathfrak{M}(\overline{\Theta}_s^\lambda, \overline{\Theta}_{\lfloor s\rfloor}^\lambda)\right) \mathrm{d}B_s^\lambda\right|^2\right] \\
&= 2\lambda\beta^{-1}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[\left|H(\overline{\Theta}_s^\lambda) - H(\overline{\Theta}_{\lfloor s\rfloor}^\lambda) - \mathfrak{M}(\overline{\Theta}_s^\lambda, \overline{\Theta}_{\lfloor s\rfloor}^\lambda)\right|_{\mathsf{F}}^2\right] \mathrm{d}s \\
&\leq \lambda\beta^{-1} 2d^2\overline{L}_1^2 \int_{\lfloor t\rfloor}^t \mathbb{E}\left[|\overline{\Theta}_s^\lambda - \overline{\Theta}_{\lfloor s\rfloor}^\lambda|^{2+2q}\right] \mathrm{d}s \\
&\leq \lambda^{2+q}\beta^{-1}2d^2\overline{L}_1^2\left(e^{-\lambda\overline{a}\lfloor t\rfloor}\overline{\mathsf{C}}_{\mathbf{S}0,1+q}\mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}0,1+q}\right) \\
&\leq \lambda^{2+q}\left(e^{-\overline{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2}\right),
\end{aligned}
$$

where the third inequality holds due to Lemma D.3 and D.10, and $\overline{\mathsf{C}}_{\mathbf{S}2}$, $\widetilde{\mathsf{C}}_{\mathbf{S}2}$ are given in (151).

(ii) To establish the second inequality (155), we recall Definition D.13 and use Remark 3.2 to obtain

$$
\begin{aligned}
&\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}}\int_{\lfloor t\rfloor}^t \mathfrak{M}(\overline{\Theta}_s^\lambda, \overline{\Theta}_{\lfloor s\rfloor}^\lambda)\,\mathrm{d}B_s^\lambda\right|^2\right] \\
&= 2\lambda\beta^{-1}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[\sum_{i,j=1}^d \left|\left\langle \nabla H^{(i,j)}(\overline{\Theta}_{\lfloor s\rfloor}^\lambda), \overline{\Theta}_s^\lambda - \overline{\Theta}_{\lfloor s\rfloor}^\lambda\right\rangle\right|^2\right] \mathrm{d}s \\
&\leq 2^{3-2q}\lambda\beta^{-1}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[d^2\overline{\mathsf{K}}_0^2(1+|\overline{\Theta}_{\lfloor s\rfloor}^\lambda|)^{2q}|\overline{\Theta}_s^\lambda - \overline{\Theta}_{\lfloor s\rfloor}^\lambda|^2\right] \mathrm{d}s \\
&\leq \lambda\beta^{-1}2^5d^2\overline{\mathsf{K}}_0^2\int_{\lfloor t\rfloor}^t \left(\mathbb{E}\left[1+|\overline{\Theta}_{\lfloor s\rfloor}^\lambda|^4\right]\right)^{1/2}\left(\mathbb{E}\left[|\overline{\Theta}_s^\lambda - \overline{\Theta}_{\lfloor s\rfloor}^\lambda|^4\right]\right)^{1/2} \mathrm{d}s \\
&\leq \lambda\beta^{-1}2^5d^2\overline{\mathsf{K}}_0^2\left(e^{-\lambda\overline{a}\lfloor t\rfloor}\mathbb{E}\left[|\theta_0|^4\right] + \mathsf{c}_2\left(1+1/\overline{a}\right)+1\right)^{1/2} \\
&\qquad \times \lambda\left(e^{-\lambda\overline{a}\lfloor t\rfloor}\overline{\mathsf{C}}_{\mathbf{S}0,2}\mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}0,2}\right)^{1/2} \\
&\leq \lambda^2\left(e^{-\overline{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2}\right),
\end{aligned}
$$

where the third inequality holds due to Lemma D.3 and D.10, and $\overline{\mathsf{C}}_{\mathbf{S}2}$, $\widetilde{\mathsf{C}}_{\mathbf{S}2}$ are given in (151).

(iii) To obtain the third inequality (156), we use Definition D.13 and write the following

$$
\mathbb{E}\left[2\lambda\beta^{-1}\left\langle \int_{\lfloor t\rfloor}^t \int_{\lfloor s\rfloor}^s \left(H(\widetilde{\zeta}_{\lfloor r\rfloor}^{\lambda,n}) - H(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)\right) \mathrm{d}B_r^\lambda\,\mathrm{d}s, \int_{\lfloor t\rfloor}^t \mathfrak{M}(\overline{\Theta}_s^\lambda, \overline{\Theta}_{\lfloor s\rfloor}^\lambda)\,\mathrm{d}B_s^\lambda\right\rangle\right]
$$

$$
\begin{aligned}
&= 2\lambda\beta^{-1}\mathbb{E}\left[\left\langle \int_{\lfloor t\rfloor}^{t}\int_{\lfloor s\rfloor}^{s}\left(H(\widetilde{\zeta}^{\lambda,n}_{\lfloor r\rfloor}) - H(\overline{\Theta}^{\lambda}_{\lfloor r\rfloor})\right)\,\mathrm{d}B^{\lambda}_r\,\mathrm{d}s,\right.\right.\\
&\qquad\left.\left.\int_{\lfloor t\rfloor}^{t}\left\langle \nabla H(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor}), \int_{\lfloor s\rfloor}^{s} -\lambda h(\overline{\Theta}^{\lambda}_{\lfloor r\rfloor})\,\mathrm{d}r\right\rangle\,\mathrm{d}B^{\lambda}_s\right\rangle\right]\\
&\quad + 2\lambda\beta^{-1}\mathbb{E}\left[\left\langle \int_{\lfloor t\rfloor}^{t}\int_{\lfloor s\rfloor}^{s}\left(H(\widetilde{\zeta}^{\lambda,n}_{\lfloor r\rfloor}) - H(\overline{\Theta}^{\lambda}_{\lfloor r\rfloor})\right)\,\mathrm{d}B^{\lambda}_r\,\mathrm{d}s,\right.\right.\\
&\qquad\left.\left.\int_{\lfloor t\rfloor}^{t}\left\langle \nabla H(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor}), \int_{\lfloor s\rfloor}^{s}\int_{\lfloor r\rfloor}^{r} \lambda^2 H(\overline{\Theta}^{\lambda}_{\lfloor \nu\rfloor})h(\overline{\Theta}^{\lambda}_{\lfloor \nu\rfloor})\,\mathrm{d}\nu\,\mathrm{d}r\right\rangle\,\mathrm{d}B^{\lambda}_s\right\rangle\right]\\
&\quad + 2\lambda\beta^{-1}\mathbb{E}\left[\left\langle \int_{\lfloor t\rfloor}^{t}\int_{\lfloor s\rfloor}^{s}\left(H(\widetilde{\zeta}^{\lambda,n}_{\lfloor r\rfloor}) - H(\overline{\Theta}^{\lambda}_{\lfloor r\rfloor})\right)\,\mathrm{d}B^{\lambda}_r\,\mathrm{d}s,\right.\right.\\
&\qquad\left.\left.\int_{\lfloor t\rfloor}^{t}\left\langle \nabla H(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor}), \int_{\lfloor s\rfloor}^{s}\int_{\lfloor r\rfloor}^{r} -\lambda^2\beta^{-1}\Upsilon(\overline{\Theta}^{\lambda}_{\lfloor \nu\rfloor})\,\mathrm{d}\nu\,\mathrm{d}r\right\rangle\,\mathrm{d}B^{\lambda}_s\right\rangle\right]\\
&\quad + 2\lambda\beta^{-1}\mathbb{E}\left[\left\langle \int_{\lfloor t\rfloor}^{t}\int_{\lfloor s\rfloor}^{s}\left(H(\widetilde{\zeta}^{\lambda,n}_{\lfloor r\rfloor}) - H(\overline{\Theta}^{\lambda}_{\lfloor r\rfloor})\right)\,\mathrm{d}B^{\lambda}_r\,\mathrm{d}s,\right.\right.\\
&\qquad\left.\left.\int_{\lfloor t\rfloor}^{t}\left\langle \nabla H(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor}), \int_{\lfloor s\rfloor}^{s} -\lambda\sqrt{2\lambda\beta^{-1}}\int_{\lfloor r\rfloor}^{r} H(\overline{\Theta}^{\lambda}_{\lfloor \nu\rfloor})\,\mathrm{d}B^{\lambda}_\nu\,\mathrm{d}r\right\rangle\,\mathrm{d}B^{\lambda}_s\right\rangle\right]\\
&\quad + 2\lambda\beta^{-1}\mathbb{E}\left[\left\langle \int_{\lfloor t\rfloor}^{t}\int_{\lfloor s\rfloor}^{s}\left(H(\widetilde{\zeta}^{\lambda,n}_{\lfloor r\rfloor}) - H(\overline{\Theta}^{\lambda}_{\lfloor r\rfloor})\right)\,\mathrm{d}B^{\lambda}_r\,\mathrm{d}s,\right.\right.\\
&\qquad\left.\left.\int_{\lfloor t\rfloor}^{t}\left\langle \nabla H(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor}), \sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^{s}\mathrm{d}B^{\lambda}_r\right\rangle\,\mathrm{d}B^{\lambda}_s\right\rangle\right].
\end{aligned}
\tag{158}
$$

Recall that $(\mathcal{F}^{\lambda}_t)_{t\geq 0}$ is the completed natural filtration of $(B^{\lambda}_t)_{t\geq 0}$. Then, we note that, for any $i,j,k = 1,\ldots,d$, it holds that

$$
\begin{aligned}
&\mathbb{E}\left[\left(\int_{\lfloor t\rfloor}^{t}\int_{\lfloor s\rfloor}^{s}\left(H^{(i,j)}(\widetilde{\zeta}^{\lambda,n}_{\lfloor r\rfloor}) - H^{(i,j)}(\overline{\Theta}^{\lambda}_{\lfloor r\rfloor})\right)\,\mathrm{d}(B^{\lambda}_r)^{(j)}\,\mathrm{d}s\right)\right.\\
&\qquad\left.\times\left(\int_{\lfloor t\rfloor}^{t}\left(\partial_{\theta^{(k)}}H^{(i,j)}(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor})\int_{\lfloor s\rfloor}^{s}\mathrm{d}(B^{\lambda}_r)^{(k)}\right)\,\mathrm{d}(B^{\lambda}_s)^{(j)}\right)\right]\\
&= \mathbb{E}\Bigg[\left(H^{(i,j)}(\widetilde{\zeta}^{\lambda,n}_{\lfloor t\rfloor}) - H^{(i,j)}(\overline{\Theta}^{\lambda}_{\lfloor t\rfloor})\right)\partial_{\theta^{(k)}}H^{(i,j)}(\overline{\Theta}^{\lambda}_{\lfloor t\rfloor})\\
&\qquad\times\mathbb{E}\left[\int_{\lfloor t\rfloor}^{t}\int_{\lfloor s\rfloor}^{s}\mathrm{d}(B^{\lambda}_r)^{(j)}\,\mathrm{d}s\int_{\lfloor t\rfloor}^{t}\int_{\lfloor s\rfloor}^{s}\mathrm{d}(B^{\lambda}_r)^{(k)}\,\mathrm{d}(B^{\lambda}_s)^{(j)}\,\middle|\,\mathcal{F}^{\lambda}_{\lfloor t\rfloor}\right]\Bigg]\\
&= \mathbb{E}\Bigg[\left(H^{(i,j)}(\widetilde{\zeta}^{\lambda,n}_{\lfloor t\rfloor}) - H^{(i,j)}(\overline{\Theta}^{\lambda}_{\lfloor t\rfloor})\right)\partial_{\theta^{(k)}}H^{(i,j)}(\overline{\Theta}^{\lambda}_{\lfloor t\rfloor})\\
&\qquad\times\mathbb{E}\left[\int_{\lfloor t\rfloor}^{t}(t-s)\,\mathrm{d}(B^{\lambda}_s)^{(j)}\int_{\lfloor t\rfloor}^{t}\int_{\lfloor s\rfloor}^{s}\mathrm{d}(B^{\lambda}_r)^{(k)}\,\mathrm{d}(B^{\lambda}_s)^{(j)}\,\middle|\,\mathcal{F}^{\lambda}_{\lfloor t\rfloor}\right]\Bigg]\\
&= \mathbb{E}\Bigg[\left(H^{(i,j)}(\widetilde{\zeta}^{\lambda,n}_{\lfloor t\rfloor}) - H^{(i,j)}(\overline{\Theta}^{\lambda}_{\lfloor t\rfloor})\right)\partial_{\theta^{(k)}}H^{(i,j)}(\overline{\Theta}^{\lambda}_{\lfloor t\rfloor})\\
&\qquad\times\left(t\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[\int_{\lfloor t\rfloor}^{s}\mathrm{d}(B^{\lambda}_r)^{(k)}\,\middle|\,\mathcal{F}^{\lambda}_{\lfloor t\rfloor}\right]\mathrm{d}s - \int_{\lfloor t\rfloor}^{t}s\mathbb{E}\left[\int_{\lfloor t\rfloor}^{s}\mathrm{d}(B^{\lambda}_r)^{(k)}\,\middle|\,\mathcal{F}^{\lambda}_{\lfloor t\rfloor}\right]\mathrm{d}s\right)\Bigg]\\
&= 0.
\end{aligned}
$$

This implies that the last term in (158) is zero. Indeed, we have that

$$
\begin{aligned}
&2\lambda\beta^{-1}\mathbb{E}\left[\left\langle \int_{\lfloor t\rfloor}^{t}\int_{\lfloor s\rfloor}^{s}\left(H(\widetilde{\zeta}^{\lambda,n}_{\lfloor r\rfloor})-H(\overline{\Theta}^{\lambda}_{\lfloor r\rfloor})\right)\,\mathrm{d}B^{\lambda}_r\,\mathrm{d}s,\right.\right.\\
&\qquad\left.\left.\int_{\lfloor t\rfloor}^{t}\left\langle \nabla H(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor}),\sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^{s}\,\mathrm{d}B^{\lambda}_r\right\rangle\,\mathrm{d}B^{\lambda}_s\right\rangle\right]\\
&=(2\lambda\beta^{-1})^{3/2}\mathbb{E}\left[\sum_{i=1}^{d}\left(\int_{\lfloor t\rfloor}^{t}\sum_{j=1}^{d}\int_{\lfloor s\rfloor}^{s}\left(H^{(i,j)}(\widetilde{\zeta}^{\lambda,n}_{\lfloor r\rfloor})-H^{(i,j)}(\overline{\Theta}^{\lambda}_{\lfloor r\rfloor})\right)\,\mathrm{d}(B^{\lambda}_r)^{(j)}\,\mathrm{d}s\right)\right.\\
&\qquad\left.\times\left(\sum_{j=1}^{d}\int_{\lfloor t\rfloor}^{t}\left(\sum_{k=1}^{d}\partial_{\theta^{(k)}}H^{(i,j)}(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor})\int_{\lfloor s\rfloor}^{s}\,\mathrm{d}(B^{\lambda}_r)^{(k)}\right)\,\mathrm{d}(B^{\lambda}_s)^{(j)}\right)\right]=0.
\end{aligned}
$$

Then, by using Remark 3.2 and (158) with the result above, we obtain that

$$
\begin{aligned}
&\mathbb{E}\left[2\lambda\beta^{-1}\left\langle \int_{\lfloor t\rfloor}^{t}\int_{\lfloor s\rfloor}^{s}\left(H(\widetilde{\zeta}^{\lambda,n}_{\lfloor r\rfloor})-H(\overline{\Theta}^{\lambda}_{\lfloor r\rfloor})\right)\,\mathrm{d}B^{\lambda}_r\,\mathrm{d}s,\int_{\lfloor t\rfloor}^{t}\mathfrak{M}(\overline{\Theta}^{\lambda}_s,\overline{\Theta}^{\lambda}_{\lfloor s\rfloor})\,\mathrm{d}B^{\lambda}_s\right\rangle\right]\\
&\le 4\lambda^2\beta^{-2}\mathbb{E}\left[\left|\int_{\lfloor t\rfloor}^{t}\int_{\lfloor s\rfloor}^{s}\left(H(\widetilde{\zeta}^{\lambda,n}_{\lfloor r\rfloor})-H(\overline{\Theta}^{\lambda}_{\lfloor r\rfloor})\right)\,\mathrm{d}B^{\lambda}_r\,\mathrm{d}s\right|^2\right]\\
&\quad+\lambda^2\mathbb{E}\left[\left|\int_{\lfloor t\rfloor}^{t}\left\langle \nabla H(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor}),\int_{\lfloor s\rfloor}^{s}-h(\overline{\Theta}^{\lambda}_{\lfloor r\rfloor})\,\mathrm{d}r\right\rangle\,\mathrm{d}B^{\lambda}_s\right|^2\right]\\
&\quad+\lambda^4\mathbb{E}\left[\left|\int_{\lfloor t\rfloor}^{t}\left\langle \nabla H(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor}),\int_{\lfloor s\rfloor}^{s}\int_{\lfloor r\rfloor}^{r}H(\overline{\Theta}^{\lambda}_{\lfloor \nu\rfloor})h(\overline{\Theta}^{\lambda}_{\lfloor \nu\rfloor})\,\mathrm{d}\nu\,\mathrm{d}r\right\rangle\,\mathrm{d}B^{\lambda}_s\right|^2\right]\\
&\quad+\lambda^4\beta^{-2}\mathbb{E}\left[\left|\int_{\lfloor t\rfloor}^{t}\left\langle \nabla H(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor}),\int_{\lfloor s\rfloor}^{s}\int_{\lfloor r\rfloor}^{r}-\Upsilon(\overline{\Theta}^{\lambda}_{\lfloor \nu\rfloor})\,\mathrm{d}\nu\,\mathrm{d}r\right\rangle\,\mathrm{d}B^{\lambda}_s\right|^2\right]\\
&\quad+2\lambda^3\beta^{-1}\mathbb{E}\left[\left|\int_{\lfloor t\rfloor}^{t}\left\langle \nabla H(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor}),\int_{\lfloor s\rfloor}^{s}-\int_{\lfloor r\rfloor}^{r}H(\overline{\Theta}^{\lambda}_{\lfloor \nu\rfloor})\,\mathrm{d}B^{\lambda}_\nu\,\mathrm{d}r\right\rangle\,\mathrm{d}B^{\lambda}_s\right|^2\right]\\
&\le 4\lambda^2\beta^{-2}\int_{\lfloor t\rfloor}^{t}\int_{\lfloor s\rfloor}^{s}\mathbb{E}\left[|H(\widetilde{\zeta}^{\lambda,n}_{\lfloor r\rfloor})-H(\overline{\Theta}^{\lambda}_{\lfloor r\rfloor})|_{\mathsf{F}}^2\right]\,\mathrm{d}r\,\mathrm{d}s\\
&\quad+\lambda^2\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[\sum_{i,j=1}^{d}\left|\left\langle \nabla H^{(i,j)}(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor}),h(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor})\right\rangle\right|^2\right]\,\mathrm{d}s\\
&\quad+\lambda^4\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[\sum_{i,j=1}^{d}\left|\left\langle \nabla H^{(i,j)}(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor}),H(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor})h(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor})\right\rangle\right|^2\right]\,\mathrm{d}s\\
&\quad+\lambda^4\beta^{-2}\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[\sum_{i,j=1}^{d}\left|\left\langle \nabla H^{(i,j)}(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor}),\Upsilon(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor})\right\rangle\right|^2\right]\,\mathrm{d}s\\
&\quad+2\lambda^3\beta^{-1}\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[\sum_{i,j=1}^{d}\left|\left\langle \nabla H^{(i,j)}(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor}),H(\overline{\Theta}^{\lambda}_{\lfloor s\rfloor})\int_{\lfloor s\rfloor}^{s}\int_{\lfloor r\rfloor}^{r}\,\mathrm{d}B^{\lambda}_\nu\,\mathrm{d}r\right\rangle\right|^2\right]\,\mathrm{d}s\\
&\le 8\lambda^2\beta^{-2}d\overline{L}_2^2\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[|\widetilde{\zeta}^{\lambda,n}_{\lfloor s\rfloor}|^2+|\overline{\Theta}^{\lambda}_{\lfloor s\rfloor}|^2\right]\,\mathrm{d}s\\
&\quad+\lambda^2\int_{\lfloor t\rfloor}^{t}\mathbb{E}\left[d^2\overline{\mathsf{K}}_0^2(1+|\overline{\Theta}^{\lambda}_{\lfloor s\rfloor}|^q)^2\overline{\mathsf{K}}_1^2(1+|\overline{\Theta}^{\lambda}_{\lfloor s\rfloor}|)^2\right]\,\mathrm{d}s
\end{aligned}
$$

$$
\begin{aligned}
&+\lambda^4\int_{\lfloor t\rfloor}^t \mathbb{E}\left[d^2\overline{\mathsf{K}}_0^2(1+|\overline{\Theta}^\lambda_{\lfloor s\rfloor}|^q)^2\overline{L}_3^2\overline{\mathsf{K}}_1^2(1+|\overline{\Theta}^\lambda_{\lfloor s\rfloor}|)^2\right]\,\mathrm{d}s\\
&+\lambda^4\beta^{-2}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[d^2\overline{\mathsf{K}}_0^2(1+|\overline{\Theta}^\lambda_{\lfloor s\rfloor}|^q)^2d^2\overline{L}_2^2\right]\,\mathrm{d}s\\
&+2\lambda^3\beta^{-1}\int_{\lfloor t\rfloor}^t \mathbb{E}\left[d^2\overline{\mathsf{K}}_0^2(1+|\overline{\Theta}^\lambda_{\lfloor s\rfloor}|^q)^2\overline{L}_3^2\left|\int_{\lfloor s\rfloor}^s\int_{\lfloor r\rfloor}^r \mathrm{d}B^\lambda_\nu\,\mathrm{d}r\right|^2\right]\,\mathrm{d}s\\
&\leq 8\lambda^2\beta^{-2}d\overline{L}_2^2\int_{\lfloor t\rfloor}^t \mathbb{E}\left[V_4(\widetilde{\zeta}^{\lambda,n}_{\lfloor s\rfloor})+1+|\overline{\Theta}^\lambda_{\lfloor s\rfloor}|^4\right]\,\mathrm{d}s\\
&+4\lambda^2(\overline{\mathsf{K}}_1^2+\overline{L}_3^2\overline{\mathsf{K}}_1^2+d^2\beta^{-2}\overline{L}_2^2)d^2\overline{\mathsf{K}}_0^2\int_{\lfloor t\rfloor}^t \mathbb{E}\left[1+|\overline{\Theta}^\lambda_{\lfloor s\rfloor}|^4\right]\,\mathrm{d}s\\
&+8\lambda^2\beta^{-1}d^2\overline{\mathsf{K}}_0^2\overline{L}_3^2\int_{\lfloor t\rfloor}^t\left(\mathbb{E}\left[1+|\overline{\Theta}^\lambda_{\lfloor s\rfloor}|^4\right]\right)^{1/2}(3(d+4))\,\mathrm{d}s\\
&\leq 24\lambda^2\beta^{-2}d\overline{L}_2^2\left(e^{-\lambda\bar{a}\lfloor t\rfloor/2}\mathbb{E}[|\theta_0|^4]+(\bar{\mathsf{c}}_2\,(1+1/\bar{a})+1)+\mathrm{v}_4(\overline{\mathsf{M}}_V(4)\right)\\
&+2^7\lambda^2(\overline{\mathsf{K}}_1^2+\overline{L}_3^2\overline{\mathsf{K}}_1^2+\beta^{-2}\overline{L}_2^2+\beta^{-1}\overline{L}_3^2)d^4\overline{\mathsf{K}}_0^2\\
&\quad\times\left(e^{-\lambda\bar{a}\lfloor t\rfloor/2}\mathbb{E}[|\theta_0|^4]+\bar{\mathsf{c}}_2\,(1+1/\bar{a})+1\right)\\
&\leq\lambda^2\left(e^{-\bar{a}n/4}5\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4]+5\widetilde{\mathsf{C}}_{\mathbf{S}2}\right),
\end{aligned}
$$

where the fifth inequality holds due to Lemma D.3 and D.5, and where $\overline{\mathsf{C}}_{\mathbf{S}2}$, $\widetilde{\mathsf{C}}_{\mathbf{S}2}$ are given in (151).

This completes the proof. □

Proof of Lemma D.6 . By using the definitions of $\overline{\Theta}^\lambda_t$ in (109) and $\widetilde{\zeta}^{\lambda,n}_t$ in Definition D.2, and by applying Itô's formula, we obtain, for any $n\in\mathbb{N}_0$, $t\in(nT,(n+1)T]$,

$$
\begin{aligned}
&W_2^2(\mathcal{L}(\overline{\Theta}^\lambda_t),\mathcal{L}(\widetilde{\zeta}^{\lambda,n}_t))\\
&\leq\mathbb{E}\left[\left|\overline{\Theta}^\lambda_t-\widetilde{\zeta}^{\lambda,n}_t\right|^2\right]\\
&=-2\lambda\mathbb{E}\left[\int_{nT}^t\left\langle\overline{\Theta}^\lambda_s-\widetilde{\zeta}^{\lambda,n}_s,h(\overline{\Theta}^\lambda_{\lfloor s\rfloor})-h(\widetilde{\zeta}^{\lambda,n}_s)\right\rangle\,\mathrm{d}s\right]\\
&+2\lambda\mathbb{E}\left[\int_{nT}^t\left\langle\overline{\Theta}^\lambda_s-\widetilde{\zeta}^{\lambda,n}_s,\lambda\int_{\lfloor s\rfloor}^s\left(H(\overline{\Theta}^\lambda_{\lfloor r\rfloor})h(\overline{\Theta}^\lambda_{\lfloor r\rfloor})-\beta^{-1}\Upsilon(\overline{\Theta}^\lambda_{\lfloor r\rfloor})\right)\,\mathrm{d}r\right\rangle\,\mathrm{d}s\right]\\
&-2\lambda\mathbb{E}\left[\int_{nT}^t\left\langle\overline{\Theta}^\lambda_s-\widetilde{\zeta}^{\lambda,n}_s,\sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^s H(\overline{\Theta}^\lambda_{\lfloor r\rfloor})\,\mathrm{d}B^\lambda_r\right\rangle\,\mathrm{d}s\right]\\
&=-2\lambda\mathbb{E}\left[\int_{nT}^t\left\langle\overline{\Theta}^\lambda_s-\widetilde{\zeta}^{\lambda,n}_s,h(\overline{\Theta}^\lambda_s)-h(\widetilde{\zeta}^{\lambda,n}_s)\right\rangle\,\mathrm{d}s\right]\\
&-2\lambda\mathbb{E}\left[\int_{nT}^t\left\langle\overline{\Theta}^\lambda_s-\widetilde{\zeta}^{\lambda,n}_s,h(\overline{\Theta}^\lambda_{\lfloor s\rfloor})-h(\overline{\Theta}^\lambda_s)\right\rangle\,\mathrm{d}s\right]\\
&+2\lambda\mathbb{E}\left[\int_{nT}^t\left\langle\overline{\Theta}^\lambda_s-\widetilde{\zeta}^{\lambda,n}_s,\lambda\int_{\lfloor s\rfloor}^s\left(H(\overline{\Theta}^\lambda_{\lfloor r\rfloor})h(\overline{\Theta}^\lambda_{\lfloor r\rfloor})-\beta^{-1}\Upsilon(\overline{\Theta}^\lambda_{\lfloor r\rfloor})\right)\,\mathrm{d}r\right\rangle\,\mathrm{d}s\right]\\
&-2\lambda\mathbb{E}\left[\int_{nT}^t\left\langle\overline{\Theta}^\lambda_s-\widetilde{\zeta}^{\lambda,n}_s,\sqrt{2\lambda\beta^{-1}}\int_{\lfloor s\rfloor}^s H(\overline{\Theta}^\lambda_{\lfloor r\rfloor})\,\mathrm{d}B^\lambda_r\right\rangle\,\mathrm{d}s\right].
\end{aligned}
\tag{159}
$$

By applying Itô's formula to $h(\overline{\Theta}_s^\lambda)$, we obtain (152). Substituting (152) into (159), using Assumption 5 and Young's inequality yield

$$
\begin{aligned}
&\mathbb{E}\left[\left|\overline{\Theta}_t^\lambda - \widetilde{\zeta}_t^{\lambda,n}\right|^2\right] \\
&\leq 2\lambda \overline{L}_3 \int_{nT}^t \mathbb{E}\left[|\overline{\Theta}_s^\lambda - \widetilde{\zeta}_s^{\lambda,n}|^2\right] \mathrm{d}s \\
&\quad + 2\lambda \mathbb{E}\left[\int_{nT}^t \left\langle \overline{\Theta}_s^\lambda - \widetilde{\zeta}_s^{\lambda,n}, -\lambda \int_{\lfloor s \rfloor}^s \left(H(\overline{\Theta}_r^\lambda) - H(\overline{\Theta}_{\lfloor r \rfloor}^\lambda)\right) h(\overline{\Theta}_{\lfloor r \rfloor}^\lambda)\,\mathrm{d}r \right\rangle \mathrm{d}s\right] \\
&\quad + 2\lambda \mathbb{E}\left[\int_{nT}^t \left\langle \overline{\Theta}_s^\lambda - \widetilde{\zeta}_s^{\lambda,n}, \lambda^2 \int_{\lfloor s \rfloor}^s H(\overline{\Theta}_r^\lambda) \int_{\lfloor r \rfloor}^r H(\overline{\Theta}_{\lfloor \nu \rfloor}^\lambda) h(\overline{\Theta}_{\lfloor \nu \rfloor}^\lambda)\,\mathrm{d}\nu\,\mathrm{d}r \right\rangle \mathrm{d}s\right] \\
&\quad + 2\lambda \mathbb{E}\left[\int_{nT}^t \left\langle \overline{\Theta}_s^\lambda - \widetilde{\zeta}_s^{\lambda,n}, -\lambda^2\beta^{-1} \int_{\lfloor s \rfloor}^s H(\overline{\Theta}_r^\lambda) \int_{\lfloor r \rfloor}^r \Upsilon(\overline{\Theta}_{\lfloor \nu \rfloor}^\lambda)\,\mathrm{d}\nu\,\mathrm{d}r \right\rangle \mathrm{d}s\right] \\
&\quad + 2\lambda \mathbb{E}\left[\int_{nT}^t \left\langle \overline{\Theta}_s^\lambda - \widetilde{\zeta}_s^{\lambda,n}, -\lambda\sqrt{2\lambda\beta^{-1}} \int_{\lfloor s \rfloor}^s H(\overline{\Theta}_r^\lambda) \int_{\lfloor r \rfloor}^r H(\overline{\Theta}_{\lfloor \nu \rfloor}^\lambda)\,\mathrm{d}B_\nu^\lambda\,\mathrm{d}r \right\rangle \mathrm{d}s\right] \\
&\quad + 2\lambda \mathbb{E}\left[\int_{nT}^t \left\langle \overline{\Theta}_s^\lambda - \widetilde{\zeta}_s^{\lambda,n}, \sqrt{2\lambda\beta^{-1}} \int_{\lfloor s \rfloor}^s \left(H(\overline{\Theta}_r^\lambda) - H(\overline{\Theta}_{\lfloor r \rfloor}^\lambda)\right) \mathrm{d}B_r^\lambda \right\rangle \mathrm{d}s\right] \\
&\quad + 2\lambda \mathbb{E}\left[\int_{nT}^t \left\langle \overline{\Theta}_s^\lambda - \widetilde{\zeta}_s^{\lambda,n}, \lambda\beta^{-1} \int_{\lfloor s \rfloor}^s \left(\Upsilon(\overline{\Theta}_r^\lambda) - \Upsilon(\overline{\Theta}_{\lfloor r \rfloor}^\lambda)\right) \mathrm{d}r \right\rangle \mathrm{d}s\right] \\
&\leq \lambda(2\overline{L}_3 + 5) \int_{nT}^t \mathbb{E}\left[|\overline{\Theta}_s^\lambda - \widetilde{\zeta}_s^{\lambda,n}|^2\right] \mathrm{d}s \\
&\quad + \lambda \int_{nT}^t \mathbb{E}\left[\left|-\lambda \int_{\lfloor s \rfloor}^s \left(H(\overline{\Theta}_r^\lambda) - H(\overline{\Theta}_{\lfloor r \rfloor}^\lambda)\right) h(\overline{\Theta}_{\lfloor r \rfloor}^\lambda)\,\mathrm{d}r\right|^2\right] \mathrm{d}s \\
&\quad + \lambda \int_{nT}^t \mathbb{E}\left[\left|\lambda^2 \int_{\lfloor s \rfloor}^s H(\overline{\Theta}_r^\lambda) \int_{\lfloor r \rfloor}^r H(\overline{\Theta}_{\lfloor \nu \rfloor}^\lambda) h(\overline{\Theta}_{\lfloor \nu \rfloor}^\lambda)\,\mathrm{d}\nu\,\mathrm{d}r\right|^2\right] \mathrm{d}s \\
&\quad + \lambda \int_{nT}^t \mathbb{E}\left[\left|-\lambda^2\beta^{-1} \int_{\lfloor s \rfloor}^s H(\overline{\Theta}_r^\lambda) \int_{\lfloor r \rfloor}^r \Upsilon(\overline{\Theta}_{\lfloor \nu \rfloor}^\lambda)\,\mathrm{d}\nu\,\mathrm{d}r\right|^2\right] \mathrm{d}s \qquad (160) \\
&\quad + \lambda \int_{nT}^t \mathbb{E}\left[\left|-\lambda\sqrt{2\lambda\beta^{-1}} \int_{\lfloor s \rfloor}^s H(\overline{\Theta}_r^\lambda) \int_{\lfloor r \rfloor}^r H(\overline{\Theta}_{\lfloor \nu \rfloor}^\lambda)\,\mathrm{d}B_\nu^\lambda\,\mathrm{d}r\right|^2\right] \mathrm{d}s \\
&\quad + \lambda \int_{nT}^t \mathbb{E}\left[\left|\lambda\beta^{-1} \int_{\lfloor s \rfloor}^s \left(\Upsilon(\overline{\Theta}_r^\lambda) - \Upsilon(\overline{\Theta}_{\lfloor r \rfloor}^\lambda)\right) \mathrm{d}r\right|^2\right] \mathrm{d}s \\
&\quad + 2\lambda \mathbb{E}\left[\int_{nT}^t \left\langle \overline{\Theta}_s^\lambda - \widetilde{\zeta}_s^{\lambda,n}, \sqrt{2\lambda\beta^{-1}} \int_{\lfloor s \rfloor}^s \left(H(\overline{\Theta}_r^\lambda) - H(\overline{\Theta}_{\lfloor r \rfloor}^\lambda)\right) \mathrm{d}B_r^\lambda \right\rangle \mathrm{d}s\right].
\end{aligned}
$$

By using Lemma D.11 and (160) becomes

$$
\begin{aligned}
\mathbb{E}\left[\left|\overline{\Theta}_t^\lambda - \widetilde{\zeta}_t^{\lambda,n}\right|^2\right] &\leq \lambda(2\overline{L}_3 + 5) \int_{nT}^t \mathbb{E}\left[|\overline{\Theta}_s^\lambda - \widetilde{\zeta}_s^{\lambda,n}|^2\right] \mathrm{d}s \\
&\quad + 5\lambda^{2+q}\left(e^{-\overline{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2}\right) \qquad (161) \\
&\quad + \overline{\mathfrak{I}}_1 + \overline{\mathfrak{I}}_2,
\end{aligned}
$$

where

$$\overline{\mathfrak{J}}_1^\lambda(t) := 2\lambda\mathbb{E}\left[\int_{nT}^t \left\langle \overline{\Theta}_s^\lambda - \widetilde{\zeta}_s^{\lambda,n}, \sqrt{2\lambda\beta^{-1}} \int_{\lfloor s\rfloor}^s \left(H(\overline{\Theta}_r^\lambda) - H_\lambda(\overline{\Theta}_{\lfloor r\rfloor}^\lambda) - \mathfrak{M}(\overline{\Theta}_r^\lambda, \overline{\Theta}_{\lfloor r\rfloor}^\lambda)\right) \mathrm{d}B_r^\lambda \right\rangle \mathrm{d}s\right],$$

$$\overline{\mathfrak{J}}_2^\lambda(t) := 2\lambda\mathbb{E}\left[\int_{nT}^t \left\langle \overline{\Theta}_s^\lambda - \widetilde{\zeta}_s^{\lambda,n}, \sqrt{2\lambda\beta^{-1}} \int_{\lfloor s\rfloor}^s \mathfrak{M}(\overline{\Theta}_r^\lambda, \overline{\Theta}_{\lfloor r\rfloor}^\lambda) \,\mathrm{d}B_r^\lambda \right\rangle \mathrm{d}s\right]$$

with $\mathfrak{M}$ defined in Definition D.13. By using Young's inequality and Lemma D.14, we have that

$$\overline{\mathfrak{J}}_1^\lambda(t) \le \lambda \int_{nT}^t \mathbb{E}\left[|\overline{\Theta}_s^\lambda - \widetilde{\zeta}_s^{\lambda,n}|^2\right] \mathrm{d}s + \lambda^{2+q}\left(e^{-\bar{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2}\right). \tag{162}$$

To establish an upper bound for $\overline{\mathfrak{J}}_2^\lambda(t)$, we recall the definitions of $(\overline{\Theta}_t^\lambda)_{t\ge 0}$ and $(\widetilde{\zeta}_t^{\lambda,n})_{t\ge 0}$ given in (109) and Definition D.2, respectively, and consider the following splitting:

$$\begin{aligned}
\overline{\mathfrak{J}}_2^\lambda(t) &= 2\lambda \int_{nT}^t \mathbb{E}\left[\left\langle \overline{\Theta}_s^\lambda - \overline{\Theta}_{\lfloor s\rfloor}^\lambda - (\widetilde{\zeta}_s^{\lambda,n} - \widetilde{\zeta}_{\lfloor s\rfloor}^{\lambda,n}), \sqrt{2\lambda\beta^{-1}} \int_{\lfloor s\rfloor}^s \mathfrak{M}(\overline{\Theta}_r^\lambda, \overline{\Theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\right\rangle\right] \mathrm{d}s \\
&= 2\lambda \int_{nT}^t \mathbb{E}\left[\left\langle \int_{\lfloor s\rfloor}^s \left(\lambda(h(\widetilde{\zeta}_r^{\lambda,n}) - h(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)) + \lambda^2 \int_{\lfloor r\rfloor}^r \left(H(\overline{\Theta}_{\lfloor\nu\rfloor}^\lambda)h(\overline{\Theta}_{\lfloor\nu\rfloor}^\lambda) - \beta^{-1}\Upsilon(\overline{\Theta}_{\lfloor\nu\rfloor}^\lambda)\right) \mathrm{d}\nu\right.\right.\right. \\
&\qquad \left.\left.\left. -\lambda\sqrt{2\lambda\beta^{-1}} \int_{\lfloor r\rfloor}^r H(\overline{\Theta}_{\lfloor\nu\rfloor}^\lambda)\,\mathrm{d}B_\nu^\lambda\right) \mathrm{d}r, \sqrt{2\lambda\beta^{-1}} \int_{\lfloor s\rfloor}^s \mathfrak{M}(\overline{\Theta}_r^\lambda, \overline{\Theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\right\rangle\right] \mathrm{d}s \\
&= 2\lambda \int_{nT}^t \mathbb{E}\left[\left\langle \int_{\lfloor s\rfloor}^s \left(\lambda(h(\overline{\Theta}_r^\lambda) - h(\overline{\Theta}_{\lfloor r\rfloor}^\lambda)) + \lambda^2 \int_{\lfloor r\rfloor}^r \left(H(\overline{\Theta}_{\lfloor\nu\rfloor}^\lambda)h(\overline{\Theta}_{\lfloor\nu\rfloor}^\lambda) - \beta^{-1}\Upsilon(\overline{\Theta}_{\lfloor\nu\rfloor}^\lambda)\right) \mathrm{d}\nu\right.\right.\right. \\
&\qquad \left.\left.\left. -\lambda\sqrt{2\lambda\beta^{-1}} \int_{\lfloor r\rfloor}^r H(\overline{\Theta}_{\lfloor\nu\rfloor}^\lambda)\,\mathrm{d}B_\nu^\lambda\right) \mathrm{d}r, \sqrt{2\lambda\beta^{-1}} \int_{\lfloor s\rfloor}^s \mathfrak{M}(\overline{\Theta}_r^\lambda, \overline{\Theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\right\rangle\right] \mathrm{d}s + \overline{\mathfrak{J}}_{2,1}^\lambda(t),
\end{aligned}$$

where the first equality holds due to the following:

$$\begin{aligned}
&\mathbb{E}\left[\left\langle \overline{\Theta}_{\lfloor s\rfloor}^\lambda + \widetilde{\zeta}_{\lfloor s\rfloor}^{\lambda,n}, \sqrt{2\lambda\beta^{-1}} \int_{\lfloor s\rfloor}^s \mathfrak{M}(\overline{\Theta}_r^\lambda, \overline{\Theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\right\rangle\right] \\
&= \mathbb{E}\left[\left\langle \overline{\Theta}_{\lfloor s\rfloor}^\lambda + \widetilde{\zeta}_{\lfloor s\rfloor}^{\lambda,n}, \mathbb{E}\left[\sqrt{2\lambda\beta^{-1}} \int_{\lfloor s\rfloor}^s \mathfrak{M}(\overline{\Theta}_r^\lambda, \overline{\Theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda \,\middle|\, \mathcal{F}_{\lfloor s\rfloor}^\lambda\right]\right\rangle\right] = 0,
\end{aligned} \tag{163}$$

and where

$$\overline{\mathfrak{J}}_{2,1}^\lambda(t) = 2\lambda \int_{nT}^t \mathbb{E}\left[\left\langle \int_{\lfloor s\rfloor}^s \lambda(h(\widetilde{\zeta}_r^{\lambda,n}) - h(\overline{\Theta}_r^\lambda))\,\mathrm{d}r, \sqrt{2\lambda\beta^{-1}} \int_{\lfloor s\rfloor}^s \mathfrak{M}(\overline{\Theta}_r^\lambda, \overline{\Theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\right\rangle\right] \mathrm{d}s.$$

By applying Cauchy-Schwarz inequality, Corollary D.12 and Lemma D.14, we further obtain that

$$\begin{aligned}
\overline{\mathfrak{J}}_2^\lambda(t) &\le 2\lambda \int_{nT}^t \left(\mathbb{E}\left[\left|\int_{\lfloor s\rfloor}^s \lambda\left(h(\overline{\Theta}_r^\lambda) - h(\overline{\Theta}_{\lfloor r\rfloor}^\lambda) + \lambda \int_{\lfloor r\rfloor}^r \left(H(\overline{\Theta}_{\lfloor\nu\rfloor}^\lambda)h(\overline{\Theta}_{\lfloor\nu\rfloor}^\lambda) - \beta^{-1}\Upsilon(\overline{\Theta}_{\lfloor\nu\rfloor}^\lambda)\right) \mathrm{d}\nu\right.\right.\right. \\
&\qquad \left.\left.\left.\left. -\sqrt{2\lambda\beta^{-1}} \int_{\lfloor r\rfloor}^r H(\overline{\Theta}_{\lfloor\nu\rfloor}^\lambda)\,\mathrm{d}B_\nu^\lambda\right) \mathrm{d}r\right|^2\right]\right)^{1/2} \\
&\qquad \times \left(\mathbb{E}\left[\left|\sqrt{2\lambda\beta^{-1}} \int_{\lfloor s\rfloor}^s \mathfrak{M}(\overline{\Theta}_r^\lambda, \overline{\Theta}_{\lfloor r\rfloor}^\lambda)\,\mathrm{d}B_r^\lambda\right|^2\right]\right)^{1/2} \mathrm{d}s + \overline{\mathfrak{J}}_{2,1}^\lambda(t) \\
&\le 2\lambda \int_{nT}^t \left(\lambda^4\left(e^{-\bar{a}n/2}36\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4] + 36\widetilde{\mathsf{C}}_{\mathbf{S}2}\right)\right)^{1/2} \\
&\qquad \times \left(\lambda^2\left(e^{-\bar{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2}\right)\right)^{1/2} \mathrm{d}s + \overline{\mathfrak{J}}_{2,1}^\lambda(t) \\
&\le 12\lambda^3\left(e^{-\bar{a}n/2}\overline{\mathsf{C}}_{\mathbf{S}2}\mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2}\right) + \overline{\mathfrak{J}}_{2,1}^\lambda(t).
\end{aligned} \tag{164}$$

At this stage, the task reduces to upper bound $\overline{\mathfrak{I}}_{2,1}^{\lambda}(t)$. To achieve this, we apply Cauchy-Schwarz inequality, Corollary D.12 and Lemma D.14 to obtain

$$\begin{aligned}
\overline{\mathfrak{I}}_{2,1}^{\lambda}(t) &= 2\lambda \int_{nT}^{t} \mathbb{E}\left[\left\langle \int_{\lfloor s\rfloor}^{s} \lambda \left(h(\widetilde{\zeta}_r^{\lambda,n}) - h(\widetilde{\zeta}_{\lfloor r\rfloor}^{\lambda,n}) - \sqrt{2\lambda\beta^{-1}} \int_{\lfloor r\rfloor}^{r} H(\widetilde{\zeta}_{\lfloor \nu\rfloor}^{\lambda,n})\, \mathrm{d}B_\nu^{\lambda} - \left(h(\overline{\Theta}_r^{\lambda}) \right.\right.\right.\right. \\
&\qquad \left.\left.\left.\left. - h(\overline{\Theta}_{\lfloor r\rfloor}^{\lambda}) - \sqrt{2\lambda\beta^{-1}} \int_{\lfloor r\rfloor}^{r} H(\overline{\Theta}_{\lfloor \nu\rfloor}^{\lambda})\, \mathrm{d}B_\nu^{\lambda}\right)\right) \mathrm{d}r, \sqrt{2\lambda\beta^{-1}} \int_{\lfloor s\rfloor}^{s} \mathfrak{M}(\overline{\Theta}_r^{\lambda}, \overline{\Theta}_{\lfloor r\rfloor}^{\lambda})\, \mathrm{d}B_r^{\lambda} \right\rangle\right] \mathrm{d}s \\
&\quad + 2\lambda \int_{nT}^{t} \mathbb{E}\left[\left\langle \int_{\lfloor s\rfloor}^{s} \lambda\sqrt{2\lambda\beta^{-1}} \int_{\lfloor r\rfloor}^{r} (H(\widetilde{\zeta}_{\lfloor \nu\rfloor}^{\lambda,n}) - H(\overline{\Theta}_{\lfloor \nu\rfloor}^{\lambda}))\, \mathrm{d}B_\nu^{\lambda}\, \mathrm{d}r, \right.\right. \\
&\qquad \left.\left. \sqrt{2\lambda\beta^{-1}} \int_{\lfloor s\rfloor}^{s} \mathfrak{M}(\overline{\Theta}_r^{\lambda}, \overline{\Theta}_{\lfloor r\rfloor}^{\lambda})\, \mathrm{d}B_r^{\lambda} \right\rangle\right] \mathrm{d}s \\
&\le 2\lambda \int_{nT}^{t} \left(\mathbb{E}\left[\left| \int_{\lfloor s\rfloor}^{s} \lambda \left(h(\widetilde{\zeta}_r^{\lambda,n}) - h(\widetilde{\zeta}_{\lfloor r\rfloor}^{\lambda,n}) - \sqrt{2\lambda\beta^{-1}} \int_{\lfloor r\rfloor}^{r} H(\widetilde{\zeta}_{\lfloor \nu\rfloor}^{\lambda,n})\, \mathrm{d}B_\nu^{\lambda} - \left(h(\overline{\Theta}_r^{\lambda}) \right.\right.\right.\right.\right. \\
&\qquad \left.\left.\left.\left.\left. - h(\overline{\Theta}_{\lfloor r\rfloor}^{\lambda}) - \sqrt{2\lambda\beta^{-1}} \int_{\lfloor r\rfloor}^{r} H(\overline{\Theta}_{\lfloor \nu\rfloor}^{\lambda})\, \mathrm{d}B_\nu^{\lambda}\right)\right) \mathrm{d}r \right|^2\right]\right)^{1/2} \\
&\qquad \times \left(\mathbb{E}\left[\left| \sqrt{2\lambda\beta^{-1}} \int_{\lfloor s\rfloor}^{s} \mathfrak{M}(\overline{\Theta}_r^{\lambda}, \overline{\Theta}_{\lfloor r\rfloor}^{\lambda})\, \mathrm{d}B_r^{\lambda} \right|^2\right]\right)^{1/2} \mathrm{d}s \\
&\quad + 2\lambda^3 \left(e^{-\bar{a}n/4} 5\overline{\mathsf{C}}_{\mathbf{S}2} \mathbb{E}[|\theta_0|^4] + 5\widetilde{\mathsf{C}}_{\mathbf{S}2} \right) \\
&\le 2\lambda \int_{nT}^{t} \left(\lambda^4 \left(e^{-\bar{a}n/4} 72\overline{\mathsf{C}}_{\mathbf{S}2} \mathbb{E}[|\theta_0|^4] + 72\widetilde{\mathsf{C}}_{\mathbf{S}2} \right)\right)^{1/2} \left(\lambda^2 \left(e^{-\bar{a}n/2} \overline{\mathsf{C}}_{\mathbf{S}2} \mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2} \right)\right)^{1/2} \mathrm{d}s \\
&\quad + 10\lambda^3 \left(e^{-\bar{a}n/4} \overline{\mathsf{C}}_{\mathbf{S}2} \mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2} \right) \\
&\le 34\lambda^3 \left(e^{-\bar{a}n/4} \overline{\mathsf{C}}_{\mathbf{S}2} \mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2} \right). \qquad (165)
\end{aligned}$$

Substituting (165) into (164) yields

$$\overline{\mathfrak{I}}_{2}^{\lambda}(t) \le 46\lambda^3 \left(e^{-\bar{a}n/4} \overline{\mathsf{C}}_{\mathbf{S}2} \mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2} \right). \qquad (166)$$

By applying (162) and (166) to (161), we obtain that

$$\mathbb{E}\left[\left|\overline{\Theta}_t^{\lambda} - \widetilde{\zeta}_t^{\lambda,n}\right|^2\right] \le \lambda(2\overline{L}_3 + 6) \int_{nT}^{t} \mathbb{E}\left[|\overline{\Theta}_s^{\lambda} - \widetilde{\zeta}_s^{\lambda,n}|^2\right] \mathrm{d}s + 52\lambda^{2+q} \left(e^{-\bar{a}n/4} \overline{\mathsf{C}}_{\mathbf{S}2} \mathbb{E}[|\theta_0|^4] + \widetilde{\mathsf{C}}_{\mathbf{S}2} \right),$$

which, by applying Grönwall's lemma, yields

$$W_2^2(\mathcal{L}(\overline{\Theta}_t^{\lambda}), \mathcal{L}(\widetilde{\zeta}_t^{\lambda,n})) \le \mathbb{E}\left[\left|\overline{\Theta}_t^{\lambda} - \widetilde{\zeta}_t^{\lambda,n}\right|^2\right] \le \lambda^{2+q} \left(e^{-\bar{a}n/4} \mathsf{C}_{\mathsf{Lin},0} \mathbb{E}[|\theta_0|^4] + \mathsf{C}_{\mathsf{Lin},1} \right),$$

where

$$\mathsf{C}_{\mathsf{Lin},0} := 52 e^{2\overline{L}_3 + 6} \overline{\mathsf{C}}_{\mathbf{S}2}, \quad \mathsf{C}_{\mathsf{Lin},1} := 52 e^{2\overline{L}_3 + 6} \widetilde{\mathsf{C}}_{\mathbf{S}2} \qquad (167)$$

with $\overline{\mathsf{C}}_{\mathbf{S}2}, \widetilde{\mathsf{C}}_{\mathbf{S}2}$ given in (151). □

REFERENCES

[1] Mathias Barkhagen, Ngoc Huy Chau, Éric Moulines, Miklós Rásonyi, Sotirios Sabanis, and Ying Zhang. On stochastic gradient Langevin dynamics with dependent data streams in the logconcave case. *Bernoulli*, 27(1):1–33, 2021.

[2] Nicolas Brosse, Alain Durmus, and Eric Moulines. The promises and pitfalls of stochastic gradient Langevin dynamics. In *Advances in Neural Information Processing Systems*, pages 8268–8278, 2018.

[3] Nicolas Brosse, Alain Durmus, Éric Moulines, and Sotirios Sabanis. The Tamed Unadjusted Langevin Algorithm. *Stochastic Processes and their Applications*, 129(10):3638–3663, 2019.

[4] Ngoc Huy Chau, Éric Moulines, Miklos Rásonyi, Sotirios Sabanis, and Ying Zhang. On stochastic gradient Langevin dynamics with dependent data streams: The fully nonconvex case. *SIAM Journal on Mathematics of Data Science*, 3(3):959–986, 2021.

[5] Xiang Cheng, Niladri S Chatterji, Yasin Abbasi-Yadkori, Peter L Bartlett, and Michael I Jordan. Sharp convergence rates for Langevin dynamics in the nonconvex setting. *arXiv preprint arXiv:1805.01648*, 2018.

[6] Arnak S Dalalyan. Theoretical guarantees for approximate sampling from smooth and log-concave densities. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(3): 651–676, 2017.

[7] Arnak S Dalalyan and Avetik Karagulyan. User-friendly guarantees for the Langevin Monte Carlo with inaccurate gradient. *Stochastic Processes and their Applications*, 2019.

[8] Alain Durmus and Eric Moulines. High-dimensional Bayesian inference via the unadjusted Langevin algorithm. *Bernoulli*, 25(4A):2854–2882, 2019.

[9] Alain Durmus, Szymon Majewski, and Błażej Miasojedow. Analysis of Langevin Monte Carlo via convex optimization. *The Journal of Machine Learning Research*, 20(1):2666–2711, 2019.

[10] Andreas Eberle. Reflection couplings and contraction rates for diffusions. *Probability theory and related fields*, 166(3-4):851–886, 2016.

[11] Andreas Eberle, Arnaud Guillin, and Raphael Zimmer. Quantitative Harris-type theorems for diffusions and McKean–Vlasov processes. *Transactions of the American Mathematical Society*, 371 (10):7135–7173, 2019.

[12] Martin Hutzenthaler, Arnulf Jentzen, and Peter E Kloeden. Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 467(2130):1563–1576, 2011.

[13] Martin Hutzenthaler, Arnulf Jentzen, and Peter E Kloeden. Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients. *The Annals of Applied Probability*, 22(4):1611–1641, 2012.

[14] Peter E Kloeden and Eckhard Platen. *Numerical Solution of Stochastic Differential Equations*, volume 23. Springer Science & Business Media, 2013.

[15] N. V. Krylov. A Simple Proof of the Existence of a Solution of Itô's Equation with Monotone Coefficients. *Theory of Probability & Its Applications*, 35(3):583–587, 1991. doi: 10.1137/1135082.

[16] Dong-Young Lim and Sotirios Sabanis. Polygonal Unadjusted Langevin Algorithms: Creating stable and efficient adaptive algorithms for neural networks. *Journal of Machine Learning Reserach*, 2024.

[17] Dong-Young Lim, Ariel Neufeld, Sotirios Sabanis, and Ying Zhang. Langevin dynamics based algorithm e-THεO POULA for stochastic optimization problems with discontinuous stochastic gradient. *arXiv preprint arXiv:2210.13193*, 2022.

[18] Dong-Young Lim, Ariel Neufeld, Sotirios Sabanis, and Ying Zhang. Non-asymptotic estimates for TUSLA algorithm for non-convex learning with applications to neural networks with ReLU activation function. *IMA Journal of Numerical Analysis*, 2023.

[19] Wei Liu and Xuerong Mao. Strong convergence of the stopped Euler–Maruyama method for nonlinear stochastic differential equations. *Applied Mathematics and Computation*, 223:389–400, 2013.

[20] Attila Lovas, Iosif Lytras, Miklós Rásonyi, and Sotirios Sabanis. Taming neural networks with tusla: Nonconvex learning via adaptive stochastic gradient langevin algorithms. *SIAM Journal on Mathematics of Data Science*, 5(2):323–345, 2023.

[21] Xuerong Mao. The truncated Euler–Maruyama method for stochastic differential equations. *Journal of Computational and Applied Mathematics*, 290:370–384, 2015.

[22] Ariel Neufeld, Matthew Ng Cheng En, and Ying Zhang. Non-asymptotic convergence bounds for modified tamed unadjusted Langevin algorithm in non-convex setting. *arXiv preprint arXiv:2207.02600*, 2022.

[23] Ariel Neufeld, Matthew Ng Cheng En, and Ying Zhang. Robust SGLD algorithm for solving non-convex distributionally robust optimisation problems. *arXiv preprint arXiv:2403.09532*, 2024.

[24] Maxim Raginsky, Alexander Rakhlin, and Matus Telgarsky. Non-convex learning via Stochastic Gradient Langevin Dynamics: a nonasymptotic analysis. In *Proceedings of the 2017 Conference*

on Learning Theory, volume 65 of *Proceedings of Machine Learning Research*, pages 1674–1703, Amsterdam, Netherlands, 07–10 Jul 2017. PMLR.

[25] Sotirios Sabanis. A note on tamed Euler approximations. *Electronic Communications in Probability*, 18:1–10, 2013.

[26] Sotirios Sabanis and Ying Zhang. Higher order Langevin Monte Carlo algorithm. *Electronic Journal of Statistics*, 13(2):3805 – 3850, 2019. doi: 10.1214/19-EJS1615. URL https://doi.org/10.1214/19-EJS1615.

[27] Sotirios Sabanis and Ying Zhang. A fully data-driven approach to minimizing CVaR for portfolio of assets via SGLD with discontinuous updating. *arXiv preprint arXiv:2007.01672*, 2020.

[28] Pan Xu, Jinghui Chen, Difan Zou, and Quanquan Gu. Global convergence of Langevin dynamics based algorithms for nonconvex optimization. In *Advances in Neural Information Processing Systems*, pages 3122–3133, 2018.

[29] Ying Zhang, Ömer Deniz Akyildiz, Theodoros Damoulas, and Sotirios Sabanis. Nonasymptotic estimates for stochastic gradient Langevin dynamics under local conditions in nonconvex optimization. *Applied Mathematics & Optimization*, 87(2), 2023.

Division of Mathematical Sciences, Nanyang Technological University, 21 Nanyang Link, 637371 Singapore

Email address: ariel.neufeld@ntu.edu.sg

Financial Technology Thrust, Society Hub, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China

Email address: yingzhang@hkust-gz.edu.cn