
\mathcal{L}_1 -DRAC: DISTRIBUTIONALLY ROBUST ADAPTIVE CONTROL

Global Results

Aditya Gahlawat

Mechanical Science and Engineering
University of Illinois at Urbana-Champaign
Urbana, IL 61801
gahlawat@illinois.edu

Sambhu H. Karumanchi

Mechanical Science and Engineering
University of Illinois at Urbana-Champaign
Urbana, IL 61801
shk9@illinois.edu

Naira Hovakimyan

Mechanical Science and Engineering
University of Illinois at Urbana-Champaign
Urbana, IL 61801
nhovakim@illinois.edu

ABSTRACT

Data-driven machine learning methodologies have attracted considerable attention for the control and estimation of dynamical systems. However, such implementations suffer from a lack of predictability and robustness. Thus, adoption of data-driven tools has been minimal for safety-aware applications despite their impressive empirical results. While classical tools like robust adaptive control can ensure predictable performance, their consolidation with data-driven methods remains a challenge and, when attempted, leads to conservative results. The difficulty of consolidation stems from the inherently different ‘spaces’ that robust control and data-driven methods occupy. Data-driven methods suffer from the *distribution-shift* problem, which current robust adaptive controllers can only tackle if using over-simplified learning models and unverifiable assumptions. In this paper, we present \mathcal{L}_1 *distributionally robust adaptive control* (\mathcal{L}_1 -DRAC): a control methodology for uncertain stochastic processes that guarantees robustness certificates in terms of uniform (finite-time) and maximal distributional deviation. We leverage the \mathcal{L}_1 adaptive control methodology to ensure the existence of Wasserstein *ambiguity* set around a nominal distribution, which is guaranteed to contain the true distribution. The uniform ambiguity set produces an *ambiguity tube* of distributions centered on the nominal temporally-varying nominal distribution. The designed controller generates the ambiguity tube in response to both *epistemic* (model uncertainties) and *aleatoric* (inherent randomness and disturbances) uncertainties.

Keywords stochastic control, \mathcal{L}_1 -adaptive control, distributionally robust control, controlled stochastic processes, risk aware control.

1 Introduction

Consider a spectrum of control methodologies with classical tools like robust adaptive control [1, 2] on one end and data-driven approaches relying on data, computation, and deep-learning on the other end [3, 4, 5, 5, 6]. Robust adaptive control was born out of a need for certifiable robustness margins and has had a long and rich history of use with safety-critical applications such as aviation [2]. On the other end, while data-driven methods have displayed impressive empirical performance and can often learn control policies purely from data (e.g., model-free reinforcement learning [7]), such methods lack robustness guarantees and safety guarantees [8].

Ideally, one would like to consolidate the robustness guarantees of the classic methodologies with the empirical performance of the data-driven methodologies. However, such a consolidation is far from a straightforward exercise and,

indeed, is a significant hurdle in the adoption of data-driven controllers/methods for safety-critical systems. In an attempt to provide safety guarantees for systems operating with data-driven learned components in the loop, recent attempts have focused on considering a robust worst-case analysis and bounded disturbances, which lead to overly conservative results, see e.g. [9, 10, 11] and references therein. Instead, one can study average-case (or high-probability) stochastic safety guarantees to alleviate the conservativeness since then one analyzes the distributions and their associated statistical properties instead of purely their supports. Of course, the study of dynamical systems subject to stochastic perturbations brings forth further challenges in their analysis, and thus, a majority of existing work relies on relatively simple models that allow for amenable statistical properties like, e.g., linear systems that preserve the Gaussian nature of perturbations under integration or assumptions that are difficult to verify [12, 13].

In addition to safety guarantees with reduced conservatism, another appealing feature of systems operating under stochastic perturbations is their representation as time-evolving distributions (measures on probability spaces). For example, under moderate regularity conditions, the probability density functions associated with the transition probabilities of solutions to certain stochastic differential equations (SDEs) exist and evolve as per the Kolmogorov forward equation (also known as the Fokker-Planck equation) [14, Chp. 2], [15], [15, Chp. 8]. Of course, along with the distributional representation of systems with stochastic perturbations, we still retain their representation in terms of trajectories (sample paths) as in deterministic systems. The distributional representation of stochastic systems is beneficial for using data-driven learned systems and control policies. Rooted in statistical learning theory [16], one can use concepts from generalization theory like (empirical) risk minimization [17] to train models that achieve low-error on average over the training data distribution. Therefore, while stochastic systems can be represented distributionally, data-driven models are trained over distributions, and this commonality can be exploited to consolidate safe and robust control with the use of data-driven models. It should be noted that deterministic evolution modeled by ordinary/partial differential equations also admit a distributional representation via the degenerate dirac measures concentrated on the system states. Deterministic evolution is therefore not a viable candidate to represent general measures since it has only the single distributional representative. In contrast, even the relatively simple class of SDEs with state-independent diffusion vector fields can represent general distributions as evidenced by their usage in generative modelling [18].

While the distributional nature shared by stochastic systems and data-driven learned models is helpful, one needs to consider the issues that learned models contend with since these will affect the downstream task of their use in the control of systems. A central assumption in the learning-based method is that the training and testing datasets (testing refers to the actual system implementation) are samples from the same distribution [19]. The assumption of the same training and testing (implementation) data distributions allows one to use generalization theory [17] to obtain results like the one in [20] where the authors were able to provide the validity of learned stability certificates to new trajectories initialized from unseen initial conditions, but from the *same* distribution over initial conditions that generated the training data. However, as one may expect, the assumption of encountering the same distribution of scenarios that the learned models are trained on will seldom hold in real-life applications. Thus, a significant issue that learned systems have to contend with is the *distribution shift problem*: the real-life scenario wherein a learned model has to provide predictions and actions in response to an input from a distribution that is different from the distribution it was trained on [21, 22]. Indeed, distribution shift offers a significant hurdle in using learned models and policies in safety-critical systems, see [23, 24] and references therein. Distribution shifts can occur, for example, when deploying a controller learned in a simulator. The controller is learned purely based on the quality (accuracy) of the simulator and its representational capabilities to offer sufficiently rich and realistic scenarios to the learning agent during training. The distribution shift problem that a simulator-based learned controller encounters in its deployment on a real system can be broadly classified under the sim2real transfer problem [25, 26, 27]. Distribution shifts can also occur due to different distributions of initial conditions and the shift between the control policy that generates the training data and the optimized learned policy deployed on the system. While far from trivial to address, researchers have made significant progress in tackling the distribution shifts due to change between data logging and learned policy, as is common in imitation learning (IL), using robustness properties of systems [28].

Perhaps the most significant source of distribution shift is uncertainties in system dynamics models, also known as *epistemic uncertainties*. Note that epistemic uncertainties contribute to the reality gap in the sim2real transfer problem since their effect manifests as the lack of the simulator in representing reality. Furthermore, if one considers stochastic systems subject to random perturbations (*aleatoric uncertainties*), the effects of epistemic uncertainties can be worsened due to the contribution of the aleatoric uncertainties. Of course, as we mentioned above, classical control tools like robust adaptive methodologies were explicitly developed to counter the effects of such uncertainties in a predictable and guaranteed manner. However, the often unrealistic need for explicitly parameterized uncertainties and bounded and deterministic sets to which uncertainties belong causes one to discard the distributional nature of learned models to use the classical tools. As we mentioned previously, in doing so, we obtain conservative results and also lose the distributional representation of learned models and rely solely on bounded and deterministic representations, which contain far less valuable and actionable information. Therefore, we would like to develop a general control

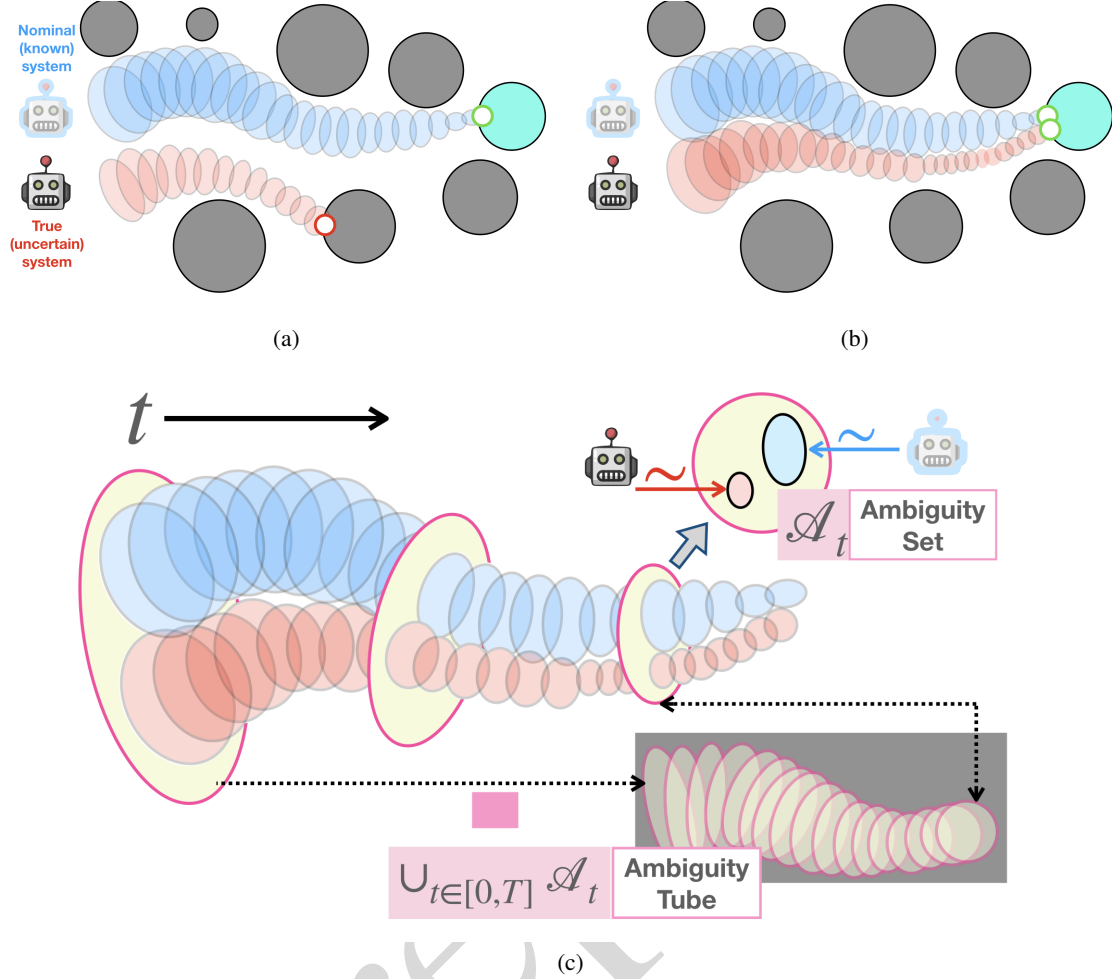


Figure 1: Consider the problem of safely navigating an uncertain stochastic system to a goal set (green circle), avoiding unsafe subsets of the state-space (grey circles). One constructs a control policy for the uncertainty-free nominal (known) system (the robot with faded colors and a blue outline) as it represents the best knowledge available for the true (uncertain) system (solid-colored robot). (a) While the control policy successfully guides the nominal (known) system to the goal set, as we illustrate with temporal state distributions in light blue, applying the same policy to the true (uncertain) system leads to unquantifiable and undesirable behaviors due to the presence of uncertainties (illustrated with light-red temporal state-distributions). (b) Thus, one attempts to design an additional feedback policy to handle uncertainties such that the original policy can still guide the true (uncertain) system predictably and safely. (c) We provide one such approach, the \mathcal{L}_1 -DRAC control, for the design of robust adaptive feedback such that we are assured of having the *a priori* deviations between the nominal (known) system’s state distribution and the true (uncertain) system’s state distribution. These guarantees are in the form of *ambiguity sets* (\mathcal{A}_t) within which both the nominal and true state-distributions are guaranteed to lie, $\forall t \geq 0$. Due to the uniform (finite-time) guarantees on the system’s distributional transients, one can extrude such ambiguity tubes in time to obtain *ambiguity tubes* ($\bigcup_{t \in [0, T]} \mathcal{A}_t$, for any $T \in (0, \infty)$), which enables safe predictive planning.

methodology for stochastic systems — subject to both epistemic and aleatoric uncertainties — to produce certificates of robustness that are distributional in nature. Such approaches generalize the notion of robustness since the certificates are defined, not on the underlying state space, but instead on the set of probability measures on the state spaces. We posit that the development of the stated methodology can facilitate seamless and non-conservative integration of robust control with data-driven learned models that can be verified to produce predictable and, hence, safe behavior. In other words, instead of restricting the capabilities of data-driven learned models to make them amenable to consolidation with classical control tools, we aim to raise the abstractions of robust-adaptive control tools so that such tools are conducive to working with learned models by design.

We present \mathcal{L}_1 -distributionally robust adaptive control (\mathcal{L}_1 -DRAC): a robust adaptive methodology to control uncertain stochastic processes whose evolution is governed by stochastic differential equations (SDEs). We design \mathcal{L}_1 -DRAC such that we can quantify and control the ‘distance’ between the state distributions of the known (nominal) system and the true (uncertain) system. We use the Wasserstein norm [29, Chp. 6] as a metric on the space of Borel measure to quantify the distance between the nominal and true distributions. We use *ambiguity sets*¹ defined via the Wasserstein distance between the nominal and true state distributions as the robustness certificate against distribution shifts due to epistemic and aleatoric uncertainties. We develop \mathcal{L}_1 -DRAC using the architecture of \mathcal{L}_1 -adaptive control [30], a robust adaptive methodology that decouples estimation from control and provides transient guarantees in response to uncertainties. The \mathcal{L}_1 -adaptive control has been successfully implemented on NASA’s AirStar 5.5% sub-scale generic transport aircraft model [31], Calspan’s Learjet [32], and uncrewed aerial vehicles [33, 34]. A high-level illustration of the goals and capabilities of \mathcal{L}_1 -DRAC are illustrated in Fig. 1.

1.1 Prior art

We now discuss the existing results in the literature that provide results for problems similar to the one we consider in this manuscript.

(i) **Uniform (finite-time) guarantees:** We begin by discussing results that provide uniform (finite-time) guarantees for controlled stochastic systems. The authors in [35] and [36] consider asymptotic reference tracking for discrete-time stochastic systems. However, the safe operation of (stochastic and uncertain) systems requires uniform (finite-time) guarantees instead of asymptotic guarantees. The closed-loop system’s behavior should remain predictable $\forall t \geq 0$, not just when $t \rightarrow 0$. Results on finite-time (uniform) guarantees for stochastic systems in discrete-time can be found in [37], and for continuous-time in [38] and [39]. Using a different analysis, the authors in [39] were able to provide bounds on higher-order moments leading to tighter tracking error bounds. Furthermore, the results in [36] and [40] require linearity of the systems under consideration.

(ii) **Control-theoretic approaches:** Control Lyapunov function (CLF)-based approaches [41, 42] are the most well-known and applied methods for controlling perturbed deterministic systems. Such methods thus also led to the development of Lyapunov-based approaches for stochastic systems [43, 44, 45, 46] wherein notions like that of globally asymptotically stable in probability are used [47, Chp. 2], [48, Chp. 5]. Asymptotic stability in probability is a property of the trivial solution (similar to the deterministic counterpart) and hence applies only to systems whose noise vector field vanishes at zero [48, Sec. 5.1]. The absence of such constraints prevents the stabilization of the trivial solution and thus requires robust approaches, e.g., see [43, Sec. 4]. Still, however, being Lyapunov-based approaches, as in the references mentioned above, one attempts to compute a stochastic Lyapunov function which depends explicitly on the noise (diffusion) vector field [47, Lem. 2.1], [48, Thm. 5.3]. Such requirements exacerbate the already challenging problem of synthesizing (control) Lyapunov functions.

An approach that avoids the synthesis of stochastic Lyapunov functions is to consider stochastic systems whose robust stability can be determined owing to the stability of the deterministic system counterpart (noiseless stochastic system). For example, the authors in [39] derive the stability of a stochastic system using a Lyapunov function for the deterministic counterpart system. However, this approach requires synthesizing a CLF for the deterministic system, which is still no trivial task for nonlinear systems. As an alternative, contraction theory-based solutions offer a computationally tractable convex formulation for searching CLFs for nonlinear systems [49]. In fact, contraction theory offers necessary *and* sufficient characterization for the stability of trajectories of nonlinear systems [50]. A few examples of contraction theory-based solutions for control and estimation of nonlinear deterministic systems can be found in, e.g. [51, 52] and in [53, 38] for stochastic systems. The tutorial paper [54] provides a comprehensive and exhaustive discussion on using contraction theory to control and estimate nonlinear systems. Using the robustness of the contraction-based controllers, the authors in [55, 54] prove the robustness of the stochastic counterpart systems. However, as the authors in [39] observe, the transfer of the stability properties of the deterministic systems to their stochastic counterparts only holds under certain conditions [15]. This fact holds for the robust controllers in [55, 54], which apply only to stochastic systems whose noise (diffusion) vector fields are uniformly spatially and temporally bounded. The authors in [56] developed a learning-based synthesis of contraction theory-based control for nonlinear stochastic systems; however, similar to [55], the robustness guarantees only hold for systems with bounded noise (diffusion) vector fields.

In addition to the references provided above, a further few examples of adaptive control for uncertain nonlinear stochastic systems can be found in, e.g. [57] where the authors consider a strict-feedback system with time-delays and utilize a stochastic Lyapunov function to show global asymptotic stability in probability, and [58], wherein the authors consider output-feedback tracking for a system with linearly parameterized uncertainties and bounded noise (diffusion)

¹Sets of distributions (Borel measures on vector spaces over the field of reals).

vector field. Using deep-learning methods, the authors in [59] (see also [54, Sec. 8]) provide adaptive control schemes with robustness guarantees for affine and multiplicatively-separable parametric uncertainties and uniformly bounded disturbances. The assumption on bounded disturbances precludes the use of very general classes of random processes like the Lévy processes with Brownian motion as an example with continuous paths [60]. Similarly, the assumption on structured parametric uncertainties can also be too restrictive to verify. Another example of learning-based adaptive control synthesis can be found in [61], where the authors pose the synthesis of adaptive controllers as a meta-learning problem. The authors used model ensembling to counter epistemic uncertainties and displayed robustness to distribution shifts. However, this was achieved empirically, thus preventing its use for *a priori* predictive planning. Furthermore, the distribution shift was shown for distributions with compact supports thus unable to harness the generality of unbounded support of Lévy processes like the Brownian motion. The result in [62] also uses \mathcal{L}_1 -adaptive control, albeit for linear systems, with additive Brownian motion and for linearly parametrized uncertainties affecting solely the drift vector field.

(iii) Learning-based control: One can find recent results on handling epistemic uncertainties and the associated distribution shift under the umbrella term of learning-based control. For example, the authors in [63] consider discrete-time deterministic systems and combine density models with Lyapunov functions. Hence, the learned agent remains in, or close to, the training distribution during implementation (testing). While the work in [63] and ours share the same goal, the approach taken by the authors is primarily a learning-based data-driven approach where the system is controlled to remain near the training data distribution, thus avoiding high-uncertainty subsets of the state space. Reachability analysis for uncertain systems is a popular and effective approach for the safe control of uncertain systems, e.g. [64] wherein the authors use the Hamilton-Jacobi reachability analysis for systems under unknown but bounded disturbances. Similarly, the authors in [65] analyze goal reachability for a class of uncertain stochastic systems under unknown but bounded stochastic disturbances. As we stated before, using bounded stochastic disturbances disregards beneficial statistical properties of the disturbances and uncertainties, and hence the state distributions. Thus, one must then rely solely on considering the supports of the distributions leading to conservative analyses. While the assumption of bounded disturbances simplifies the analysis relatively since, e.g., as in [65], the system state is differentiable in time even in the presence of stochastic perturbations, such assumptions exclude the use of models driven by Lévy processes like Brownian motion [66]. Reachability analysis of systems perturbed by noise with non-compactly supported distributions is a challenging prospect as evidenced by recent results in [67, 68, 69] wherein the authors consider linear systems with additive random disturbances that are possibly correlated, and [70] wherein the authors consider both linear and nonlinear discrete-time systems with the vector fields assumed to be known. The work in [71] considers reachability analysis for nonlinear discrete-time stochastic systems in terms of probability measures corresponding to the state distributions. Finally, we refer the interested reader to [72] and [73] where the authors develop non-conservative and computationally efficient approaches for reachability of known continuous-time stochastic systems driven by Brownian motion.

(iv) Distributionally robust control: Distributionally robust optimization (DRO) is a sub-field of mathematical optimization that considers obtaining extrema of distributionally ambiguous (uncertain) cost functions, see e.g. [74, 75, 76] and references therein. Therefore, it stands to reason that DRO can assist in achieving the same goals for systems with distributional uncertainties, similar to how one synthesizes robust controllers against systems with compactly supported uncertainties (\in compact sets). Indeed, several recent results use DRO for distributionally robust control (DRC). The results in [77, 78] consider disturbances with unknown distributions in discrete time; however, the dynamics are assumed to be linear with the disturbances affecting the systems additively. Under the same assumptions of linear dynamics and additive disturbances, the authors in [79] provided a distributionally robust integration of perception, planning, and control. While the authors in [80] consider DRC for nonlinear systems, the noise is assumed to perturb the system additively. The authors in [81, 82] consider nonlinear systems with disturbances of unknown distributions also affecting the system in a nonlinear fashion; however, the authors assume the availability of a (finite) number of samples from the actual disturbance distribution along with the accurate knowledge of the dynamics. The authors in [83] recently demonstrated the use of DRC for partially observable, albeit linear systems. When samples from the true distributions are unavailable, assumptions on the existence of an ambiguity set of distributions is assumed as in [84]. Instead of considering systems with distributionally uncertain disturbances, the DRO formulation can also be used for the case of uncertain environments, see e.g. [85, 86, 87, 88].

1.2 Contributions

The following are the contributions and the features of the \mathcal{L}_1 -DRAC control compared to existing results above in Sec. 1.1.

(i) The design of \mathcal{L}_1 -DRAC control ensures the existence of ***a priori* computable uniform (finite-time) guarantees** for the closed-loop system in terms of **maximal deviation between the distributions** (probability measures) of the

true (uncertain) stochastic system and its nominal (known) version. The uniform bounds of distributional deviation are in the form of **Wasserstein metric** that define the **guaranteed ambiguity sets (and tubes)**, see Fig. 1(c). Moreover, \mathcal{L}_1 -DRAC's guarantees mirror those of the nominal (known) version; \mathcal{L}_1 -DRAC can ensure the existence of *a priori* Wasserstein ambiguity sets (tubes) **up to any order identical to the nominal (known) system**, thus leading to, e.g., tighter tracking errors. Unlike the existing results we mention in Sec. 1.1, we provide finite-time guarantees for **uncertain nonlinear stochastic systems**, thus avoiding the assumption of accurate knowledge of the system dynamics or their structure, e.g., linear dynamics and additive disturbances.

(ii) The \mathcal{L}_1 -DRAC control relies only on the stability of the deterministic counterpart of the known (nominal) subsystem as in [39, 56, 55]. However, when compared to the existing control-theoretic approaches, \mathcal{L}_1 -DRAC is applicable to **nonlinear systems with (globally) unbounded and state-dependent diffusion and drift uncertainties** which leads to **instability of the true (uncertain) system** despite the stability of the nominal (known) system. The **uncertainties are not required to have any parametric structures** and require mild assumptions on their growth but not their global boundedness. Despite the unbounded and non-parametric uncertainties, \mathcal{L}_1 -DRAC is applicable to systems driven by Brownian motion.

(iii) Unlike the existing learning-based approaches, the design of \mathcal{L}_1 -DRAC **does not use data-driven learning** to produce guarantees of uniform and maximal distribution shift and instead relies on an **adaptive mechanism** to compensate for the uncertainties. While the nominal (known) system and a controller to stabilize it might be learned from data, the distributional robustness that the \mathcal{L}_1 -DRAC controller guarantees does not require learning. Since the \mathcal{L}_1 -DRAC control ensures uniform guarantees of maximal distribution-shift, we consider \mathcal{L}_1 -DRAC **as an approach where one uses control to enable safe use of data-driven learning**, as opposed to using learning to enable safe control.

A significant benefit of using \mathcal{L}_1 -DRAC control is the **ease of ensuring the safety of uncertain nonlinear stochastic systems** when compared to the existing state-of-the-art. For example, \mathcal{L}_1 -DRAC control can provide a **computationally inexpensive (optimization-free)** approach to compute **reachable sets or collision-free plans for safe operation of uncertain nonlinear stochastic systems**. Indeed, the *a priori* existence of ambiguity sets that \mathcal{L}_1 -DRAC control ensures (see Fig. 1) provides a reachable set of state-distributions (probability measures) for any $t \geq 0$. Thus, one only needs to analyze the reachability of the *nominal (known) system*, and the ambiguity set guarantees “snap-on” to guarantee reachability for the true (uncertain) system without any additional computation and optimization.

(iv) The results on distributionally robust control (DRC) we reference above consider accurate dynamics but are perturbed by disturbances of unknown distributions, which implies uncertain distributions over states despite the accurate knowledge of the dynamics. Thus, one may assume the accurate dynamics to correspond to the known (nominal) version of the true (uncertain) systems. At the same time, the unknown distribution of the disturbance subsumes the effects of epistemic uncertainties and noise. We instead consider **uncertain nonlinear systems driven by Brownian motion**. Even though the distribution of Brownian motion is well-known, e.g., normally distributed and independent increments [89, 60], the presence of state-dependent uncertainties in *both* drift and diffusion terms imply that the **random perturbations affecting the systems possess unknown distributions, are correlated, appear multiplicatively, can affect the control channel, and thus lead to uncertain state distributions**. Therefore, while the DRC results and our work consider different systems regarding the knowledge of the dynamics and the distribution of disturbance/noise, both lead to uncertain state distributions.

However, unlike the state-of-the-art for DRC, \mathcal{L}_1 -DRAC **does not require *a priori* existence or knowledge of ambiguity sets for the true (uncertain) state distributions, nor the *a priori* availability of (finite) samples from the uncertain distributions**. Indeed, the existence and knowledge of ambiguity sets are not guaranteed for the uncertain systems we consider. We do not make the unverifiable assumption of *a priori* availability of samples from the true (uncertain) state distributions because the \mathcal{L}_1 -DRAC controller uses state-feedback, and hence, it **receives exactly one sample from the time-varying uncertain state distribution at each point in time**. We emphasize the point that \mathcal{L}_1 -DRAC retains a standard controller implementation operating under state-feedback while offering robustness guarantees in the space of distributions. Thus, unlike a majority of approaches where the control law's feedback space and the certification space are identical (e.g. feedback \in state space \mathbb{R}^n and invariant sets also $\in \mathbb{R}^n$), the \mathcal{L}_1 -DRAC

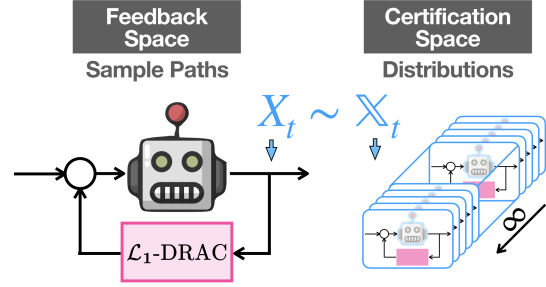


Figure 2: The \mathcal{L}_1 -DRAC controller is designed for a standard implementation and operates under state feedback (the state is denoted by X_t). However, \mathcal{L}_1 -DRAC provides certificates on the state distribution \mathbb{X}_t ($X_t \sim \mathbb{X}_t$).

is designed for the case when the feedback space (state space) \neq certification space (probability measures on the state space), see Fig. 2. Under slightly stronger conditions than the ones required for well-posedness for nonlinear SDEs, we show 1) the **existence of the ambiguity sets (tubes) for the true (uncertain) distributions**, 2) the existence of the ***a priori* computable radius of ambiguity sets over temporal state measures**, and 3) a certain amount of control on the radius of the ambiguity sets **allowing us to exercise a degree of control, under reasonable tradeoffs, on limiting the distribution shift**.

1.3 Article Organization

We present the systems under consideration in Sec. 2.1, followed by the required assumptions, their justifications, and consequences in Sec. 2.2. The problem statement is stated in Sec. 2.3. Section 3 presents the core results on the design and analysis of the \mathcal{L}_1 -DRAC law. Section 3.1 provides the architecture of \mathcal{L}_1 -DRAC, and Sec. 3.2 provides the design choice of the pertinent control parameters. The performance analysis of \mathcal{L}_1 -DRAC is provided in Sec. 4, and the associated discussion is presented in Sec. 4.3. The manuscript is concluded with Sec. 5.

A majority of derivations are presented in the appendix: Appendix A collects all the pertinent definitions as a quick reference, Appendix B contains the supporting technical results, and Appendices C and D provide derivations for the analysis in Sec. 4.1 and 4.2, respectively.

1.4 Notation

We define $a \wedge b \doteq \min\{a, b\}$, for any $a, b \in \mathbb{R}$, and $\mathbb{1}_{x \in A}(x)$ is the indicator function of set A evaluated at x . We denote by $\mathbb{R}_{>0}$ ($\mathbb{R}_{\geq 0}$) the set of positive (non-negative) reals. For any $\alpha \in \mathbb{N}$, we denote by $\mathbb{N}_{\geq \alpha}$ the set of naturals $\geq \alpha$. The set of linear operators $\mathbb{R}^n \rightarrow \mathbb{R}^m$, $n, m \in \mathbb{N}$, are denoted by $\mathbb{R}^{n \times m}$, while \mathbb{S}^n denotes self-adjoint operators $\mathbb{R}^n \rightarrow \mathbb{R}^n$. We denote by $\mathbb{S}_{>0}^n$ ($\mathbb{S}_{\geq 0}^n$) the set of positive (semi)definite symmetric maps on \mathbb{R}^n . Unless explicitly stated, $\|\cdot\|$ denotes the Euclidean 2-norm on \mathbb{R}^n , $n \in \mathbb{N}$. On the space $\mathbb{R}^{n \times m}$, $n, m \in \mathbb{N}$, $\|\cdot\|$ denotes the induced 2-norm, while we denote the Frobenius norm on the same space by $\|\cdot\|_F$. The set of all continuous maps from some \mathcal{X} to \mathcal{Y} is denoted by $\mathcal{C}(\mathcal{X}; \mathcal{Y})$. Similarly, $\mathcal{C}^k(\mathcal{X}; \mathcal{Y})$, $k \in \{\mathbb{N}, \infty\}$, denotes k -times continuously differentiable maps. For any $f \in \mathcal{C}^2(\mathbb{R}^n; \mathbb{R})$

$$\begin{aligned}\nabla f(x) &\doteq [\partial_{x_1} f(x) \quad \cdots \partial_{x_n} f(x)]^\top \in \mathbb{R}^n, \\ \nabla^2 f(x) &\doteq \nabla \cdot (\nabla f)^\top = [\nabla \partial_{x_1} f(x) \quad \cdots \nabla \partial_{x_n} f(x)] \in \mathbb{S}^n.\end{aligned}$$

Similarly,

$$\nabla g(x) \doteq [\nabla g_1(x) \quad \cdots \quad \nabla g_m(x)] \in \mathbb{R}^{n \times m}, \quad g \in \mathcal{C}^1(\mathbb{R}^n; \mathbb{R}^m).$$

For $T, p \in (0, \infty)$ and any finite-dimensional space \mathcal{X} , we define

$$\mathfrak{L}_p([0, T]; \mathcal{X}) = \left\{ z : [0, T] \rightarrow \mathcal{X} \mid \|z\|_{\mathfrak{L}_p} = \left(\int_0^T \|z(\nu)\|_{\mathcal{X}}^p d\nu \right)^{\frac{1}{p}} < \infty \right\},$$

and we say $z \in \mathfrak{L}_p^{loc}(\mathcal{X})$ if $z \in \mathfrak{L}_p([0, T]; \mathcal{X})$ for every $T > 0$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space that we consider as the underlying space throughout the manuscript. For a filtration \mathcal{F}_t on the probability space above, we define²

$$L_p([0, T]; \mathcal{X} \mid \mathcal{F}_t) = \left\{ \mathfrak{F}_t\text{-adapted } z(t, \cdot) : \Omega \rightarrow \mathcal{X} \mid z(\cdot, \omega) \in \mathfrak{L}_p([0, T]; \mathcal{X}), \quad \omega \in \Omega \right\},$$

and $z \in L_p^{loc}(\mathcal{X} \mid \mathcal{F}_t)$ if $z \in L_p([0, T]; \mathcal{X} \mid \mathcal{F}_t)$ at every $T > 0$ assuming the filtration exists. We similarly define

$$\mathcal{M}_p([0, T]; \mathcal{X} \mid \mathcal{F}_t) = \left\{ z \in L_p([0, T]; \mathcal{X} \mid \mathcal{F}_t) \mid \mathbb{E} \left[\|z\|_{\mathfrak{L}_p}^p \right]^{\frac{1}{p}} = \left(\int_{\Omega} \|z(\cdot, \omega)\|_{\mathfrak{L}_p}^p \mathbb{P}(d\omega) \right)^{\frac{1}{p}} < \infty \right\},$$

and $\mathcal{M}_p^{loc}(\mathcal{X} \mid \mathcal{F}_t)$ is defined analogously to $L_p^{loc}(\mathcal{X} \mid \mathcal{F}_t)$. With an abuse of notation, for any \mathfrak{F}_t -adapted $z \in \mathcal{C}([0, T]; \mathcal{X})$ we write

$$\|z(t)\|_{L_p} = \mathbb{E} \left[\|z\|_{\mathfrak{L}_p}^p \right]^{\frac{1}{p}} = \left(\int_{\Omega} \|z(\cdot, \omega)\|_{\mathfrak{L}_p}^p \mathbb{P}(d\omega) \right)^{\frac{1}{p}} < \infty, \quad t \in [0, T].$$

²Identical definitions can be found in, e.g., \mathcal{L}_{ad} in [90, Notation 5.1.1] and \mathcal{L}^p in [91].

Note that $\|\cdot\|_{L_p}$ is **not a norm** on the space L_p defined above since the integrand in $\|\cdot\|_{L_p}$ uses the finite-dimensional norm $\|\cdot\|_{\mathcal{X}}$, and not the infinite-dimensional $\|\cdot\|_{\mathcal{L}_p}$.

Remark 1.1 While it is standard in the literature to represent the sample points $\in \Omega$ by ω , we use the upright version $\omega \in \Omega$ instead. This allows us to use the standard version ω to represent the control parameter of filter bandwidth (yet to be defined). Moreover, as is standard, we suppress the argument $\omega \in \Omega$ for functions with domain Ω . Thus, from this point onward and unless explicitly stated, ω denotes a control parameter, and not a sample point of Ω .

The Borel σ -algebra over a topological space \mathcal{Y} is denoted by $\mathcal{B}(\mathcal{Y})$, and $\sigma(z)$ denotes the σ -algebra generated by any measurable $z : (\Omega, \mathcal{F}) \rightarrow (\mathcal{Y}, \mathcal{B}(\mathcal{Y}))$. The notation $\mathcal{F}_1 \vee \mathcal{F}_2$ denotes the join of σ -algebras \mathcal{F}_1 and \mathcal{F}_2 .

The Wasserstein metric of order $p \in [0, \infty)$ between probability measures ς_1 and ς_2 on a Polish space (complete and separable metric space [92, Defn. 18.1]) $(\mathcal{Z}, d_{\mathcal{Z}})$ is given as follows [29, Defn. 6.1]:

$$W_p(\varsigma_1, \varsigma_2) = \left(\inf_{\pi \in \Pi(\varsigma_1, \varsigma_2)} \int_{\mathcal{Z}} d_{\mathcal{Z}}(z_1, z_2)^p \pi(dz_1, dz_2) \right)^{\frac{1}{p}},$$

where $d_{\mathcal{Z}}$ is the distance function on \mathcal{Z} , and $\Pi(\varsigma_1, \varsigma_2)$ is the set of all couplings between measures ς_1 and ς_2 .

2 Systems, Assumptions, and Problem Statement

We begin with the definitions of the stochastic processes under consideration.

2.1 The Systems

The following defines the known and unknown drift and diffusion vector fields.

Definition 1 (Vector Fields) Consider the **known functions** $f : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, $g : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^{n \times m}$, and $p : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times d}$, for $n, m, d \in \mathbb{N}$. Consider also the **unknown functions** $\Lambda_{\mu} : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $\Lambda_{\sigma} : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times d}$. For any $(a, b, t) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_{\geq 0}$, we denote by

$$F_{\mu}(t, a, b) \doteq f(t, a) + g(t)b + \Lambda_{\mu}(t, a) \in \mathbb{R}^n, \quad F_{\sigma}(t, a) \doteq p(t, a) + \Lambda_{\sigma}(t, a) \in \mathbb{R}^{n \times d} \quad (1)$$

the **true (uncertain) drift and diffusion vector fields**, respectively. Similarly, for any $(a, t) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}$, we denote by

$$\bar{F}_{\mu}(t, a) \doteq f(t, a) \in \mathbb{R}^n, \quad \bar{F}_{\sigma}(t, a) \doteq p(t, a) \in \mathbb{R}^{n \times d} \quad (2)$$

the **nominal (known) drift and diffusion vector fields**, respectively. Note that we have the following decompositions

$$F_{\mu}(t, a, b) = \bar{F}_{\mu}(t, a) + g(t)b + \Lambda_{\mu}(t, a) \in \mathbb{R}^n, \quad F_{\sigma}(t, a) = \bar{F}_{\sigma}(t, a) + \Lambda_{\sigma}(t, a) \in \mathbb{R}^{n \times d}. \quad (3)$$

We now use the vector fields in Definition 1 to define the true and nominal processes that we study with respect to each other in this manuscript. The notations in the following mostly follow the convention in [15]. The underlying complete probability space is $(\Omega, \mathcal{F}, \mathbb{P})$ that we introduced in Sec. 1.4.

Definition 2 (Processes) We denote by W_t and W_t^* any two **independent** \mathbb{P} -Brownian motions. The filtrations generated by W_t and W_t^* are denoted by \mathfrak{W}_t and \mathfrak{W}_t^* , respectively, and we define $\mathfrak{W}_{\infty} = \sigma(\cup_{t \geq 0} \mathfrak{W}_t)$ and $\mathfrak{W}_{\infty}^* = \sigma(\cup_{t \geq 0} \mathfrak{W}_t^*)$. Let $x_0 \sim \xi_0$ and $x_0^* \sim \xi_0^*$ be two \mathbb{R}^n -valued random variables that are respectively independent of \mathfrak{W}_{∞} and \mathfrak{W}_{∞}^* , where ξ_0 and ξ_0^* are probability measures on $\mathcal{B}(\mathbb{R}^n)$. We further define $\mathfrak{W}_{0,t} = \sigma(\xi_0) \vee \mathfrak{W}_t$ and $\mathfrak{W}_{0,t}^* = \sigma(\xi_0^*) \vee \mathfrak{W}_t^*$.

For any $T \in (0, \infty)$, we say that $X, X^* \in \mathcal{C}([0, T]; \mathbb{R}^n)$ are the **true (uncertain) and nominal (known) processes**, respectively, if they are respectively adapted to the filtrations $\mathfrak{W}_{0,t}$ and $\mathfrak{W}_{0,t}^*$, and are the unique strong solutions [60, Defn. 5.2.1] to the following Itô stochastic differential equations (SDEs), for all $t \in [0, T]$:

$$dX_t = F_{\mu}(t, X_t, U_t) dt + F_{\sigma}(t, X_t) dW_t, \quad X_0 = x_0 \sim \xi_0 \text{ (}\mathbb{P}\text{-a.s.)}, \quad (4a)$$

$$dX_t^* = \bar{F}_{\mu}(t, X_t^*) dt + \bar{F}_{\sigma}(t, X_t^*) dW_t^*, \quad X_0^* = x_0^* \sim \xi_0^* \text{ (}\mathbb{P}\text{-a.s.)}, \quad (4b)$$

where the vector fields $F_{\{\mu, \sigma\}}$ and $\bar{F}_{\{\mu, \sigma\}}$, are presented in Definition 1, and U_t is some yet-to-be-defined process. We refer to U_t as the **feedback process**.³

We define the **true law** \mathbb{X}_t as

$$\mathbb{X}_t(B) \doteq \mathbb{P}(X_t^{-1}(B)), \quad (t, B) \in [0, T] \times \mathcal{B}(\mathbb{R}^n).$$

The **nominal law** \mathbb{X}_t^* is defined analogously for X_t^* . Thus, $X_t \sim \mathbb{X}_t$ and $X_t^* \sim \mathbb{X}_t^*$.

³We will drop the quantifier \mathbb{P} -a.s. from here on.

The independence of the Brownian motions W_t and W_t^* follows from the fact that distinct real-world systems will be affected by distinct realizations of the random noise. We will later discuss the consequences on the main results of the manuscript. See [38] for a further discussion.

Remark 2.1 Since $X \in \mathcal{C}([0, T]; \mathbb{R}^n)$, instead of considering X_t as a pointwise-in-time random variable on \mathbb{R}^n , one may consider the trajectory (sample path) $X_{[0, T]} \doteq (X_t)_{0 \leq t \leq T}$ as a random variable on $\mathcal{C}([0, T]; \mathbb{R}^n)$. This then motivates the definition of the (path) law $\mathbb{X}_{[0, T]}$, $X_{[0, T]} \sim \mathbb{X}_{[0, T]}$, as follows [60, Sec. 2.4, p. 60]:

$$\mathbb{X}_{[0, T]}(B) = \mathbb{P}\left(X_{[0, T]}^{-1}(B)\right), \quad (t, B) \in [0, T] \times \mathcal{B}(\mathcal{C}([0, T]; \mathbb{R}^n)).$$

Along with the separability of strong solutions [93, Thm. 5.2.1], the Kolmogorov extension theorem [60, Thm. 2.2] implies that the finite-dimensional cylinder sets form a determining class for the probability measures on $\mathcal{B}(\mathcal{C}([0, T]; \mathbb{R}^n))$ [94, Thm 2.0]. Consequently, the path law can be represented alternatively as follows [15, Sec. 7.1]:

$$\mathbb{X}_{t_1 \dots t_k}(B_1 \times \dots \times B_k) = \mathbb{P}(X_{t_1}^{-1}(B_1), \dots, X_{t_k}^{-1}(B_k)),$$

for all choice of temporal instances $0 < t_1 < \dots < t_k \leq T$, and all $B_1, \dots, B_k \in \mathcal{B}(\mathbb{R}^n)$ [94, Sec.2.1, p. 12]. We thus see that the law \mathbb{X}_t , $t \in [0, T]$, in Definition 2 can be interpreted as the probability measure defined on a cylinder set with $k = 1$.

2.2 Assumptions

We begin with the assumptions for the nominal (known) system in (4b).

Assumption 1 (Nominal (Known) System) The known drift function $f(t, a)$ in Definition 1 is locally Lipschitz in $a \in \mathbb{R}^n$, uniformly in $t \in \mathbb{R}_{\geq 0}$, and there exists a known $\Delta_f \in \mathbb{R}_{>0}$ such that

$$\|f(t, a)\|^2 \leq \Delta_f^2 (1 + \|a\|^2), \quad \forall (t, a) \in \mathbb{R}_{\geq 0} \times \mathbb{R}^n.$$

Furthermore, the known diffusion function $p(t, a)$ in Definition 1 is uniformly bounded and globally Lipschitz in its arguments, and there exist known $\Delta_p, \hat{L}_p, L_p \in \mathbb{R}_{>0}$ such that

$$\|p(t, a)\|_F \leq \Delta_p, \quad \|p(t, a) - p(t', a')\|_F \leq \hat{L}_p |t - t'| + L_p \|a - a'\|^{\frac{1}{2}}, \quad \forall (t, t', a, a') \in \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}^n \times \mathbb{R}^n.$$

The input operator $g : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^{n \times m}$ has full column rank $\forall t \in \mathbb{R}_{\geq 0}$ and satisfies

$$g \in \mathcal{C}^1([0, \infty); \mathbb{R}^{n \times m}), \quad \|g(t)\|_F \leq \Delta_g, \quad \|\dot{g}(t)\|_F \leq \Delta_{\dot{g}}, \quad \forall t \in \mathbb{R}_{\geq 0},$$

where $\Delta_g, \Delta_{\dot{g}} \in \mathbb{R}_{>0}$ are assumed to be known. Since $g(t)$ is full rank, we can construct a $g^\perp : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^{n \times n-m}$ such that $\text{Im } g(t)^\perp = \ker g(t)^\top$ and $\text{rank } (\bar{g}(t)) = n$, $\forall t \in \mathbb{R}_{\geq 0}$ where

$$\bar{g}(t) \doteq [g(t) \quad g(t)^\perp] \in \mathbb{R}^{n \times n}.$$

We assume that $g^\perp(t) \in \mathbb{R}^{n \times n-m}$ is uniformly Lipschitz continuous in $t \in \mathbb{R}_{\geq 0}$, and $\|g(t)^\perp\|_F \leq \Delta_g^\perp$, $\forall t \in \mathbb{R}_{\geq 0}$, where $\Delta_g^\perp \in \mathbb{R}_{>0}$ is assumed to be known. We further assume there exists a known scalar $\Delta_\Theta \in \mathbb{R}_{>0}$ such that

$$\|\Theta_{ad}(t)\|_F \leq \Delta_\Theta, \quad \forall t \in \mathbb{R}_{\geq 0}, \quad \text{where } \Theta_{ad}(t) \doteq [\mathbb{I}_m \quad 0_{m, n-m}] \bar{g}(t)^{-1} \in \mathbb{R}^{m \times n}.$$

Finally, the initial condition $x_0^* \sim \xi_0^*$ satisfies $\|x_0^*\|_{L_{2p^*}} < \infty$, for some $p^* \in \mathbb{N}_{\geq 1}$.

Remark 2.2 We assume the uniform boundedness of the known diffusion term $p(t, x)$, instead of linear growth, since it is a sufficient condition to transfer the stability of the deterministic counterpart of (4b) ($W_t^* \equiv 0$) to the stochastic version via robustness arguments, see e.g., [38, 54, 55]. However, as we will see below (see Assumption 4 and the subsequent comments in Remark 4), we do not place any such assumptions on the diffusion term of the true (uncertain) system in (4a). Therefore, the stability of (4a) is not guaranteed by the stability of its deterministic counterpart.

Remark 2.3 The assumption on the initial condition implies bounded second moments as a minimum. Along with the initial condition's independence from \mathcal{W}_∞^* (see Defn. 2), this assumption is standard for the well-posedness of SDEs, see e.g. [Thm. 5.2.1][15].

We now place assumptions on the stability of the deterministic counterpart ($W_t^* \equiv 0$) of the nominal (known) system in (4b). We need the following definition of the induced logarithmic norm of $A \in \mathbb{R}^{n \times n}$:

$$\mu(A) = \lim_{h \rightarrow 0^+} \frac{\|\mathbb{I}_n + hA\| - 1}{h}.$$

The induced logarithmic norm (*log norm* for brevity) is guaranteed to exist and can be interpreted as the derivative of e^{At} in the direction of A and evaluated at \mathbb{I}_n . Refer to [95, Sec. 2.3] for further details.

Assumption 2 (Nominal (Known) System Stability) *There exists a known $\lambda \in \mathbb{R}_{>0}$ such that*

$$\mu(\nabla_x \bar{F}_\mu(t, x)) \leq -\lambda, \quad \forall (t, x) \in \mathbb{R}_{\geq 0} \times \mathbb{R}^n, \quad (5)$$

where \bar{F}_μ is defined in (2).

The condition in (5) allows one to use contraction theory to conclude the *incremental exponential stability (IES)* of the deterministic system $\dot{x} = \bar{F}_\mu(t, x)$ [95, Thm. 3.9]. The following lemma provides a vital consequence of the above.

Lemma 2.1 [[96]] *Given the deterministic system $\dot{x} = \bar{F}_\mu(t, x)$, the statement in (5) is equivalent to*

$$(x - y)^\top (\bar{F}_\mu(t, x) - \bar{F}_\mu(t, y)) \leq -\lambda \|x - y\|^2, \quad \forall (t, x, y) \in \mathbb{R}_{\geq 0} \times \mathbb{R}^n \times \mathbb{R}^n. \quad (6)$$

The use of the matrix measure (log norm) is one of the many approaches for characterizing the contractive property of the vector field $\bar{F}_\mu(t, x)$ [97, 98, 96]. Additional approaches include: incremental Lyapunov Functions [50, 99], Finsler-Lyapunov functions [100, 101]; and Riemannian metric conditions [49, 102]. From a synthesis perspective, the concept of a *control contraction metric (CCM)*, a Riemannian metric, has shown immense utility for controlled systems, e.g., [52, 11, 54].

Since we aim to establish the stability of the true (uncertain) process (4a) with respect to the nominal (known) process (4b), it is a reasonable requirement that the latter satisfies certain properties that represent the desired behavior that we wish to endow upon the uncertain system. We state these in the next assumption on the statistical nature of (4b).

Assumption 3 *There exists a known $\Delta_\star \in \mathbb{R}_{>0}$, such that, for any $T \in (0, \infty)$, the strong solution X_t^\star of the nominal (known) process (4b) exists and satisfies*

$$\|X_t^\star\|_{L_{2p^\star}} \leq \Delta_\star, \quad \forall t \in [0, T],$$

where $p^\star \in \mathbb{N}_{\geq 1}$ is stated in Assumption 1.

Remark 2.4 *The well-posedness of (4b) is straightforward to establish under general conditions, see e.g. [60, Definition 5.2.1], [14, Sec. 4.5], and [15, Sec. 5.2].*

Next, we state the assumptions that we place on the true (uncertain) system in (4a).

Assumption 4 (True (Uncertain) System) *The unknown functions Λ_μ and Λ_σ , presented in Definition 1, satisfy*

$$\|\Lambda_\mu(t, a)\|^2 \leq \Delta_\mu^2 (1 + \|a\|^2), \quad \|\Lambda_\sigma(t, a)\|_F^2 \leq \Delta_\sigma^2 (1 + \|a\|^2)^{\frac{1}{2}}, \quad \forall (t, a) \in \mathbb{R}_{\geq 0} \times \mathbb{R}^n,$$

where $\Delta_\mu, \Delta_\sigma \in \mathbb{R}_{>0}$ are known. Furthermore, define $\Lambda_\mu^\parallel : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\Lambda_\mu^\perp : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}^{n-m}$, $\Lambda_\sigma^\parallel : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}^{m \times d}$, and $\Lambda_\sigma^\perp : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}^{n-m \times d}$ as

$$\begin{bmatrix} \Lambda_\mu^\parallel(t, a) \\ \Lambda_\mu^\perp(t, a) \end{bmatrix} = [g(t) \quad g(t)^\perp]^{-1} \Lambda_\mu(t, a), \quad \begin{bmatrix} \Lambda_\sigma^\parallel(t, a) \\ \Lambda_\sigma^\perp(t, a) \end{bmatrix} = [g(t) \quad g(t)^\perp]^{-1} \Lambda_\sigma(t, a), \quad \forall (t, a) \in \mathbb{R}_{\geq 0} \times \mathbb{R}^n, \quad (7)$$

where the input operator g is presented in Definition 1, and g^\perp is defined in Assumption 1. As a consequence of the growth bounds, the functions above satisfy

$$\|\Lambda_\mu^{\{\parallel, \perp\}}(t, a)\|^2 \leq (\Delta_\mu^{\{\parallel, \perp\}})^2 (1 + \|a\|^2), \quad \|\Lambda_\sigma^{\{\parallel, \perp\}}(t, a)\|_F^2 \leq (\Delta_\sigma^{\{\parallel, \perp\}})^2 (1 + \|a\|^2)^{\frac{1}{2}}, \quad \forall (t, a) \in \mathbb{R}_{\geq 0} \times \mathbb{R}^n,$$

for known $\Delta_\mu^\parallel, \Delta_\mu^\perp, \Delta_\sigma^\parallel, \Delta_\sigma^\perp \in \mathbb{R}_{>0}$.

Remark 2.5 *The growth bound on drift uncertainty Λ_μ is the standard general linear growth condition. For diffusion uncertainty Λ_σ , our analysis can compensate for uncertainties growing sub-linearly; we do not constrain Λ_σ to be uniformly bounded. An implication of this is that the stability of the nominal (known) system in (4b) cannot be extended to the uncertain (true) system in (4a) via robustness arguments as done in [55, 54]. The \mathcal{L}_1 -DRAC control can, however, accommodate unbounded drift uncertainties satisfying the sub-linear growth condition.*

Similar to the decomposition of Λ_μ and Λ_σ in the last assumption, we decompose the known diffusion term p as presented below.

Assumption 5 *For the known diffusion term p in Definition 1, we define $p^\parallel : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}^{m \times d}$, and $p^\perp : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}^{n-m \times d}$ as*

$$\begin{bmatrix} p^\parallel(t, a) \\ p^\perp(t, a) \end{bmatrix} = [g(t) \quad g(t)^\perp]^{-1} p(t, a), \quad \forall (t, a) \in \mathbb{R}_{\geq 0} \times \mathbb{R}^n, \quad (8)$$

where, we assume that p^\parallel and p^\perp are Lipschitz continuous, locally in $a \in \mathbb{R}^n$ and uniformly in $t \in \mathbb{R}_{\geq 0}$.

Due to the uniform boundedness and Hölder continuity of p in Assumption 1, we assume that there exist known $\Delta_p^\parallel, \Delta_p^\perp, L_p^\parallel, \hat{L}_p^\parallel, L_p^\perp, \hat{L}_p^\perp \in \mathbb{R}_{>0}$ such that

$$\begin{aligned} \|p^\parallel(t, a)\|_F &\leq \Delta_p^\parallel, \quad \|p^\perp(t, a)\|_F \leq \Delta_p^\perp, \quad \|p^\parallel(t, a) - p^\parallel(t', a')\|_F \leq \hat{L}_p^\parallel |t - t'| + L_p^\parallel \|a - a'\|^{\frac{1}{2}}, \\ \|p^\perp(t, a) - p^\perp(t', a')\|_F &\leq \hat{L}_p^\perp |t - t'| + L_p^\perp \|a - a'\|^{\frac{1}{2}}, \end{aligned}$$

for all $(t, t', a, a') \in \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}^n \times \mathbb{R}^n$.

For notational simplicity, and in light of Assumptions 4-5, we define the following:

Definition 3 We define functions $F_\sigma^\parallel : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}^{m \times d}$ and $F_\sigma^\perp : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}^{n-m \times d}$ as follows:

$$F_\sigma^\parallel(t, a) \doteq p^\parallel(t, a) + \Lambda_\sigma^\parallel(t, a), \quad F_\sigma^\perp(t, a) \doteq p^\perp(t, a) + \Lambda_\sigma^\perp(t, a), \quad \forall (t, a) \in \mathbb{R}_{\geq 0} \times \mathbb{R}^n.$$

The next assumption characterizes the continuity of the uncertain vector fields. Note that these are standard for well-posedness, and can be relaxed for local results.

Assumption 6 The drift uncertainty $\Lambda_\mu : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, and its decomposed components $\Lambda_\mu^\parallel : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $\Lambda_\mu^\perp : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}^{n-m}$, are globally Lipschitz continuous in their arguments, and there exist known positive scalars $L_\mu, \hat{L}_\mu, L_\mu^\parallel, \hat{L}_\mu^\parallel, L_\mu^\perp, \hat{L}_\mu^\perp$ such that

$$\|\Lambda_\mu^{\{\cdot, \parallel, \perp\}}(t, a) - \Lambda_\mu^{\{\cdot, \parallel, \perp\}}(t', a')\| \leq \hat{L}_\mu^{\{\cdot, \parallel, \perp\}} |t - t'| + L_\mu^{\{\cdot, \parallel, \perp\}} \|a - a'\|,$$

for all $(t, t', a, a') \in \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}^n \times \mathbb{R}^n$.

The diffusion uncertainty $\Lambda_\sigma : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times d}$, and its decomposed components $\Lambda_\sigma^\parallel : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}^{m \times d}$ and $\Lambda_\sigma^\perp : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}^{(n-m) \times d}$, are globally Lipschitz continuous in their temporal variable, and Hölder continuous with exponent 1/2 in their state-dependent arguments, and there exist known positive scalars $L_\sigma, \hat{L}_\sigma, L_\sigma^\parallel, \hat{L}_\sigma^\parallel, L_\sigma^\perp, \hat{L}_\sigma^\perp$ such that

$$\|\Lambda_\sigma^{\{\cdot, \parallel, \perp\}}(t, a) - \Lambda_\sigma^{\{\cdot, \parallel, \perp\}}(t', a')\|_F \leq \hat{L}_\sigma^{\{\cdot, \parallel, \perp\}} |t - t'| + L_\sigma^{\{\cdot, \parallel, \perp\}} \|a - a'\|^{\frac{1}{2}},$$

for all $(t, t', a, a') \in \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}^n \times \mathbb{R}^n$.

Remark 2.6 The global Lipschitz continuity in Assumption 6 implies the linear growth of the drift uncertainty Λ_μ in its state-dependent argument, as stated in Assumption 4. However, the diffusion uncertainty Λ_σ is assumed to satisfy the stronger condition of Hölder continuity with exponent 1/2 in its state-dependent argument, which is a sufficient condition for sub-linear growth, as stated in Assumption 4.

The following assumption is regarding the stability of the deterministic counterpart of the true (uncertain) system in (4a) in response to the unmatched drift uncertainty.

Assumption 7 The constant $\lambda \in \mathbb{R}_{>0}$ in Assumption 2 satisfies

$$\lambda > \Delta_g^\perp \cdot \max\{\Delta_\mu^\perp, L_\mu^\perp\},$$

where the bounds $\Delta_g^\perp, \Delta_\mu^\perp, L_\mu^\perp \in \mathbb{R}_{>0}$ are presented in Assumptions 1, 4, and 6, respectively.

Remark 2.7 The condition in Assumption 7 ensures that the true (uncertain) system in (4a) in the absence of diffusion terms ($F_\sigma \equiv 0$) and the matched drift uncertainty $\Lambda_\mu^\parallel \equiv 0$, i.e., deterministic counterpart of the true (uncertain) system with only the unmatched uncertainty Λ_μ^\perp , maintains its incremental exponential stability. The assumption is required since, by definition, the control action cannot “reach” the unmatched uncertainties. This assumption is similar to Assumption 2 in that it pertains only to the deterministic counterparts of the Itô SDEs that we consider.

Lastly, we can obtain stronger results if instead of the local Lipschitz continuity of the known drift vector field in Assumption 1, we assume the global Lipschitz continuity as below.

Assumption 8 (Stronger Regularity of Known Drift Vector Field) There exist known $L_f, \hat{L}_f \in \mathbb{R}_{>0}$ such that the known drift $f(t, a)$ in Definition 1 satisfy:

$$\|f(t, a) - f(t', a')\| \leq \hat{L}_f |t - t'| + L_f \|a - a'\|, \quad \forall (t, t', a, a') \in \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}^n \times \mathbb{R}^n.$$

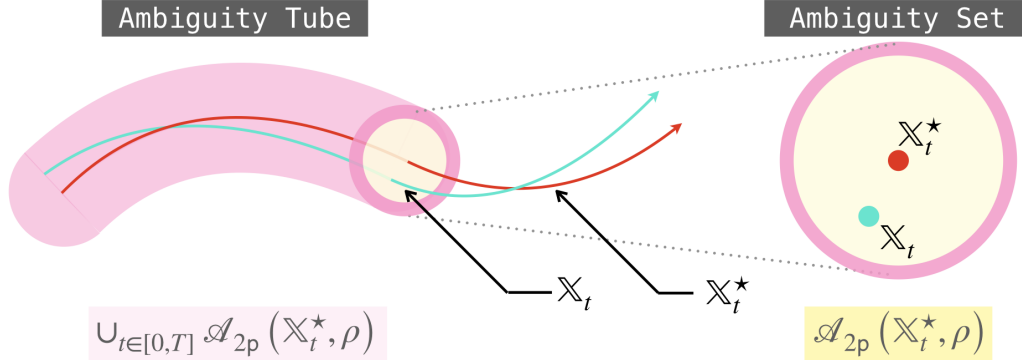


Figure 3: The uniform ambiguity tube (left) and ambiguity set (right) of laws defined in (11) as sets of radius ρ around the nominal (known) distribution \mathbb{X}_t^* . The \mathcal{L}_1 -DRAC feedback is designed to ensure that the true (uncertain) distribution \mathbb{X}_t remains inside the ambiguity set for all $t \in [0, T]$.

2.3 Problem Statement

Consider the **closed-loop true (uncertain) Itô SDE**, for any $T \in (0, \infty)$:

$$dX_t = F_\mu(t, X_t, U_{\mathcal{L}_1, t}) dt + F_\sigma(t, X_t) dW_t, \quad X_0 = x_0 \sim \xi_0, \forall t \in [0, T], \quad (9)$$

that is induced by closing the feedback loop on the true (uncertain) Itô SDE in (4a) with the **feedback process** $U_t = U_{\mathcal{L}_1, t}$, defined as

$$U_{\mathcal{L}_1, t} \doteq \mathcal{F}_{\mathcal{L}_1}(X)(t), \quad \mathcal{F}_{\mathcal{L}_1} : \mathcal{C}([0, T]; \mathbb{R}^n) \rightarrow \mathcal{C}([0, T]; \mathbb{R}^m), \quad \forall t \in [0, T], \quad (10)$$

where $\mathcal{F}_{\mathcal{L}_1}$ is the \mathcal{L}_1 -**DRAC feedback operator**. We wish to synthesize the \mathcal{L}_1 -DRAC feedback operator $\mathcal{F}_{\mathcal{L}_1}$ such that the following conditions are satisfied:

Condition 1: The true uncertain process $X_t \sim \mathbb{X}_t$ exists and is a unique strong solution of the closed-loop true (uncertain) Itô SDE in (9) for any $T \in (0, \infty)$.

Condition 2: There exists an a priori known $\rho \in \mathbb{R}_{>0}$ such that

$$\mathbb{X}_t \in \mathcal{A}_{2p}(\mathbb{X}_t^*, \rho) \doteq \{\text{Probability measures } \nu \text{ on } \mathcal{B}(\mathbb{R}^n) \mid W_{2p}(\mathbb{X}_t^*, \nu) < \rho\}, \quad \forall t \in [0, T], \quad (11)$$

where \mathbb{X}_t^* is the law of the nominal (known) processes (4b), and $p \in \{1, \dots, p^*\}$ for $p^* \in \mathbb{N}_{\geq 1}$ introduced in Assumptions 1 and 3. We refer to $\mathcal{A}_{2p}(\mathbb{X}_t^*, \rho)$ as the **uniform ambiguity set of laws**. Similarly, we refer to $\cup_{t \in [0, T]} \mathcal{A}_{2p}(\mathbb{X}_t^*, \rho)$ as the **uniform ambiguity tube of laws**. The uniform ambiguity set and the uniform ambiguity tube are illustrated in Fig. 3.

Condition 3: Furthermore, there exists an a priori known $\mathbb{R}_{>0} \ni \hat{\rho}(\hat{t}) < \rho$, $\hat{t} \in (0, T)$, such that

$$\mathbb{X}_t \in \mathcal{A}_{2p}(\mathbb{X}_t^*, \hat{\rho}(\hat{t})) \subset \mathcal{A}_{2p}(\mathbb{X}_t^*, \rho), \quad \forall t \in [\hat{t}, T]. \quad (12)$$

Analogous to the previous nomenclature, we refer to $\mathcal{A}_{2p}(\mathbb{X}_t^*, \hat{\rho}(\hat{t}))$ and $\cup_{t \in [0, T]} \mathcal{A}_{2p}(\mathbb{X}_t^*, \hat{\rho}(\hat{t}))$ as the **uniform ultimate ambiguity set of laws** and **uniform ultimate ambiguity tube of laws**, respectively.

Remark 2.8 The ambiguity tube (union of ambiguity sets of laws over $\mathcal{B}(\mathbb{R}^n)$) is not the same as ambiguity set of laws on $\mathcal{B}(\mathcal{C}([0, T]; \mathbb{R}^n))$. The ambiguity sets we consider are defined using the Wasserstein metric on the Polish space \mathbb{R}^n under the Euclidean metric. Laws on the trajectories (paths) are instead defined on the Polish space $\mathcal{C}([0, T]; \mathbb{R}^n)$ equipped with the sup-metric or some other appropriately chosen metrics [92, Sec. 18.1]. We have previously discussed the difference between the underlying spaces for the laws in Remark 2.1.

3 \mathcal{L}_1 -DRAC Control: Design

In this section we define and analyze the closed loop true (uncertain) system in (9) with the \mathcal{L}_1 -DRAC feedback control. We begin with the definition of \mathcal{L}_1 -DRAC feedback operator $\mathcal{F}_{\mathcal{L}_1}$ that we introduced in (10).

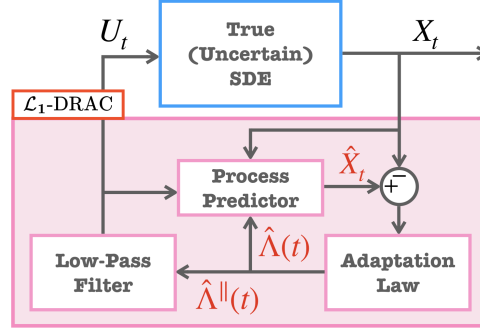


Figure 4: The architecture of the \mathcal{L}_1 -DRAC controller. The controller has three components: a *process predictor* with output \hat{X}_t , an *adaptation law*, and a *low pass filter*.

3.1 Architecture and Definition

The design of $\mathcal{F}_{\mathcal{L}_1}$ is based on the \mathcal{L}_1 -adaptive control methodology [30]. We will design the \mathcal{L}_1 -DRAC feedback operator $\mathcal{F}_{\mathcal{L}_1}$ such that the closed-loop true (uncertain) Itô SDE in (9) satisfies the conditions we set forth in Sec. 2.3. Following the architecture of \mathcal{L}_1 -adaptive control [30], the \mathcal{L}_1 -DRAC feedback operator $\mathcal{F}_{\mathcal{L}_1}$ consists of a process predictor, an adaptation law, and a low-pass filter as illustrated in Fig. 4.

We define the \mathcal{L}_1 -DRAC feedback operator $\mathcal{F}_{\mathcal{L}_1} : \mathcal{C}([0, T] : \mathbb{R}^n) \rightarrow \mathcal{C}([0, T] : \mathbb{R}^n)$ as follows:

$$\mathcal{F}_{\mathcal{L}_1}(y) \doteq \mathcal{F}_{\omega} \circ \mathcal{F}_{T_s} \circ \mathcal{F}_{\lambda_s}(y), \quad y \in \mathcal{C}([0, T] : \mathbb{R}^n), \quad (13)$$

which we can alternatively represent as

$$\mathcal{F}_{\mathcal{L}_1}(y) = \mathcal{F}_{\omega}(\hat{\Lambda}^{\parallel}), \quad \hat{\Lambda}^{\parallel} = \mathcal{F}_{T_s}^{\parallel}(\hat{\Lambda}), \quad \hat{\Lambda} = \mathcal{F}_{T_s}(\hat{y}, y), \quad \hat{y} = \mathcal{F}_{\lambda_s}(y), \quad (14)$$

where

$$\mathcal{F}_{\omega}(\hat{\Lambda}^{\parallel})(t) \doteq -\omega \int_0^t e^{-\omega(t-\nu)} \hat{\Lambda}^{\parallel}(\nu) d\nu, \quad (\text{Low-pass filter}) \quad (15a)$$

$$\begin{aligned} \hat{\Lambda}^{\parallel}(t) &= \mathcal{F}_{T_s}^{\parallel}(\hat{\Lambda})(t) = \sum_{i=0}^{\lfloor \frac{t}{T_s} \rfloor} \Theta_{ad}(iT_s) \hat{\Lambda}(t) \mathbb{1}_{[iT_s, (i+1)T_s)}(t), \\ \hat{\Lambda}(t) &= \mathcal{F}_{T_s}(\hat{y}, y)(t) = 0_n \mathbb{1}_{[0, T_s)}(t) \\ &\quad + \lambda_s (1 - e^{\lambda_s T_s})^{-1} \sum_{i=1}^{\lfloor \frac{t}{T_s} \rfloor} (\hat{y}(iT_s) - y(iT_s)) \mathbb{1}_{[iT_s, (i+1)T_s)}(t), \end{aligned} \quad (\text{Adaptation Law}) \quad (15b)$$

$\hat{y}(t) = \mathcal{F}_{\lambda_s}(y)(t) \Rightarrow$ solution to the integral equation:

$$\hat{y}(t) = x_0 + \int_0^t \left(-\lambda_s \mathbb{1}_n (\hat{y}(\nu) - y(\nu)) + f(\nu, y(\nu)) + g(\nu) \mathcal{F}_{\mathcal{L}_1}(y)(\nu) + \hat{\Lambda}(\nu) \right) d\nu, \quad (\text{Process Predictor}) \quad (15c)$$

for $t \in [0, T]$, where $\omega, T_s, \lambda_s \in \mathbb{R}_{>0}$ are the control parameters. The parameters ω and T_s are referred to as the **filter bandwidth** and the **sampling period**, respectively. Additionally, $\Theta_{ad}(t) = [\mathbb{1}_m \quad 0_{m, n-m}] \bar{g}(t)^{-1} \in \mathbb{R}^{m \times n}$, where $\bar{g}(t) = [g(t) \quad g(t)^{\perp}] \in \mathbb{R}^{n \times n}$, and here g^{\perp} is defined in Assumption 4. Finally, $x_0 \sim \xi_0$ in (15c) is the initial condition of the true (uncertain) process (4a). By initializing the process predictor with the initial condition of the true (uncertain) process, we simplify the analysis, however, this is not a strict requirement. See [103] for a more general treatment of the process predictor initialization for deterministic systems.

3.2 \mathcal{L}_1 -DRAC Parameters

We now present the choice of the control parameters $\omega \in \mathbb{R}_{>0}$ and $T_s \in \mathbb{R}_{>0}$, the bandwidth for the low-pass filter in (15a) and the sampling period for the adaptation law in (15b), respectively. Suppose that the assumptions in Sec. 2.2 hold. For an arbitrarily chosen $\epsilon_r, \epsilon_a \in \mathbb{R}_{>0}$, and any $p \in \{1, \dots, p^*\}$, select $\mathbf{q}_r(p), \mathbf{q}_a(p) \in \mathbb{R}_{>0}$ such that

$$\left(1 - \frac{\Delta_g^{\perp} \Delta_{\mu}^{\perp}}{\lambda} \right) \mathbf{q}_r(p)^2 \geq \alpha (X_0^r, X_0^*)_{2p}^2 + \Gamma_r(\mathbf{q}_r(p), p, \omega) + \epsilon_r, \quad (16a)$$

$$\left(1 - \frac{\Delta_g^\perp L_\mu^\perp}{\lambda}\right) \varrho_a(\mathbf{p})^2 \geq \Gamma_a(\varrho_a(\mathbf{p}), \mathbf{p}, \omega) + \epsilon_a, \quad (16b)$$

and define

$$\varrho(\mathbf{p}) = \varrho_r(\mathbf{p}) + \varrho_a(\mathbf{p}), \quad (17)$$

where

$$\alpha(X_0^r, X_0^*)_{2p} \in \left\{ \|X_0^r - X_0^*\|_{L_{2p}}, W_{2p}(\xi_0, \xi_0^*) \right\}, \quad (18a)$$

$$\begin{aligned} \Gamma_r(\varrho_r(\mathbf{p}), \mathbf{p}, \omega) &= \frac{\Delta_{\odot_1}^r}{2\lambda} + \frac{\omega \Delta_{\odot_4}^r(\mathbf{p})}{|2\lambda - \omega|} + \left(\frac{\Delta_{\odot_1}^r}{2\lambda} + \frac{\omega \Delta_{\odot_4}^r(\mathbf{p})}{|2\lambda - \omega|} \right) (\varrho_r(\mathbf{p}) + \Delta_\star)^{\frac{1}{2}} \\ &\quad + \left(\frac{\Delta_{\odot_1}^r}{2\lambda} + \frac{\omega \Delta_{\odot_8}^r(\mathbf{p})}{|2\lambda - \omega|} \right) (\varrho_r(\mathbf{p}) + \Delta_\star) \\ &\quad + \left(\frac{\Delta_{\odot_2}^r}{2\lambda} + \frac{\Delta_{\odot_3}^r(\mathbf{p})}{2\sqrt{\lambda}} + \frac{\Delta_{\odot_1}^r}{2\lambda} \Delta_\star \right) \varrho_r(\mathbf{p}) + \frac{\Delta_{\odot_1}^r(\mathbf{p})}{2\sqrt{\lambda}} (\varrho_r(\mathbf{p}) + \Delta_\star)^{\frac{3}{2}}, \end{aligned} \quad (18b)$$

$$\Gamma_a(\varrho_a(\mathbf{p}), \mathbf{p}, \omega) = \left(\frac{\Delta_{\odot_1}^r}{2\lambda} + \frac{\omega \Delta_{\odot_4}^r(\mathbf{p})}{|2\lambda - \omega|} + \frac{\Delta_{\odot_1}^r(\mathbf{p})}{2\sqrt{\lambda}} (\varrho_a(\mathbf{p}))^{\frac{1}{2}} \right) \varrho_a(\mathbf{p}), \quad (18c)$$

and where the definitions of the constants $\Delta_i^r(\mathbf{p})$, $\Delta_i(\mathbf{p})$, $i \in \{\odot, \ominus, \otimes, \ast\}$, are provided in Appendix A.

Remark 3.1 (Existence of $\varrho_r(\mathbf{p})$ and $\varrho_a(\mathbf{p})$) The condition in Assumption 7 implies $\Delta_g^\perp \Delta_\mu^\perp / \lambda \in (0, 1)$. It is also straightforward to see that

$$\Gamma_r(\varrho_r(\mathbf{p}), \mathbf{p}, \omega) \in \mathcal{O}(\varrho_r(\mathbf{p})^{\frac{3}{2}}), \text{ as } \varrho_r(\mathbf{p}) \rightarrow \infty, \text{ and } \Gamma_r(\varrho_r(\mathbf{p}), \mathbf{p}, \omega) \in \mathcal{O}(1), \text{ as } \omega \rightarrow \infty.$$

It then follows

$$\left(1 - \frac{\Delta_g^\perp \Delta_\mu^\perp}{\lambda}\right) \varrho_r(\mathbf{p})^2 - \Gamma_r(\varrho_r(\mathbf{p}), \mathbf{p}, \omega) \in \mathcal{O}(\varrho_r(\mathbf{p})^2), \text{ as } \varrho_r(\mathbf{p}) \rightarrow \infty,$$

and

$$\left(1 - \frac{\Delta_g^\perp \Delta_\mu^\perp}{\lambda}\right) \varrho_r(\mathbf{p})^2 - \Gamma_r(\varrho_r(\mathbf{p}), \mathbf{p}, \omega) \in \mathcal{O}(1), \text{ as } \omega \rightarrow \infty. \quad (19)$$

Therefore, one can always choose a $\varrho_r(\mathbf{p}) \in \mathbb{R}_{>0}$ that satisfies (16a). Importantly, the choice of $\varrho_r(\mathbf{p}) \in \mathbb{R}_{>0}$ remains **independent** of the value of $\omega \in \mathbb{R}_{>0}$. Similar arguments apply for the existence of $\varrho_a(\mathbf{p})$ satisfying (16b).

Since $\Delta_g^\perp \Delta_\mu^\perp / \lambda \in (0, 1)$ and $\Gamma_r(\varrho_r(\mathbf{p}), \mathbf{p}, \omega) \geq 0$, $\forall \{\varrho_r(\mathbf{p}), \mathbf{p}, \omega\} \in \mathbb{R}_{>0} \times \{1, \dots, \mathbf{p}^*\} \times \mathbb{R}_{>0}$, a consequence of (16a) is that

$$\varrho_r(\mathbf{p}) > \alpha(X_0^r, X_0^*)_{2p}, \quad \alpha(X_0^r, X_0^*)_{2p} \in \left\{ \|X_0^r - X_0^*\|_{L_{2p}}, W_{2p}(\xi_0, \xi_0^*) \right\}. \quad (20)$$

3.2.1 Filter Bandwidth

Given $\varrho_r(\mathbf{p})$ and $\varrho_a(\mathbf{p})$ that satisfy (16), the bandwidth $\omega \in \mathbb{R}_{>0}$ of the filter \mathcal{F}_ω in (15a) is chosen so as to verify the following conditions:

$$\frac{1}{|2\lambda - \omega|} \Theta_r(\varrho_r(\mathbf{p}), \mathbf{p}, \omega) < \left(1 - \frac{\Delta_g^\perp \Delta_\mu^\perp}{\lambda}\right) \varrho_r(\mathbf{p})^2 - \alpha(X_0^r, X_0^*)_{2p}^2 - \Gamma_r(\varrho_r(\mathbf{p}), \mathbf{p}, \omega), \quad (21a)$$

$$\frac{1}{|2\lambda - \omega|} \Theta_a(\varrho_a(\mathbf{p}), \varrho_r(\mathbf{p}), \mathbf{p}, \omega) < \left(1 - \frac{\Delta_g^\perp L_\mu^\perp}{\lambda}\right) \varrho_a(\mathbf{p})^2 - \Gamma_a(\varrho_a(\mathbf{p}), \mathbf{p}, \omega), \quad (21b)$$

where

$$\begin{aligned} \Theta_r(\varrho_r(\mathbf{p}), \mathbf{p}, \omega) &= \Delta_{\odot_2}^r(\mathbf{p}) + \sqrt{\omega} \Delta_{\odot_3}^r(\mathbf{p}) + (\Delta_{\odot_2}^r(\mathbf{p}) + \sqrt{\omega} \Delta_{\odot_3}^r(\mathbf{p})) (\varrho_r(\mathbf{p}) + \Delta_\star)^{\frac{1}{2}} \\ &\quad + (\Delta_{\odot_4}^r(\mathbf{p}) + \sqrt{\omega} \Delta_{\odot_6}^r(\mathbf{p})) (\varrho_r(\mathbf{p}) + \Delta_\star) + (\Delta_{\odot_5}^r(\mathbf{p}) + \sqrt{\omega} \Delta_{\odot_7}^r(\mathbf{p})) \varrho_r(\mathbf{p}) \\ &\quad + [(\Delta_{\odot_2}^r(\mathbf{p}) + \sqrt{\omega} \Delta_{\odot_4}^r(\mathbf{p})) (\varrho_r(\mathbf{p}) + \Delta_\star) + (\Delta_{\odot_3}^r(\mathbf{p}) + \sqrt{\omega} \Delta_{\odot_5}^r(\mathbf{p})) \varrho_r(\mathbf{p})] (\varrho_r(\mathbf{p}) + \Delta_\star)^{\frac{1}{2}} \end{aligned}$$

$$+ [\Delta_{\otimes_2}^r(\mathbf{p})(\mathbf{q}_r(\mathbf{p}) + \Delta_\star) + \Delta_{\otimes_3}^r(\mathbf{p})\mathbf{q}_r(\mathbf{p})](\mathbf{q}_r(\mathbf{p}) + \Delta_\star), \quad (22)$$

and

$$\begin{aligned} \Theta_a(\mathbf{q}_a(\mathbf{p}), \mathbf{q}_r(\mathbf{p}), \mathbf{p}, \omega) = & \left[\sqrt{\omega} \Delta_{\otimes_1}(\mathbf{p}) + \sqrt{\omega} \Delta_{\otimes_3}(\mathbf{p}) \mathbf{q}''(\mathbf{p}) + (\Delta_{\otimes_2}(\mathbf{p}) + \sqrt{\omega} \Delta_{\otimes_4}(\mathbf{p})) \mathbf{q}_a(\mathbf{p}) \right] (\mathbf{q}_a(\mathbf{p}))^{\frac{1}{2}} \\ & + [\Delta_{\otimes_2}(\mathbf{p}) + \sqrt{\omega} \Delta_{\otimes_3}(\mathbf{p}) + \Delta_{\otimes_2} \mathbf{q}''(\mathbf{p}) + \Delta_{\otimes_3} \mathbf{q}_a(\mathbf{p})] \mathbf{q}_a(\mathbf{p}), \end{aligned} \quad (23)$$

with $\mathbf{q}''(\mathbf{p}) = \mathbf{q}_a(\mathbf{p}) + 2(\mathbf{q}_r(\mathbf{p}) + \Delta_\star)$.

Remark 3.2 (Feasibility of (21)) *It is easily shown that*

$$\left\{ \frac{1}{|2\lambda - \omega|} \Theta_r(\mathbf{q}_r(\mathbf{p}), \mathbf{p}, \omega), \frac{1}{|2\lambda - \omega|} \Theta_a(\mathbf{q}_a(\mathbf{p}), \mathbf{q}_r(\mathbf{p}), \mathbf{p}, \omega) \right\} \in \mathcal{O}\left(\frac{1}{\sqrt{\omega}}\right), \text{ as } \omega \rightarrow \infty.$$

Moreover, (16a) and (19) imply that

$$\left(1 - \frac{\Delta_g^\perp \Delta_\mu^\perp}{\lambda}\right) \mathbf{q}_r(\mathbf{p})^2 - \alpha(X_0^r, X_0^\star)_{2\mathbf{p}}^2 - \Gamma_r(\mathbf{q}_r(\mathbf{p}), \mathbf{p}, \omega) \geq \epsilon_r > 0,$$

independent of $\omega \in \mathbb{R}_{>0}$, where $\alpha(X_0^r, X_0^\star)_{2\mathbf{p}} \in \left\{ \|X_0^r - X_0^\star\|_{L_{2\mathbf{p}}}, W_{2\mathbf{p}}(\xi_0, \xi_0^\star) \right\}$. As a consequence, one can always choose a filter bandwidth such that

$$0 < \frac{1}{|2\lambda - \omega|} \Theta_r(\mathbf{q}_r(\mathbf{p}), \mathbf{p}, \omega) < \left(1 - \frac{\Delta_g^\perp \Delta_\mu^\perp}{\lambda}\right) \mathbf{q}_r(\mathbf{p})^2 - \alpha(X_0^r, X_0^\star)_{2\mathbf{p}}^2 - \Gamma_r(\mathbf{q}_r(\mathbf{p}), \mathbf{p}, \omega),$$

thus proving the feasibility of (21a). The feasibility of (21b) follows similarly.

3.2.2 Sampling Period

Given $\mathbf{q}_r(\mathbf{p})$ and $\mathbf{q}_a(\mathbf{p})$ that satisfy (16), and filter bandwidth ω that satisfies (21), the sampling period $T_s \in \mathbb{R}_{>0}$ of the adaptation law \mathcal{F}_{T_s} in (15b) is chosen such that

$$\begin{aligned} \sum_{i=1}^3 \Upsilon_{a_i}(\mathbf{q}'(\mathbf{p}), \mathbf{q}_a(\mathbf{p}); \mathbf{p}, \omega, T_s) \leq & \left(1 - \frac{\Delta_g^\perp L_\mu^\perp}{\lambda}\right) \mathbf{q}_a(\mathbf{p})^2 - \Gamma_a(\mathbf{q}_a(\mathbf{p}), \mathbf{p}, \omega) \\ & - \frac{1}{|2\lambda - \omega|} \Theta_a(\mathbf{q}_a(\mathbf{p}), \mathbf{q}_r(\mathbf{p}), \mathbf{p}, \omega), \end{aligned} \quad (24)$$

where $\mathbf{q}'(\mathbf{p}) = \mathbf{q}_a(\mathbf{p}) + \mathbf{q}_r(\mathbf{p}) + \Delta_\star$, and we have defined the following for any $\xi_1, \xi_2 \in \mathbb{R}_{\geq 0}$:

$$\Upsilon_{a_1}(\xi_1, \xi_2; \mathbf{p}, \omega, T_s) = \tilde{\Upsilon}^-(\xi_1; \mathbf{p}, \omega, T_s) \xi_2 + \frac{\omega}{\lambda |2\lambda - \omega|} \hat{\Upsilon}(\xi_2; \mathbf{p}, \omega) \Upsilon^-(\xi_1; \mathbf{p}, \omega, T_s), \quad (25a)$$

$$\Upsilon_{a_2}(\xi_1, \xi_2; \mathbf{p}, \omega, T_s) = \frac{1}{\lambda} \left(\Delta_g \xi_2 + \frac{\sqrt{\omega}}{|2\lambda - \omega|} \hat{\Upsilon}(\xi_2; \mathbf{p}, \omega) \right) (\Upsilon'_1(\xi_1; \mathbf{p}, \omega, T_s) + \sqrt{\omega} \Upsilon'_2(\xi_1; \mathbf{p}, \omega, T_s)), \quad (25b)$$

$$\Upsilon_{a_3}(\xi_1, \xi_2; \mathbf{p}, \omega, T_s) = \frac{1}{|2\lambda - \omega|} \left(2\Delta_g \xi_2 + \frac{\sqrt{\omega}}{\lambda} \hat{\Upsilon}(\xi_2; \mathbf{p}, \omega) \right) \Upsilon'_3(\xi_1; \mathbf{p}, \omega, T_s), \quad (25c)$$

and where the maps $\Upsilon'_{\{1,2,3\}}$, $\tilde{\Upsilon}^-$, Υ^- , and $\hat{\Upsilon}$ are defined in (A.18) and (A.19) in Section A.2.1.

Remark 3.3 (Feasibility of (24)) *As a consequence of (A.21) in Proposition A.1, one sees that*

$$\lim_{T_s \downarrow 0} \sum_{i=1}^3 \Upsilon_{a_i}(\mathbf{q}'(\mathbf{p}), \mathbf{q}_a(\mathbf{p}); \mathbf{p}, \omega, T_s) = 0,$$

for any chosen $(\mathbf{q}_r(\mathbf{p}), \mathbf{q}_a(\mathbf{p}), \omega, \mathbf{p}) \in \mathbb{R}_{>0} \times \mathbb{R}_{>0} \times \mathbb{R}_{>0} \times \mathbb{N}_{\geq 1}$. Furthermore, (16b) and (21b) imply that

$$\left(1 - \frac{\Delta_g^\perp L_\mu^\perp}{\lambda}\right) \mathbf{q}_a(\mathbf{p})^2 - \Gamma_a(\mathbf{q}_a(\mathbf{p}), \mathbf{p}, \omega) - \frac{1}{|2\lambda - \omega|} \Theta_a(\mathbf{q}_a(\mathbf{p}), \mathbf{q}_r(\mathbf{p}), \mathbf{p}, \omega) > 0.$$

It then follows from the above two facts that one can always choose a sampling period $T_s \in \mathbb{R}_{>0}$ that renders (24) true.

4 \mathcal{L}_1 -DRAC Control: Analysis

We now present the analysis to obtain the performance bounds for the \mathcal{L}_1 -DRAC true (uncertain) Itô SDE in (9). We follow a two-step approach as used for deterministic \mathcal{L}_1 adaptive control analysis, see e.g. [30, 104]. First, in Sec. 4.1 we introduce an intermediate and non-realizable process that we call the *reference process* and obtain the performance bounds between it and the nominal (known) process. Then, in Sec. 4.2 we analyze the performance of the \mathcal{L}_1 -DRAC true (uncertain) process relative to the reference process. The last step leads us to the relative performance bounds between the true (uncertain) process and the nominal (known) process via the triangle inequality of the Wasserstein metric.

4.1 Performance Analysis: Reference Process

We begin with the definition of the reference process.

Definition 4 (Reference Process) We say that X_t^r , $t \in [0, T]$, for any $T \in (0, \infty)$, is the **reference process**, if X_t^r is a unique strong solution to the following **reference Itô SDE**:

$$dX_t^r = F_\mu(t, X_t^r, U_t^r) dt + F_\sigma(t, X_t^r) dW_t, \quad X_0^r = x_0 \sim \xi_0, \quad \forall t \in [0, T], \quad (26)$$

where the initial condition $x_0 \sim \xi_0$ and the driving Brownian motion W_t are identical to those for the true (uncertain) process in (9) (and (4a)). The **reference feedback process** U_t^r is defined via the **reference feedback operator** \mathcal{F}_r as follows:

$$U_t^r = \mathcal{F}_r(X^r)(t), \quad \mathcal{F}_r(X^r) \doteq \mathcal{F}_\omega(\Lambda_\mu^\parallel(\cdot, X^r)) + \mathcal{F}_{\mathcal{N}, \omega}(F_\sigma^\parallel(\cdot, X^r), W), \quad t \in [0, T], \quad (27)$$

where the operator \mathcal{F}_ω is the low-pass filter defined in (15a), the vector field F_σ^\parallel is introduced in Definition 3, and

$$\mathcal{F}_{\mathcal{N}, \omega}(F_\sigma^\parallel(\cdot, X^r), W)(t) \doteq -\omega \int_0^t e^{-\omega(t-\nu)} F_\sigma^\parallel(\nu, X_\nu^r) dW_\nu. \quad (28)$$

Moreover, the functions $\Lambda_\mu^\parallel(t, X_t^r) \in \mathbb{R}^m$ and $\Lambda_\sigma^\parallel(t, X_t^r) \in \mathbb{R}^{m \times d}$ are defined in Assumption 4, and $p^\parallel(t, X_t^r) \in \mathbb{R}^{m \times d}$ is defined in Assumption 5.

Finally, similar to \mathbb{X}_t and \mathbb{X}_t^* in Definition 2, we denote by \mathbb{X}_t^r the **reference law (path/trajectory probability measures)** induced by the process X_t^r on $\mathcal{B}(\mathbb{R}^n)$, i.e., $X_t^r \sim \mathbb{X}_t^r$.

The reference process is obtained by closing the loop of the true (uncertain) process in (4a) with the feedback process U_t^r that is composed of the filtered matched drift uncertainty $\mathcal{F}_\omega(\Lambda_\mu^\parallel(\cdot, X^r))(t)$ and the filtered **totality** of the matched diffusion vector field $\mathcal{F}_{\mathcal{N}, \omega}(p^\parallel(\cdot, X^r) + \Lambda_\sigma^\parallel(\cdot, X^r), W)(t)$ expressed as an Itô integral with respect to the driving true Brownian motion. Thus, the reference process is **non-realizable** since, by definition, we do not have knowledge of the epistemic uncertainties Λ_μ and Λ_σ , and the aleatoric uncertainty W . The reference process represents the **best achievable performance** since it quantifies, as a function of the low-pass filter bandwidth ω , how the system operates under the non-realizable assumption of perfect knowledge of the uncertainties.

Remark 4.1 In addition to the epistemic and aleatoric uncertainties Λ_μ , Λ_σ , and W , the reference feedback process U_t^r also includes the known diffusion term $p(\cdot, X^r)$ in its definition. Therefore, the reference feedback process U_t^r further attempts to remove the effects of the known diffusion term $p(t, X_t^r)$ from the reference process X_t^r . The reason for the inclusion of the known diffusion term p is that due to the state-multiplicative nature of the term $p(t, X_t^r) dW_t$, one cannot in general disambiguate the effects of the known term $p(t, X_t^r)$ from the total uncertainty $\int_0^t \Lambda_\mu(\nu, X_\nu^r) d\nu + \int_0^t (p(\nu, X_\nu^r) + \Lambda_\sigma(\nu, X_\nu^r)) dW_\nu$.

For the performance analysis of the reference process with respect to the nominal (known) process, we define the following:

Definition 5 (Joint Known (Nominal)-Reference Process) We say that $Y_t \in \mathbb{R}^{2n}$, $t \in [0, T]$, for any $T \in (0, \infty)$, is the **joint known(nominal)-reference process**, if it is a unique strong solution of the following **joint known(nominal)-reference Itô SDE** on $(\Omega, \mathcal{F}, \mathbb{W}_{0,t} \times \mathbb{W}_{0,t}^*, \mathbb{P})$ (see Definition 2 for the filtrations):

$$dY_t = G_\mu(t, Y_t) dt + G_\sigma(t, Y_t) d\widehat{W}_t, \quad t \in [0, T], \quad Y_0 = y_0 \sim \xi_0 \times \xi_0^*, \quad Y_t \sim \mathbb{Y}_t, \quad (29)$$

where

$$y_0 \doteq \begin{bmatrix} x_0 \\ x_0^* \end{bmatrix} \in \mathbb{R}^{2n}, \quad Y_t \doteq \begin{bmatrix} X_t^r \\ X_t^* \end{bmatrix} \in \mathbb{R}^{2n}, \quad \widehat{W}_t \doteq \begin{bmatrix} W_t \\ W_t^* \end{bmatrix} \in \mathbb{R}^{2d},$$

$$G_\mu(t, Y_t) \doteq \begin{bmatrix} F_\mu(t, X_t^r, U_t^r) \\ \bar{F}_\mu(t, X_t^*) \end{bmatrix} \in \mathbb{R}^{2n}, \quad G_\sigma(t, Y_t) \doteq \begin{bmatrix} F_\sigma(t, X_t^r) & 0_{n,d} \\ 0_{n,d} & \bar{F}_\sigma(t, X_t^*) \end{bmatrix} \in \mathbb{R}^{2n \times 2d},$$

and where ξ_0 and ξ_0^* are the initial condition distributions and \mathbb{Y}_t denotes the law of (X_t^r, X_t^*) on $\mathcal{B}(\mathbb{R}^{2n})$.

We obtain the desired results using the Khasminskii-type theorem [91, Thm. 3.2]. For this purpose, we require a truncated version of (29) to build local solutions and extend them to the global solution using a limiting procedure.

Definition 6 (Truncated Joint Known (Nominal)-Reference Process) We first define

$$U_N \doteq \{a \in \mathbb{R}^{2n} : \|a\| < N\} \subset \subset \mathbb{R}^{2n}, \quad \forall N \in \mathbb{R}_{>0}, \quad (30)$$

where $\subset \subset \mathbb{R}^{2n}$ denotes compact containment in \mathbb{R}^{2n} [105, Sec. A.2]. Next, we define the **truncated joint known(nominal)-reference Itô SDE** as

$$dY_{N,t} = G_{N,\mu}(t, Y_{N,t}) dt + G_{N,\sigma}(t, Y_{N,t}) d\widehat{W}_t, \quad Y_{N,0} = Y_0, \quad (31)$$

where, the process and the drift and diffusion vector fields are defined as

$$Y_{N,t} \doteq \begin{bmatrix} X_{N,t}^r \\ X_{N,t}^* \end{bmatrix}, \quad G_{N,\mu}(t, a) (G_{N,\sigma}(t, a)) = \begin{cases} G_\mu(t, a) (G_\sigma(t, a)), & \|a\| \leq N \\ 0_{2n} (0_{2n,2d}), & \|a\| \geq 2N \end{cases}, \quad \forall (a, t) \in \mathbb{R}^{2n} \times [0, T],$$

for any $G_{N,\mu}(t, a)$ and $G_{N,\sigma}(t, a)$ that are **uniformly** bounded and Lipschitz continuous for all $a \in \mathbb{R}^{2n}$ and $t \in \mathbb{R}_{\geq 0}$. Similar to Y_t , we refer to $Y_{N,t} \in \mathbb{R}^{2n}$ as the **truncated joint known(nominal)-reference process** if it is a unique strong solution of (31).

An example of explicit construction of functions of the form $G_{N,\{\mu,\sigma\}}$ can be found in [89, p. 191].

With the setup complete, we start the analysis of the reference process by first establishing the existence and uniqueness of strong solutions of the truncated joint process.

Proposition 4.1 (Well-Posedness of (31)) If Assumptions 1 and 4 hold true, then for any $N \in \mathbb{R}_{>0}$, $Y_{N,t}$ is a unique strong solution of (31), $\forall t \in [0, T]$, for any $T \in (0, \infty)$ and is a strong Markov process $\forall t \in \mathbb{R}_{\geq 0}$. Furthermore, define⁴

$$\tau_N \doteq T \wedge \inf \{t \in \mathbb{R}_{\geq 0} : Y_{N,t} \notin U_N\}, \quad (32)$$

where the open and bounded set U_N is defined in (30), for an arbitrary $N \in \mathbb{R}_{>0}$. Then, $Y_{N,t} \in \mathcal{M}_2([0, \tau_N], \mathbb{R}^{2n} | \mathfrak{W}_{0,t} \times \mathfrak{W}_{0,t}^*)$ uniquely solves (29), in the strong sense, for all $t \in [0, \tau_N]$.

Proof. See Appendix C. □

Using the well-posedness of the truncated process, we can now define the transition probability for the process as ([48, Sec. 3.1], [91, Sec. 2.9]):

$$\mathbb{Y}_N(Y_{N,s}, s; B, t) \doteq \mathbb{P}(Y_{N,t} \in B | Y_{N,s}), \quad \forall 0 \leq s \leq t \leq T, \quad B \in \mathcal{B}(\mathbb{R}^{2n}). \quad (33)$$

Next, we derive bounds on the *uniform in time* moments between the truncated versions of the reference and known (nominal) processes.

Lemma 4.1 Let the assumptions in Sec. 2.2 hold. For an arbitrary $N \in \mathbb{R}_{>0}$, let the stopping time τ_N be as in (32) for the truncated joint process in Definition 6. For any constant $t^* \in \mathbb{R}_{>0}$ define⁵

$$\tau^* = t^* \wedge \tau_N, \quad \tau(t) = t \wedge \tau^*. \quad (34)$$

Then, with $\mathcal{V}(Y_N) \doteq \|X_N^r - X_N^*\|^2$, the truncated joint process $Y_{N,t} = (X_{N,t}^r, X_{N,t}^*)$ in Definition 6 satisfies:

$$\begin{aligned} \mathbb{E}_{y_0} [\mathcal{V}(Y_{N,\tau(t)})] &\leq \mathbb{E}_{y_0} [e^{-2\lambda\tau(t)}] \mathcal{V}(y_0) + \mathbb{E}_{y_0} [e^{-2\lambda\tau(t)} \Xi^r(\tau(t), Y_N)] \\ &\quad + \mathbb{E}_{y_0} [e^{-(2\lambda+\omega)\tau(t)} \Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega)], \quad \forall t \in \mathbb{R}_{\geq 0}, \end{aligned} \quad (35)$$

⁴The stopping time τ_N takes on the value of either the finite horizon T , or the first exit time of $Y_{N,t}$ from set U_N .

⁵The stopping time τ^* is either the constant t^* , the finite horizon T , or the first exit time of $Y_{N,t}$ from the set U_N . The variable $\tau(t)$ is identical to t up to τ^* , and zero afterwards.

and

$$\begin{aligned} \|\mathcal{V}(Y_{N,\tau(t)})\|_{L_p}^{y_0} &\leq \left\| e^{-2\lambda\tau(t)} \mathcal{V}(y_0) + \left\| e^{-2\lambda\tau(t)} \Xi^r(\tau(t), Y_N) \right\|_{L_p}^{y_0} \right. \\ &\quad \left. + \left\| e^{-(2\lambda+\omega)\tau(t)} \Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega) \right\|_{L_p}^{y_0} \right\|_{L_p}^{y_0}, \quad \forall (t, \mathbf{p}) \in \mathbb{R}_{\geq 0} \times \mathbb{N}_{\geq 2}, \end{aligned} \quad (36)$$

where, for any Borel measurable function $a : \mathbb{R}^{2n} \rightarrow \mathbb{R}$,

$$\mathbb{E}_{y_0} [a(Y_{N,t})] \doteq \mathbb{E} [a(Y_{N,t}) \mid Y_{N,0} = y_0] = \int_{\mathbb{R}^{2n}} a(y) \mathbb{Y}_N(y_0, 0; dy, t), \quad (37)$$

and where \mathbb{Y}_N is defined in (33). Similarly, $\|a(t)\|_{L_p}^{y_0}$ is defined using the probability measure $\mathbb{Y}_N(y_0, 0; \cdot, t)$. Furthermore,

$$\begin{aligned} \Xi^r(\tau(t), Y_N) &= \int_0^{\tau(t)} e^{2\lambda\nu} (\phi_\mu^r(\nu, Y_{N,\nu}) d\nu + \phi_{\sigma_\star}^r(\nu, Y_{N,\nu}) dW_\nu^\star + \phi_\sigma^r(\nu, Y_{N,\nu}) dW_\nu), \\ \Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega) &= \int_0^{\tau(t)} (\mathcal{U}_\mu^r(\tau(t), \nu, Y_N; \omega) d\nu + \mathcal{U}_\sigma^r(\tau(t), \nu, Y_N; \omega) dW_\nu), \end{aligned} \quad (38)$$

where

$$\begin{aligned} \phi_\mu^r(\nu, Y_{N,\nu}) &= \partial \mathcal{V}(Y_N)^\top g(\nu)^\perp \Lambda_\mu^\perp(\nu, X_{N,\nu}^r) + \|F_\sigma^r(\nu, X_{N,\nu}^r)\|_F^2 + \|\bar{F}_\sigma(\nu, X_{N,\nu}^\star)\|_F^2, \\ \phi_{\sigma_\star}^r(\nu, Y_{N,\nu}) &= -\partial \mathcal{V}(Y_{N,\nu})^\top \bar{F}_\sigma(\nu, X_\nu^\star), \quad \phi_\sigma^r(\nu, Y_{N,\nu}) = \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu)^\perp F_\sigma^\perp(\nu, X_{N,t}^r), \\ \mathcal{U}_\mu^r(\tau(t), \nu, Y_N; \omega) &= \psi^r(\tau(t), \nu, Y_N) \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r), \quad \mathcal{U}_\sigma^r(\tau(t), \nu, Y_N; \omega) = \psi^r(\tau(t), \nu, Y_N) F_\sigma^\parallel(\nu, X_{N,\nu}^r), \end{aligned} \quad (39)$$

and

$$\begin{aligned} \psi^r(\tau(t), \nu, Y_N) &= \frac{\omega}{2\lambda - \omega} \left(e^{\omega(\tau(t)+\nu)} \mathcal{P}^r(\tau(t), \nu) - e^{(2\lambda\tau(t)+\omega\nu)} \partial \mathcal{V}(Y_{N,\tau(t)})^\top g(\tau(t)) \right) \\ &\quad + \frac{2\lambda}{2\lambda - \omega} e^{\omega\tau(t)+2\lambda\nu} \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) \in \mathbb{R}^{1 \times m}. \end{aligned} \quad (40)$$

In the expressions above, we have defined

$$\partial \mathcal{V}(Y_N) = \partial_{X_N^r - X_N^\star} \mathcal{V}(Y_N) = 2(X_N^r - X_N^\star) \in \mathbb{R}^n, \quad H_\sigma = G_\sigma G_\sigma^\top \in \mathbb{S}^{2n},$$

and

$$\mathcal{P}^r(\tau(t), \nu) = \int_\nu^{\tau(t)} e^{(2\lambda-\omega)\beta} d_\beta \left[\partial \mathcal{V}(Y_{N,\beta})^\top g(\beta) \right] \in \mathbb{R}^{1 \times m}, \quad 0 \leq \nu \leq \tau(t), \quad (41)$$

where $d_\beta[\cdot]$ denotes the stochastic differential with respect to β .

Proof. We assume w.l.o.g. that $\mathbb{P}(y_0 \in U_N) = 1$, since the current and all consecutive proofs can be extended for the general case by following the approach in [106, Thm. 8.3.1].

Since t^\star is a constant, the fact that τ_N is a stopping time implies that τ^\star is a stopping time as well [107, Sec. 6.1]. Additionally, from Proposition 4.1, we know that $Y_{N,t}$ is a unique strong solution of (31), for all $t \in [0, T]$. Consequently, the assumptions on the regularity of the vector fields in Sec. 2.2 imply that the vector fields $G_{N,\mu}$ and $G_{N,\sigma}$, that define the truncated process (31), are uniformly bounded and globally Lipschitz continuous on \mathbb{R}^{2n} . It is thus straightforward to show that $Y_{N,t}$ is a strong Markov process by invoking [91, Thm. 2.9.3]. Hence, the process $Y_{N,\tau(t)}$ obtained by stopping the process $Y_{N,t}$ at either the first (random) instant it leaves the set U_N (or when $t = t^\star$) is a Markov process and is well-defined [48, Lem. 3.2]. We may thus apply the Itô lemma [91, Thm. 1.6.4] to $e^{2\lambda\tau(t)} \mathcal{V}(Y_{N,\tau(t)})$ using the dynamics in (31) to obtain

$$\begin{aligned} e^{2\lambda\tau(t)} \mathcal{V}(Y_{N,\tau(t)}) &= \mathcal{V}(y_0) + 2\lambda \int_0^{\tau(t)} e^{2\lambda\nu} \mathcal{V}(Y_{N,\nu}) d\nu + \int_0^{\tau(t)} e^{2\lambda\nu} \nabla \mathcal{V}(Y_{N,\nu})^\top G_{N,\sigma}(\nu, Y_{N,\nu}) d\widehat{W}_\nu \\ &\quad + \int_0^{\tau(t)} e^{2\lambda\nu} \left(\nabla \mathcal{V}(Y_{N,\nu})^\top G_{N,\mu}(\nu, Y_{N,\nu}) + \frac{1}{2} \text{Tr} [H_{N,\sigma}(\nu, Y_{N,\nu}) \nabla^2 \mathcal{V}(Y_{N,\nu})] \right) d\nu, \end{aligned}$$

for all $t \in \mathbb{R}_{\geq 0}$, where $H_{N,\sigma}(\nu, Y_{N,\nu}) = G_{N,\sigma}(\nu, Y_{N,\nu}) G_{N,\sigma}(\nu, Y_{N,\nu})^\top$. Note that since $[0, \tau^*] \subseteq [0, \tau_N]$, Proposition 4.1 allows us to conclude that $Y_{N,t}$ is also unique strong solution of (29), for all $t \in [0, \tau^*]$. Thus, we may replace $G_{N,\mu}$, $G_{N,\sigma}$, and $H_{N,\sigma}$ with G_μ , G_σ , and H_σ , respectively, in the above inequality and obtain

$$\begin{aligned} e^{2\lambda\tau(t)} \mathcal{V}(Y_{N,\tau(t)}) &= \mathcal{V}(y_0) + 2\lambda \int_0^{\tau(t)} e^{2\lambda\nu} \mathcal{V}(Y_{N,\nu}) d\nu + \int_0^{\tau(t)} e^{2\lambda\nu} \nabla \mathcal{V}(Y_{N,\nu})^\top G_\sigma(\nu, Y_{N,\nu}) d\widehat{W}_\nu \\ &\quad + \int_0^{\tau(t)} e^{2\lambda\nu} \left(\nabla \mathcal{V}(Y_{N,\nu})^\top G_\mu(\nu, Y_{N,\nu}) + \frac{1}{2} \text{Tr} [H_\sigma(\nu, Y_{N,\nu}) \nabla^2 \mathcal{V}(Y_{N,\nu})] \right) d\nu, \end{aligned}$$

for all $t \in \mathbb{R}_{\geq 0}$, where $H_\sigma(\nu, Y_{N,\nu}) = G_\sigma(\nu, Y_{N,\nu}) G_\sigma(\nu, Y_{N,\nu})^\top$. Substituting the expressions in (C.9), Proposition C.1, for the last two terms on the right hand side of the above expression and re-arranging terms leads to

$$\begin{aligned} e^{2\lambda\tau(t)} \mathcal{V}(Y_{N,\tau(t)}) &\leq \mathcal{V}(y_0) + \int_0^{\tau(t)} e^{2\lambda\nu} (\phi_\mu^r(\nu, Y_{N,\nu}) d\nu + \phi_{\sigma^*}^r(\nu, Y_{N,\nu}) dW_\nu^* + \phi_\sigma^r(\nu, Y_{N,\nu}) dW_\nu) \\ &\quad + \int_0^{\tau(t)} e^{2\lambda\nu} \left([\phi_U^r(\nu, Y_{N,\nu}) + \phi_{\mu^\parallel}^r(\nu, Y_{N,\nu})] d\nu + \phi_{\sigma^\parallel}^r(\nu, Y_{N,\nu}) dW_\nu \right), \quad (42) \end{aligned}$$

for all $t \in \mathbb{R}_{\geq 0}$, where

$$\begin{aligned} \phi_U^r(\nu, Y_{N,\nu}) &= \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) U_\nu^r, \quad \phi_{\mu^\parallel}^r(\nu, Y_{N,\nu}) = \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r), \\ \phi_{\sigma^\parallel}^r(\nu, Y_{N,\nu}) &= \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r). \end{aligned}$$

Next, we use Proposition C.2 to obtain the following expression

$$\begin{aligned} &\int_0^{\tau(t)} e^{2\lambda\nu} \left([\phi_U^r(\nu, Y_{N,\nu}) + \phi_{\mu^\parallel}^r(\nu, Y_{N,\nu})] d\nu + \phi_{\sigma^\parallel}^r(\nu, Y_{N,\nu}) dW_\nu \right) \\ &= \int_0^{\tau(t)} e^{2\lambda\nu} \left(\phi_{\mu^\parallel}^r(\nu, Y_{N,\nu}) + \phi_{U_\mu}^r(\nu, Y_{N,\nu}; \omega) \right) d\nu + \int_0^{\tau(t)} e^{2\lambda\nu} \left(\phi_{\sigma^\parallel}^r(\nu, Y_{N,\nu}) + \phi_{U_\sigma}^r(\nu, Y_{N,\nu}; \omega) \right) dW_\nu \\ &\quad + \int_0^{\tau(t)} \left(\hat{\mathcal{U}}_\mu^r(\tau(t), \nu, Y_N; \omega) d\nu + \hat{\mathcal{U}}_\sigma^r(\tau(t), \nu, Y_N; \omega) dW_\nu \right), \quad \forall t \in \mathbb{R}_{\geq 0}, \quad (43) \end{aligned}$$

where

$$\begin{aligned} \hat{\mathcal{U}}_\mu^r(\tau(t), \nu, Y_N; \omega) &= e^{-\omega\tau(t)} \frac{\omega}{2\lambda - \omega} \left(e^{\omega\tau(t)} \mathcal{P}^r(\tau(t), \nu) - e^{2\lambda\tau(t)} \partial \mathcal{V}(Y_{N,\tau(t)})^\top g(\tau(t)) \right) e^{\omega\nu} \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r), \\ \hat{\mathcal{U}}_\sigma^r(\tau(t), \nu, Y_N; \omega) &= e^{-\omega\tau(t)} \frac{\omega}{2\lambda - \omega} \left(e^{\omega\tau(t)} \mathcal{P}^r(\tau(t), \nu) - e^{2\lambda\tau(t)} \partial \mathcal{V}(Y_{N,\tau(t)})^\top g(\tau(t)) \right) e^{\omega\nu} F_\sigma^\parallel(\nu, X_{N,\nu}^r), \\ \phi_{U_\mu}^r(\nu, Y_{N,\nu}; \omega) &= \frac{\omega}{2\lambda - \omega} \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r), \\ \phi_{U_\sigma}^r(\nu, Y_{N,\nu}; \omega) &= \frac{\omega}{2\lambda - \omega} \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r). \end{aligned}$$

Using the definitions of $\phi_{\mu^\parallel}^r$ and $\phi_{\sigma^\parallel}^r$ in (42), we obtain

$$\begin{aligned} \phi_{\mu^\parallel}^r(\nu, Y_{N,\nu}) + \phi_{U_\mu}^r(\nu, Y_{N,\nu}; \omega) &= \frac{2\lambda}{2\lambda - \omega} \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r), \\ \phi_{\sigma^\parallel}^r(\nu, Y_{N,\nu}) + \phi_{U_\sigma}^r(\nu, Y_{N,\nu}; \omega) &= \frac{2\lambda}{2\lambda - \omega} \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r). \end{aligned}$$

Substituting into (43) thus produces

$$\begin{aligned} &\int_0^{\tau(t)} e^{2\lambda\nu} \left([\phi_U^r(\nu, Y_{N,\nu}) + \phi_{\mu^\parallel}^r(\nu, Y_{N,\nu})] d\nu + \phi_{\sigma^\parallel}^r(\nu, Y_{N,\nu}) dW_\nu \right) \\ &= e^{-\omega\tau(t)} \int_0^{\tau(t)} \left(\mathcal{U}_\mu^r(\tau(t), \nu, Y_N; \omega) d\nu + \mathcal{U}_\sigma^r(\tau(t), \nu, Y_N; \omega) dW_\nu \right), \quad \forall t \in \mathbb{R}_{\geq 0}, \end{aligned}$$

where \mathcal{U}_μ^r and \mathcal{U}_σ^r are defined in (39). Substituting the above expression for the last integral on the right hand side of (42) and using the definitions of Ξ^r and $\Xi_{\mathcal{U}}^r$ in (38) yields the following bound:

$$e^{2\lambda\tau(t)}\mathcal{V}(Y_{N,\tau(t)}) \leq \mathcal{V}(y_0) + \Xi^r(\tau(t), Y_N) + e^{-\omega\tau(t)}\Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega), \quad \forall t \in \mathbb{R}_{\geq 0},$$

which in turn produces

$$\mathcal{V}(Y_{N,\tau(t)}) \leq e^{-2\lambda\tau(t)}\mathcal{V}(y_0) + e^{-2\lambda\tau(t)}\Xi^r(\tau(t), Y_N) + e^{-(2\lambda+\omega)\tau(t)}\Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega), \quad \forall t \in \mathbb{R}_{\geq 0}. \quad (44)$$

Then, using the linearity of the expectation operator, we obtain

$$\begin{aligned} \mathbb{E}_{y_0} [\mathcal{V}(Y_{N,\tau(t)})] &\leq \mathbb{E}_{y_0} [e^{-2\lambda\tau(t)}] \mathcal{V}(y_0) + \mathbb{E}_{y_0} [e^{-2\lambda\tau(t)}\Xi^r(\tau(t), Y_N)] \\ &\quad + \mathbb{E}_{y_0} [e^{-(2\lambda+\omega)\tau(t)}\Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega)], \quad \forall t \in \mathbb{R}_{\geq 0}, \end{aligned}$$

which establishes (35). Finally, applying the Minkowski's inequality to (44) for $p \in \mathbb{N}_{\geq 2}$ yields the bound in (36), thus concluding the proof. \square

The next result establishes moment bounds for the truncated system up to the first exit time τ_N .

Lemma 4.2 *Let the assumptions in Sec. 2.2 hold and let $\mathbf{q}_r(p) \in \mathbb{R}_{>0}$, $p \in \{1, \dots, p^*\}$, be as defined in (16a) with $\alpha(X_0^r, X_0^*)_{2p} = \|X_0^r - X_0^*\|_{L_{2p}}$, where p^* is defined in Assumption 3. If the filter bandwidth condition in (21a) holds with $\alpha(X_0^r, X_0^*)_{2p} = \|X_0^r - X_0^*\|_{L_{2p}}$, then the truncated joint process $Y_{N,t} = (X_{N,t}^r, X_{N,t}^*)$ in Definition 6, satisfies*

$$\|X_{N,t}^r - X_{N,t}^*\|_{L_{2p}} < \mathbf{q}_r(p), \quad \forall t \in [0, \tau_N], \quad (45)$$

where the stopping time τ_N is as in (32) for an arbitrary $N \in \mathbb{R}_{>0}$.

Proof. We prove (45) by contradiction. The t -continuity of the strong solutions $X_{N,t}^r$ and $X_{N,t}^*$ on $[0, \tau_N]$ implies that $\|X_{N,t}^r - X_{N,t}^*\|_{L_{2p}}$ is t -continuous as well. Moreover, as we presented in (20), the definition of $\mathbf{q}_r(p) \in \mathbb{R}_{>0}$ in (16a) implies that

$$\mathbf{q}_r(p) > \|X_{N,0}^r - X_{N,0}^*\|_{L_{2p}}. \quad (46)$$

Then, (46) and the deterministic nature of $t \mapsto \|X_{N,t}^r - X_{N,t}^*\|_{L_{2p}}$ allow us to formulate the negation of (45) as

$$\neg (45) : \exists \text{ (deterministic) } t^* \in [0, \tau_N] \text{ such that } \begin{cases} \|X_{N,t}^r - X_{N,t}^*\|_{L_{2p}} < \mathbf{q}_r(p), & \forall t \in [0, t^*) \\ \|X_{N,t^*}^r - X_{N,t^*}^*\|_{L_{2p}} = \mathbf{q}_r(p). \end{cases} \quad (47)$$

Now, let π_0 be the probability measure on $\mathcal{B}(\mathbb{R}^{2n})$ induced by the initial condition $Y_{N,0} = Y_0 = (X_0^r, X_0^*) = (x_0, x_0^*) = y_0$ of the truncated joint process. Since the initial vector $Y_{N,0} = Y_0$ is independent of the Brownian motion \widehat{W}_t , the strong solution $Y_{N,\tau(t)}$ is adapted to the filtration generated by Y_0 and \widehat{W}_t [15, Thm. 5.2.1]. Let us denote by $\mathbb{Y}_N(\cdot, t)$ the law of the process $Y_{N,t}$ under $Y_{N,0} \sim \pi_0$. Hence, $\mathbb{Y}_N(\cdot, t)$ is a probability measure on \mathbb{R}^{2n} given as

$$\mathbb{Y}_N(B, t) = \mathbb{P}(Y_{N,t} \in B) = \int_{\mathbb{R}^{2n}} \mathbb{Y}_N(y_0, 0; B, t) \pi_0(dy_0), \quad \forall B \in \mathcal{B}(\mathbb{R}^{2n}), \quad (48)$$

where the integrand is the transition probability defined in (33). Then,

$$\begin{aligned} \mathbb{E}[\mathcal{V}(Y_{N,\tau(t)})] &= \int_{\mathbb{R}^{2n}} \mathcal{V}(y) \mathbb{Y}_N(dy, \tau(t)) = \int_{\mathbb{R}^{2n}} \int_{\mathbb{R}^{2n}} \mathcal{V}(y) \mathbb{Y}_N(y_0, 0; dy, \tau(t)) \pi_0(dy_0) \\ &= \int_{\mathbb{R}^{2n}} \mathbb{E}_{y_0}[\mathcal{V}(Y_{N,\tau(t)})] \pi_0(dy_0), \quad t \in \mathbb{R}_{\geq 0}, \end{aligned}$$

where the last expression follows from (37), and $\tau(t)$ is defined in (34). Since we are assuming that $t^* < \tau_N$, we have from (34) that $\tau^* = t^*$ and thus $\tau(t) = t \wedge t^* \in [0, t^*]$, for all $t \in \mathbb{R}_{\geq 0}$. Hence, we can write the expression above as

$$\mathbb{E}[\mathcal{V}(Y_{N,t})] = \int_{\mathbb{R}^{2n}} \mathbb{E}_{y_0}[\mathcal{V}(Y_{N,\tau(t)})] \pi_0(dy_0), \quad t \in [0, t^*]. \quad (49)$$

Similarly, we can write

$$\begin{aligned}\|\mathcal{V}(Y_{N,t})\|_{L_p} &= \mathbb{E} [\mathcal{V}(Y_{N,t})^p]^{\frac{1}{p}} = \left(\int_{\mathbb{R}^{2n}} \mathbb{E}_{y_0} [\mathcal{V}(Y_{N,\tau(t)})^p] \pi_0(dy_0) \right)^{\frac{1}{p}} \\ &= \left(\int_{\mathbb{R}^{2n}} \left(\|\mathcal{V}(Y_{N,\tau(t)})\|_{L_p}^{y_0} \right)^p \pi_0(dy_0) \right)^{\frac{1}{p}}, \quad t \in [0, t^*].\end{aligned}\quad (50)$$

As stated above $\tau(t) \in [0, t^*]$, for all $t \in \mathbb{R}_{\geq 0}$. Thus, we may use Lemma 4.1 by substituting (35) into (49), which produces

$$\begin{aligned}\mathbb{E} [\mathcal{V}(Y_{N,t})] &\leq \int_{\mathbb{R}^{2n}} \mathcal{V}(y_0) \pi_0(dy_0) + \int_{\mathbb{R}^{2n}} \mathbb{E}_{y_0} \left[e^{-2\lambda\tau(t)} \Xi^r(\tau(t), Y_N) \right] \pi_0(dy_0) \\ &\quad + \int_{\mathbb{R}^{2n}} \mathbb{E}_{y_0} \left[e^{-(2\lambda+\omega)\tau(t)} \Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega) \right] \pi_0(dy_0), \quad \forall t \in [0, t^*],\end{aligned}\quad (51)$$

where we have used the bound $e^{-2\lambda\tau(t)} \leq 1$. Now, recall the definition $\mathcal{V}(Y_N) \doteq \|X_N^r - X_N^*\|^2$ from the statement of Lemma 4.1. Hence, one sees that

$$\begin{aligned}\int_{\mathbb{R}^{2n}} \mathcal{V}(y_0) \pi_0(dy_0) &= \int_{\mathbb{R}^{2n}} \|x_0 - x_0^*\|^2 \pi_0(dy_0) \\ &= \int_{\mathbb{R}^{2n}} \|x_0 - x_0^*\|^2 \pi_0(dx_0, dx_0^*) = \mathbb{E} [\|X_0^r - X_0^*\|^2] = \|X_0^r - X_0^*\|_{L_2}^2,\end{aligned}$$

which, upon substituting into (51) produces

$$\begin{aligned}\mathbb{E} [\mathcal{V}(Y_{N,t})] &\leq \|X_0^r - X_0^*\|_{L_2}^2 + \int_{\mathbb{R}^{2n}} \mathbb{E}_{y_0} \left[e^{-2\lambda\tau(t)} \Xi^r(\tau(t), Y_N) \right] \pi_0(dy_0) \\ &\quad + \int_{\mathbb{R}^{2n}} \mathbb{E}_{y_0} \left[e^{-(2\lambda+\omega)\tau(t)} \Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega) \right] \pi_0(dy_0), \quad \forall t \in [0, t^*].\end{aligned}\quad (52)$$

Similarly, substituting (36) into (50) leads to

$$\begin{aligned}\|\mathcal{V}(Y_{N,t})\|_{L_p} &\leq \left(\int_{\mathbb{R}^{2n}} \mathcal{V}(y_0)^p \pi_0(dy_0) \right)^{\frac{1}{p}} \\ &\quad + \left(\int_{\mathbb{R}^{2n}} \left(\left\| e^{-2\lambda\tau(t)} \Xi^r(\tau(t), Y_N) \right\|_{L_p}^{y_0} + \left\| e^{-(2\lambda+\omega)\tau(t)} \Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega) \right\|_{L_p}^{y_0} \right)^p \pi_0(dy_0) \right)^{\frac{1}{p}},\end{aligned}\quad (53)$$

for all $(t, p) \in [0, t^*] \times \{2, \dots, p^*\}$, where, as above, we have used the bound $e^{-2\lambda\tau(t)} \leq 1$, along with the Minkowski's inequality. Once again, it follows from the definition of $\mathcal{V}(Y_N)$ that

$$\begin{aligned}\left(\int_{\mathbb{R}^{2n}} \mathcal{V}(y_0)^p \pi_0(dy_0) \right)^{\frac{1}{p}} &= \left(\int_{\mathbb{R}^{2n}} \|x_0 - x_0^*\|^{2p} \pi_0(dy_0) \right)^{\frac{1}{p}} \\ &= \left(\int_{\mathbb{R}^{2n}} \|x_0 - x_0^*\|^{2p} \pi_0(dx_0, dx_0^*) \right)^{\frac{1}{p}} = \|X_0^r - X_0^*\|_{L_{2p}}^2.\end{aligned}$$

Substituting the above into (53) leads to

$$\begin{aligned}\|\mathcal{V}(Y_{N,t})\|_{L_p} &\leq \|X_0^r - X_0^*\|_{L_{2p}}^2 \\ &\quad + \left(\int_{\mathbb{R}^{2n}} \left(\left\| e^{-2\lambda\tau(t)} \Xi^r(\tau(t), Y_N) \right\|_{L_p}^{y_0} + \left\| e^{-(2\lambda+\omega)\tau(t)} \Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega) \right\|_{L_p}^{y_0} \right)^p \pi_0(dy_0) \right)^{\frac{1}{p}},\end{aligned}\quad (54)$$

for all $(t, p) \in [0, t^*] \times \{2, \dots, p^*\}$. Now, let us define

$$\tilde{\mathcal{E}}'_{r,y_0}(t, 2p) \doteq \|X_{N,t}^r - X_{N,t}^*\|_{L_{2p}}^{y_0}, \quad \mathcal{E}'_{r,y_0}(t, 2p) \doteq \|X_{N,t}^r\|_{L_{2p}}^{y_0}, \quad (55a)$$

$$\tilde{\mathcal{E}}_{r,y_0}(2p) \doteq \sup_{\nu \in [0, t^*]} \|X_{N,\nu}^r - X_{N,\nu}^*\|_{L_{2p}}^{y_0}, \quad \mathcal{E}_{r,y_0}(2p) \doteq \sup_{\nu \in [0, t^*]} \|X_{N,\nu}^r\|_{L_{2p}}^{y_0}, \quad (55b)$$

for $p \in \{1, \dots, p^*\}$. Then, we can write

$$\begin{aligned}\mathbb{E} [\mathcal{V}(Y_{N,t})] &= \int_{\mathbb{R}^{2n}} \mathbb{E}_{y_0} [\mathcal{V}(Y_{N,t})] \pi_0(dy_0) = \int_{\mathbb{R}^{2n}} \tilde{\mathcal{E}}'_{r,y_0}(t, 2)^2 \pi_0(dy_0) \\ &= \mathbb{E}_{\pi_0} [\tilde{\mathcal{E}}'_{r,y_0}(t, 2)^2] = \left\| \tilde{\mathcal{E}}'_{r,y_0}(t, 2)^2 \right\|_{L_1}^{\pi_0}, \quad t \in [0, t^*],\end{aligned}$$

and

$$\begin{aligned}\|\mathcal{V}(Y_{N,t})\|_{L_p} &= \left(\int_{\mathbb{R}^{2n}} \left(\|\mathcal{V}(Y_{N,t})\|_{L_p}^{y_0} \right)^p \pi_0(dy_0) \right)^{\frac{1}{p}} \\ &= \left(\int_{\mathbb{R}^{2n}} \tilde{\mathcal{E}}'_{r,y_0}(t, 2p)^{2p} \pi_0(dy_0) \right)^{\frac{1}{p}} = \left\| \tilde{\mathcal{E}}'_{r,y_0}(t, 2p)^2 \right\|_{L_p}^{\pi_0}, \quad t \in [0, t^*].\end{aligned}$$

Hence, we can write (52) and (54) as

$$\begin{aligned}\left\| \tilde{\mathcal{E}}'_{r,y_0}(t, 2)^2 \right\|_{L_1}^{\pi_0} &\leq \|X_0^r - X_0^*\|_{L_2}^2 + \int_{\mathbb{R}^{2n}} \mathbb{E}_{y_0} \left[e^{-2\lambda\tau(t)} \Xi^r(\tau(t), Y_N) \right] \pi_0(dy_0) \\ &\quad + \int_{\mathbb{R}^{2n}} \mathbb{E}_{y_0} \left[e^{-(2\lambda+\omega)\tau(t)} \Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega) \right] \pi_0(dy_0), \quad \forall t \in [0, t^*],\end{aligned}\quad (56)$$

and

$$\begin{aligned}\left\| \tilde{\mathcal{E}}'_{r,y_0}(t, 2p)^2 \right\|_{L_p}^{\pi_0} &\leq \|X_0^r - X_0^*\|_{L_{2p}}^2 \\ &\quad + \left(\int_{\mathbb{R}^{2n}} \left(\left\| e^{-2\lambda\tau(t)} \Xi^r(\tau(t), Y_N) \right\|_{L_p}^{y_0} + \left\| e^{-(2\lambda+\omega)\tau(t)} \Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega) \right\|_{L_p}^{y_0} \right)^p \pi_0(dy_0) \right)^{\frac{1}{p}},\end{aligned}\quad (57)$$

for all $(t, p) \in [0, t^*] \times \{2, \dots, p^*\}$. Since $\tau^* = t^*$ by assumption, we can substitute the bounds derived in Lemma C.3 into the above inequalities and get

$$\left\| \tilde{\mathcal{E}}'_{r,y_0}(t, 2)^2 \right\|_{L_1}^{\pi_0} \leq \|X_0^r - X_0^*\|_{L_2}^2 + \int_{\mathbb{R}^{2n}} \Phi^r(1, \omega) \pi_0(dy_0), \quad \forall t \in [0, t^*],\quad (58)$$

and

$$\left\| \tilde{\mathcal{E}}'_{r,y_0}(t, 2p)^2 \right\|_{L_p}^{\pi_0} \leq \|X_0^r - X_0^*\|_{L_{2p}}^2 + \left(\int_{\mathbb{R}^{2n}} \Phi^r(p, \omega)^p \pi_0(dy_0) \right)^{\frac{1}{p}}, \quad \forall (t, p) \in [0, t^*] \times \{2, \dots, p^*\},\quad (59)$$

where we have defined

$$\begin{aligned}\Phi^r(p, \omega) &= \Delta_{\circ}^r(p, \omega) + \check{\Delta}_{\ominus}^r(p, \omega) (\mathcal{E}_{r,y_0}(2p))^{\frac{1}{2}} + \check{\Delta}_{\ominus}^r(p, \omega) \mathcal{E}_{r,y_0}(2p) + \Delta_{\ominus}^r(p, \omega) \tilde{\mathcal{E}}_{r,y_0}(2p) \\ &\quad + \left(\check{\Delta}_{\otimes}^r(p, \omega) \mathcal{E}_{r,y_0}(2p) + \Delta_{\otimes}^r(p, \omega) \tilde{\mathcal{E}}_{r,y_0}(2p) \right) (\mathcal{E}_{r,y_0}(2p))^{\frac{1}{2}} \\ &\quad + \left(\check{\Delta}_{\oplus}^r(p, \omega) \mathcal{E}_{r,y_0}(2p) + \Delta_{\oplus}^r(p, \omega) \tilde{\mathcal{E}}_{r,y_0}(2p) \right) (\mathcal{E}_{r,y_0}(2p)), \quad p \in \{1, \dots, p^*\}.\end{aligned}$$

The two bounds above can be combined into the following single bound over $p \in \{1, \dots, p^*\}$:

$$\left\| \tilde{\mathcal{E}}'_{r,y_0}(t, 2p)^2 \right\|_{L_p}^{\pi_0} \leq \|X_0^r - X_0^*\|_{L_{2p}}^2 + \left(\int_{\mathbb{R}^{2n}} \Phi^r(p, \omega)^p \pi_0(dy_0) \right)^{\frac{1}{p}}, \quad \forall (t, p) \in [0, t^*] \times \{1, \dots, p^*\}.\quad (60)$$

It then follows from the Minkowski's and the Cauchy-Schwarz inequalities that

$$\begin{aligned}&\left(\int_{\mathbb{R}^{2n}} \Phi^r(p, \omega)^p \pi_0(dy_0) \right)^{\frac{1}{p}} \\ &\leq \Delta_{\circ}^r(p, \omega) + \check{\Delta}_{\ominus}^r(p, \omega) \left\| \mathcal{E}_{r,y_0}(2p)^{\frac{1}{2}} \right\|_{L_p}^{\pi_0} + \check{\Delta}_{\ominus}^r(p, \omega) \left\| \mathcal{E}_{r,y_0}(2p) \right\|_{L_p}^{\pi_0} + \Delta_{\ominus}^r(p, \omega) \left\| \tilde{\mathcal{E}}_{r,y_0}(2p) \right\|_{L_p}^{\pi_0} \\ &\quad + \check{\Delta}_{\otimes}^r(p, \omega) \left\| \mathcal{E}_{r,y_0}(2p)^{\frac{3}{2}} \right\|_{L_p}^{\pi_0} + \Delta_{\otimes}^r(p, \omega) \left\| \tilde{\mathcal{E}}_{r,y_0}(2p) \right\|_{L_{2p}}^{\pi_0} \left\| \mathcal{E}_{r,y_0}(2p)^{\frac{1}{2}} \right\|_{L_{2p}}^{\pi_0}\end{aligned}$$

$$+ \check{\Delta}_{\otimes}^r(\mathbf{p}, \boldsymbol{\omega}) \|\mathcal{E}_{r,y_0}(2\mathbf{p})\|_{L_p}^{\pi_0} + \Delta_{\otimes}^r(\mathbf{p}, \boldsymbol{\omega}) \|\tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0} \|\mathcal{E}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0}, \quad (61)$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \{1, \dots, \mathbf{p}^*\}$.

Now, we have the following inequalities due to the Cauchy Schwarz (C-S) and Jensen's (J) inequalities:

$$\left\| \mathcal{E}_{r,y_0}(2\mathbf{p})^{\frac{1}{2}} \right\|_{L_p}^{\pi_0} \stackrel{\text{(C-S)}}{\leq} \left\| \mathcal{E}_{r,y_0}(2\mathbf{p})^{\frac{1}{2}} \right\|_{L_{2p}}^{\pi_0} \stackrel{\text{(J)}}{\leq} \left(\|\mathcal{E}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0} \right)^{\frac{1}{2}}, \quad \|\mathcal{E}_{r,y_0}(2\mathbf{p})\|_{L_p}^{\pi_0} \stackrel{\text{(C-S)}}{\leq} \|\mathcal{E}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0}, \quad (62a)$$

$$\left\| \mathcal{E}_{r,y_0}(2\mathbf{p})^{\frac{3}{2}} \right\|_{L_p}^{\pi_0} \stackrel{\text{(C-S)}}{\leq} \|\mathcal{E}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0} \left\| \mathcal{E}_{r,y_0}(2\mathbf{p})^{\frac{1}{2}} \right\|_{L_{2p}}^{\pi_0} \stackrel{\text{(J)}}{\leq} \left(\|\mathcal{E}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0} \right)^{\frac{3}{2}}, \quad (62b)$$

$$\|\mathcal{E}_{r,y_0}(2\mathbf{p})^2\|_{L_p}^{\pi_0} = \|\mathcal{E}_{r,y_0}(2\mathbf{p})\mathcal{E}_{r,y_0}(2\mathbf{p})\|_{L_p}^{\pi_0} \stackrel{\text{(C-S)}}{\leq} \left(\|\mathcal{E}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0} \right)^2, \quad (62c)$$

$$\|\tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p})\|_{L_p}^{\pi_0} \stackrel{\text{(C-S)}}{\leq} \|\tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0}. \quad (62d)$$

Substituting these bounds into (61) produces

$$\begin{aligned} & \left(\int_{\mathbb{R}^{2n}} \Phi^r(\mathbf{p}, \boldsymbol{\omega})^p \pi_0(dy_0) \right)^{\frac{1}{p}} \\ & \leq \Delta_{\odot}^r(\mathbf{p}, \boldsymbol{\omega}) + \check{\Delta}_{\odot}^r(\mathbf{p}, \boldsymbol{\omega}) \left(\|\mathcal{E}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0} \right)^{\frac{1}{2}} + \check{\Delta}_{\odot}^r(\mathbf{p}, \boldsymbol{\omega}) \|\mathcal{E}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0} + \Delta_{\odot}^r(\mathbf{p}, \boldsymbol{\omega}) \|\tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0} \\ & \quad + \left(\check{\Delta}_{\otimes}^r(\mathbf{p}, \boldsymbol{\omega}) \|\mathcal{E}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0} + \Delta_{\otimes}^r(\mathbf{p}, \boldsymbol{\omega}) \|\tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0} \right) \left(\|\mathcal{E}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0} \right)^{\frac{1}{2}} \\ & \quad + \left(\check{\Delta}_{\otimes}^r(\mathbf{p}, \boldsymbol{\omega}) \|\mathcal{E}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0} + \Delta_{\otimes}^r(\mathbf{p}, \boldsymbol{\omega}) \|\tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0} \right) \|\mathcal{E}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0}, \quad (63) \end{aligned}$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \{1, \dots, \mathbf{p}^*\}$.

Next, using the continuity of $\nu \mapsto \|X_{N,\nu}^r - X_{N,\nu}^*\|_{L_{2p}}^{y_0}$ over the closed interval $[0, t^*]$, we may assume the existence of some $t' \in [0, t^*]$, such that

$$\sup_{\nu \in [0, t^*]} \|X_{N,\nu}^r - X_{N,\nu}^*\|_{L_{2p}}^{y_0} = \|X_{N,t'}^r - X_{N,t'}^*\|_{L_{2p}}^{y_0}.$$

Therefore, using the definition of $\tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p})$ in (55b), we see that

$$\begin{aligned} \|\tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0} &= \left(\int_{\mathbb{R}^{2n}} \left(\sup_{\nu \in [0, t^*]} \|X_{N,\nu}^r - X_{N,\nu}^*\|_{L_{2p}}^{y_0} \right)^{2p} \pi_0(dy_0) \right)^{\frac{1}{2p}} \\ &= \left(\int_{\mathbb{R}^{2n}} \left(\|X_{N,t'}^r - X_{N,t'}^*\|_{L_{2p}}^{y_0} \right)^{2p} \pi_0(dy_0) \right)^{\frac{1}{2p}} = \left(\int_{\mathbb{R}^{2n}} \mathbb{E}_{y_0} \left[\|X_{N,t'}^r - X_{N,t'}^*\|_{L_{2p}}^{2p} \right] \pi_0(dy_0) \right)^{\frac{1}{2p}}. \end{aligned}$$

We can further develop the expression as follows

$$\begin{aligned} \|\tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0} &= \left(\int_{\mathbb{R}^{2n}} \mathbb{E}_{y_0} \left[\|X_{N,t'}^r - X_{N,t'}^*\|_{L_{2p}}^{2p} \right] \pi_0(dy_0) \right)^{\frac{1}{2p}} \\ &= \left(\int_{\mathbb{R}^{2n}} \mathbb{E} \left[\|X_{N,t'}^r - X_{N,t'}^*\|_{L_{2p}}^{2p} \mid Y_{N,0} = y_0 \right] \pi_0(dy_0) \right)^{\frac{1}{2p}} = \|X_{N,t'}^r - X_{N,t'}^*\|_{L_{2p}}, \quad (64) \end{aligned}$$

for some $t' \in [0, t^*]$, and thus

$$\|\tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0} \leq \sup_{\nu \in [0, t^*]} \|X_{N,\nu}^r - X_{N,\nu}^*\|_{L_{2p}}. \quad (65)$$

Similarly, using the definition of $\mathcal{E}_{r,y_0}(2\mathbf{p})$ in (55b) and the Minkowski's inequality, we conclude that there exists some $t'' \in [0, t^*]$ such that

$$\|\mathcal{E}_{r,y_0}(2\mathbf{p})\|_{L_{2p}}^{\pi_0} = \|X_{N,t''}^r\|_{L_{2p}} \leq \|X_{N,t''}^r - X_{N,t''}^*\|_{L_{2p}} + \|X_{N,t''}^*\|_{L_{2p}}, \quad t'' \in [0, t^*],$$

which further implies

$$\|\mathcal{E}_{r,y_0}(2p)\|_{L_{2p}}^{\pi_0} \leq \sup_{\nu \in [0, t^*]} \|X_{N,\nu}^r - X_{N,\nu}^*\|_{L_{2p}} + \sup_{\nu \in [0, t^*]} \|X_{N,\nu}^*\|_{L_{2p}}. \quad (66)$$

It then follows from (65) - (66), the contradiction statement (47), and Assumption 3

$$\begin{aligned} \|\tilde{\mathcal{E}}_{r,y_0}(2p)\|_{L_{2p}}^{\pi_0} &\leq \sup_{\nu \in [0, t^*]} \|X_{N,\nu}^r - X_{N,\nu}^*\|_{L_{2p}} \leq \underline{q}_r(p), \\ \|\mathcal{E}_{r,y_0}(2p)\|_{L_{2p}}^{\pi_0} &\leq \sup_{\nu \in [0, t^*]} \|X_{N,\nu}^r - X_{N,\nu}^*\|_{L_{2p}} + \sup_{\nu \in [0, t^*]} \|X_{N,\nu}^*\|_{L_{2p}} \leq \underline{q}_r(p) + \Delta_*, \end{aligned}$$

for $p \in \{1, \dots, p^*\}$. Substituting the above into (63), and substituting the resulting bound into (60), we get:

$$\begin{aligned} \|\tilde{\mathcal{E}}'_{r,y_0}(t, 2p)^2\|_{L_p}^{\pi_0} &\leq \|X_0^r - X_0^*\|_{L_{2p}}^2 + \Delta_{\circ}^r(p, \omega) + \check{\Delta}_{\ominus}^r(p, \omega) (\underline{q}_r(p) + \Delta_*)^{\frac{1}{2}} \\ &\quad + \check{\Delta}_{\ominus}^r(p, \omega) (\underline{q}_r(p) + \Delta_*) + \Delta_{\odot}^r(p, \omega) \underline{q}_r(p) \\ &\quad + \left(\check{\Delta}_{\otimes}^r(p, \omega) (\underline{q}_r(p) + \Delta_*) + \Delta_{\otimes}^r(p, \omega) \underline{q}_r(p) \right) (\underline{q}_r(p) + \Delta_*)^{\frac{1}{2}} \\ &\quad + \left(\check{\Delta}_{\oplus}^r(p, \omega) (\underline{q}_r(p) + \Delta_*) + \Delta_{\oplus}^r(p, \omega) \underline{q}_r(p) \right) (\underline{q}_r(p) + \Delta_*), \end{aligned}$$

for all $t \in [0, t^*]$. Substituting the definition of constants $\Delta_i^r(p, \omega)$, $i \in \{\circ, \ominus, \odot, \otimes, \oplus\}$, from Lemma C.3 into the above inequality and re-arranging the terms leads to

$$\|\tilde{\mathcal{E}}'_{r,y_0}(t, 2p)^2\|_{L_p}^{\pi_0} - \frac{\Delta_g^{\perp} \Delta_{\mu}^{\perp}}{\lambda} \underline{q}_r(p)^2 \leq \|X_0^r - X_0^*\|_{L_{2p}}^2 + \Gamma_r(\underline{q}_r(p), p, \omega) + \frac{1}{|2\lambda - \omega|} \Theta_r(\underline{q}_r(p), p, \omega), \quad (67)$$

for all $t \in [0, t^*]$, where $\Gamma_r(\underline{q}_r(p), p, \omega)$ and $\Theta_r(\underline{q}_r(p), p, \omega)$ are defined in (18b) and (22), respectively.

Now, using the definition of $\tilde{\mathcal{E}}'_{r,y_0}(t, 2p)$ from (55a), we see that

$$\begin{aligned} \|\tilde{\mathcal{E}}'_{r,y_0}(t, 2p)^2\|_{L_p}^{\pi_0} &= \left(\int_{\mathbb{R}^{2n}} \left(\|X_{N,t}^r - X_{N,t}^*\|_{L_{2p}}^{y_0} \right)^{2p} \pi_0(dy_0) \right)^{\frac{1}{p}} = \left(\int_{\mathbb{R}^{2n}} \mathbb{E}_{y_0} \left[\|X_{N,t}^r - X_{N,t}^*\|^{2p} \right] \pi_0(dy_0) \right)^{\frac{1}{p}} \\ &= \left(\int_{\mathbb{R}^{2n}} \mathbb{E}_{y_0} \left[\|X_{N,t}^r - X_{N,t}^*\|^{2p} \right] \pi_0(dy_0) \right)^{\frac{2}{2p}}, \end{aligned}$$

which implies that

$$\|\tilde{\mathcal{E}}'_{r,y_0}(t, 2p)^2\|_{L_p}^{\pi_0} = \left(\int_{\mathbb{R}^{2n}} \mathbb{E} \left[\|X_{N,t}^r - X_{N,t}^*\|^{2p} \mid Y_{N,0} = y_0 \right] \pi_0(dy_0) \right)^{\frac{2}{2p}} = \|X_{N,t}^r - X_{N,t}^*\|_{L_{2p}}^2.$$

We can thus write (67) as

$$\begin{aligned} \|X_{N,t}^r - X_{N,t}^*\|_{L_{2p}}^2 - \frac{\Delta_g^{\perp} \Delta_{\mu}^{\perp}}{\lambda} \underline{q}_r(p)^2 \\ \leq \|X_0^r - X_0^*\|_{L_{2p}}^2 + \Gamma_r(\underline{q}_r(p), p, \omega) + \frac{1}{|2\lambda - \omega|} \Theta_r(\underline{q}_r(p), p, \omega), \quad \forall t \in [0, t^*]. \end{aligned} \quad (68)$$

Now, at the temporal instance $t = t^*$, we have from the contradiction statement (47) that

$$\|X_{N,t^*}^r - X_{N,t^*}^*\|_{L_{2p}}^2 = \underline{q}_r(p)^2.$$

Hence, at $t = t^*$, the inequality in (68) reduces to

$$\left(1 - \frac{\Delta_g^{\perp} \Delta_{\mu}^{\perp}}{\lambda} \right) \underline{q}_r(p)^2 \leq \|X_0^r - X_0^*\|_{L_{2p}}^2 + \Gamma_r(\underline{q}_r(p), p, \omega) + \frac{1}{|2\lambda - \omega|} \Theta_r(\underline{q}_r(p), p, \omega),$$

which implies that

$$\frac{1}{|2\lambda - \omega|} \Theta_r(\underline{q}_r(p), p, \omega) \geq \left(1 - \frac{\Delta_g^{\perp} \Delta_{\mu}^{\perp}}{\lambda} \right) \underline{q}_r(p)^2 - \|X_0^r - X_0^*\|_{L_{2p}}^2 - \Gamma_r(\underline{q}_r(p), p, \omega). \quad (69)$$

However, one sees from (21a), with $\alpha(X_0^r, X_0^*)_{2p} = \|X_0^r - X_0^*\|_{L_{2p}}$, that the choice of ω ensures that

$$\frac{1}{|2\lambda - \omega|} \Theta_r(\mathbf{q}_r(\mathbf{p}), \mathbf{p}, \omega) < \left(1 - \frac{\Delta_g^\perp \Delta_\mu^\perp}{\lambda}\right) \mathbf{q}_r(\mathbf{p})^2 - \|X_0^r - X_0^*\|_{L_{2p}}^2 - \Gamma_r(\mathbf{q}_r(\mathbf{p}), \mathbf{p}, \omega),$$

which contradicts the conclusion above in (69) that follows from the postulation in (47). Therefore, the desired result in (45) is proved since its negation (47) leads to a contradiction. \square

Next, we extend the strong solutions of the truncated process to the original joint process in Definition 5.

Lemma 4.3 *Let the assumptions in Sec. 2.2 hold. Additionally, for $\alpha(X_0^r, X_0^*)_{2p} = \|X_0^r - X_0^*\|_{L_{2p}}$ and $\mathbf{p} \in \{1, \dots, p^*\}$, let $\mathbf{q}_r(\mathbf{p}) \in \mathbb{R}_{>0}$ be as defined in (16a) and let filter bandwidth condition in (21a) hold, where p^* is defined in Assumption 3. Then the joint process in Definition 5 admits a unique strong solution $Y_t = (X_t^r, X_t^*)$ for all $T \in (0, \infty)$. Furthermore, the strong solution satisfies*

$$\|X_t^r - X_t^*\|_{L_{2p}} < \mathbf{q}_r(\mathbf{p}), \quad \forall t \in [0, T], \quad (70)$$

for any $T \in (0, \infty)$.

Proof. We prove the existence of a unique strong solution to the reference process in Definition 4 by following the approach in [48, Thm. 3.5], [89, Thm. 5.3.1]. Then, along with well-posedness of the known (nominal) system in Assumption 3, we will be able to conclude the well-posedness of the joint known (nominal)-reference process. Let us define

$$t_N \doteq \inf \{t \in \mathbb{R}_{\geq 0} : Y_{N,t} \notin U_N\},$$

where the open and bounded set U_N is defined in (30), for an arbitrary $N \in \mathbb{R}_{>0}$. Then, we can write the definition of τ_N in (32) as

$$\tau_N \doteq T \wedge t_N,$$

where $T \in (0, \infty)$ is the horizon, and t_N is the first exit time of the truncated joint process $Y_{N,t} = (X_{N,t}^r, X_{N,t}^*)$ from the set U_N .

Now, Lemma 4.2 and Jensen's inequality allow us to conclude

$$\|X_{N,t}^r - X_{N,t}^*\|_{L_2} \leq \|X_{N,t}^r - X_{N,t}^*\|_{L_{2p}} < \mathbf{q}_r(\mathbf{p}), \quad \forall t \in [0, T \wedge t_N],$$

which can be further expressed as

$$\mathbb{E} [\|X_{N,t}^r - X_{N,t}^*\|^2] < \mathbf{q}_r(\mathbf{p})^2, \quad \forall t \in [0, T \wedge t_N]. \quad (71)$$

Furthermore, this bound and Assumption 3 allow us to conclude

$$\mathbb{E} [\|X_{N,t}^r\|^2] \stackrel{(\star)}{\leq} 2\mathbb{E} [\|X_{N,t}^r - X_{N,t}^*\|^2] + 2\mathbb{E} [\|X_{N,t}^*\|^2] \leq 2(\mathbf{q}_r(\mathbf{p})^2 + \Delta_\star^2), \quad \forall t \in [0, T \wedge t_N], \quad (72)$$

where (\star) is due to the triangle inequality followed by [108, Prop. 3.1.10 (iii)].

In order to establish the well-posedness of the reference process over the interval $[0, T]$ it suffices to show that

$$\lim_{N \rightarrow \infty} \mathbb{P}(t_N \leq T) = 0, \quad \forall T \in (0, \infty).$$

Using the definition of the expectation operator, we derive the estimate

$$\begin{aligned} \mathbb{E} [\|X_{N,t_N \wedge T}^r\|^2] &= \int_{\Omega} \|X_{N,t_N \wedge T}^r\|^2 d\mathbb{P} \\ &= \int_{\{t_N \leq T\}} \|X_{N,t_N}^r\|^2 d\mathbb{P} + \int_{\{t_N > T\}} \|X_{N,T}^r\|^2 d\mathbb{P} \geq \int_{\{t_N \leq T\}} \|X_{N,t_N}^r\|^2 d\mathbb{P}. \end{aligned}$$

We similarly obtain

$$\mathbb{E} [\|X_{N,t_N \wedge T}^r - X_{N,t_N \wedge T}^*\|^2] \geq \int_{\{t_N \leq T\}} \|X_{N,t_N}^r - X_{N,t_N}^*\|^2 d\mathbb{P}.$$

Combining the above two inequalities leads to

$$\begin{aligned} \int_{\{t_N \leq T\}} \|X_{N,t_N}^r\|^2 d\mathbb{P} + \int_{\{t_N \leq T\}} \|X_{N,t_N}^r - X_{N,t_N}^*\|^2 d\mathbb{P} \\ \leq \mathbb{E} \left[\|X_{N,t_N \wedge T}^r\|^2 \right] + \mathbb{E} \left[\|X_{N,t_N \wedge T}^r - X_{N,t_N \wedge T}^*\|^2 \right], \quad \forall T \in (0, \infty). \end{aligned}$$

It then follows from the bounds in (71) and (72)

$$\begin{aligned} \int_{\{t_N \leq T\}} \|X_{N,t_N}^r\|^2 d\mathbb{P} + \int_{\{t_N \leq T\}} \|X_{N,t_N}^r - X_{N,t_N}^*\|^2 d\mathbb{P} \\ \leq \mathbb{E} \left[\|X_{N,t_N \wedge T}^r\|^2 \right] + \mathbb{E} \left[\|X_{N,t_N \wedge T}^r - X_{N,t_N \wedge T}^*\|^2 \right] \leq 3\varrho_r(\mathbf{p})^2 + 2\Delta_\star^2, \quad \forall T \in (0, \infty). \end{aligned} \quad (73)$$

Since t_N is the first exit time of the process $Y_{N,t}$ from the set $U_N \subset \mathbb{R}^{2n}$ defined in (30), the stopped process Y_{N,t_N} satisfies

$$Y_{N,t_N} = (X_{N,t_N}^r, X_{N,t_N}^*) \in \partial U_N, \quad \partial U_N = \left\{ z = (z_1, z_2), z_1, z_2 \in \mathbb{R}^n \mid \|z_1\|^2 + \|z_2\|^2 = N^2 \right\}.$$

Therefore, we deduce from (73) that

$$\begin{aligned} \inf_{(z_1, z_2) \in \partial U_N} \left(\|z_1\|^2 + \|z_1 - z_2\|^2 \right) \mathbb{P}(t_N \leq T) \\ \leq \mathbb{E} \left[\|X_{N,t_N \wedge T}^r\|^2 \right] + \mathbb{E} \left[\|X_{N,t_N \wedge T}^r - X_{N,t_N \wedge T}^*\|^2 \right] \leq 3\varrho_r(\mathbf{p})^2 + 2\Delta_\star^2, \quad \forall T \in (0, \infty). \end{aligned} \quad (74)$$

Using the method of Lagrange multipliers, one can show that

$$\inf_{(z_1, z_2) \in \partial U_N} \left(\|z_1\|^2 + \|z_1 - z_2\|^2 \right) = \frac{1}{2} \min \{ 3 + \sqrt{5}, 3 - \sqrt{5} \} N^2 \geq \frac{N^2}{3}.$$

Hence, (74) can be developed into

$$\frac{N^2}{3} \mathbb{P}(t_N \leq T) \leq \inf_{(z_1, z_2) \in \partial U_N} \left(\|z_1\|^2 + \|z_1 - z_2\|^2 \right) \mathbb{P}(t_N \leq T) \leq 3\varrho_r(\mathbf{p})^2 + 2\Delta_\star^2, \quad \forall T \in (0, \infty),$$

which further implies that

$$\mathbb{P}(t_N \leq T) \leq \frac{3}{N^2} (3\varrho_r(\mathbf{p})^2 + 2\Delta_\star^2), \quad \forall T \in (0, \infty).$$

Consequently,

$$\lim_{N \rightarrow \infty} \mathbb{P}(t_N \leq T) = 0, \quad \forall T \in (0, \infty), \quad (75)$$

which establishes the well-posedness of the reference process. We thus define $Y_t = (X_t^r, X_t^*)$, where $X_t^r \doteq \lim_{N \rightarrow \infty} X_{\tau_N(t)}^r$, and $\tau_N(t) = t \wedge \tau_N$. Therefore, it follows from Fatou's lemma [108, Thm. 2.3.7] and Lemma 4.2

$$\|X_t^r - X_t^*\|_{L_{2p}} \leq \liminf_{N \rightarrow \infty} \|X_{\tau_N(t)}^r - X_{\tau_N(t)}^*\|_{L_{2p}} < \varrho_r(\mathbf{p}), \quad \forall t \in [0, T],$$

for any $T \in (0, \infty)$, which establishes the desired result in (70). □

We now formulate uniform ultimate bounds (UUBs) in the following proposition.

Proposition 4.2 *Let the assumptions in Sec. 2.2 hold and let $\varrho_r(\mathbf{p}) \in \mathbb{R}_{>0}$, $\mathbf{p} \in \{1, \dots, \mathbf{p}^*\}$, be as defined in (16a) with $\alpha(X_0^r, X_0^*)_{2p} = \|X_0^r - X_0^*\|_{L_{2p}}$, where \mathbf{p}^* is defined in Assumption 3. If the filter bandwidth condition in (21a) holds with $\alpha(X_0^r, X_0^*)_{2p} = \|X_0^r - X_0^*\|_{L_{2p}}$, then for any deterministic $\hat{t} \in (0, T)$, the joint process $Y_t = (X_t^r, X_t^*)$ in Definition 5 satisfies*

$$\|X_t^r - X_t^*\|_{L_{2p}} \leq \hat{\varrho}_r(\mathbf{p}, \omega, \hat{t}) < \varrho_r(\mathbf{p}), \quad \forall t \in [\hat{t}, T], \quad (76)$$

where

$$\hat{\varrho}_r(\mathbf{p}, \boldsymbol{\omega}, \hat{t}) = \left(e^{-2\lambda\hat{t}} \|X_0^r - X_0^*\|_{L_{2p}}^2 + \frac{\Delta_g^\perp \Delta_\mu^\perp}{\lambda} \varrho_r(\mathbf{p})^2 + \Gamma_r(\varrho_r(\mathbf{p}), \mathbf{p}, \boldsymbol{\omega}) + \frac{1}{|2\lambda - \boldsymbol{\omega}|} \Theta_r(\varrho_r(\mathbf{p}), \mathbf{p}, \boldsymbol{\omega}) \right)^{\frac{1}{2}}, \quad (77)$$

and where the pertinent definitions are provided in Sec 3.2.

Proof. Owing to Lemma 4.3, the results of Lemma 4.1 hold for the stopping time $T(t) \doteq t \wedge T$. That is, for any deterministic $\hat{t} \in (0, T)$, the inequalities (35) and (36) can be written as

$$\begin{aligned} \mathbb{E}_{y_0} [\mathcal{V}(Y_{T(t)})] &\leq e^{-2\lambda\hat{t}} \mathcal{V}(y_0) + \mathbb{E}_{y_0} \left[e^{-2\lambda T(t)} \Xi^r(T(t), Y) \right] \\ &\quad + \mathbb{E}_{y_0} \left[e^{-(2\lambda + \boldsymbol{\omega})T(t)} \Xi_{\mathcal{U}}^r(T(t), Y; \boldsymbol{\omega}) \right], \quad \forall t \in [\hat{t}, \infty), \end{aligned} \quad (78)$$

and

$$\begin{aligned} \|\mathcal{V}(Y_{T(t)})\|_{L_p}^{y_0} &\leq e^{-2\lambda\hat{t}} \mathcal{V}(y_0) + \left\| e^{-2\lambda T(t)} \Xi^r(T(t), Y) \right\|_{L_p}^{y_0} \\ &\quad + \left\| e^{-(2\lambda + \boldsymbol{\omega})T(t)} \Xi_{\mathcal{U}}^r(T(t), Y; \boldsymbol{\omega}) \right\|_{L_p}^{y_0}, \quad \forall (t, \mathbf{p}) \in [\hat{t}, \infty) \times \mathbb{N}_{\geq 2}, \end{aligned} \quad (79)$$

where we have further used Lemma 4.3 to replace the truncated joint process $Y_{N,t} = (X_{N,t}^r, X_{N,t}^*)$ with the unique strong solution of the joint process $Y_t = (X_t^r, X_t^*)$. Next, as in the proof of Lemma 4.2, by integrating (78) and (79) with respect to the joint distribution π_0 of the initial state Y_0 , we obtain the following bounds that are analogous to (52) and (54):

$$\begin{aligned} \mathbb{E} [\mathcal{V}(Y_t)] &\leq \alpha_2 e^{-2\lambda\hat{t}} \mathbb{E} [\|X_0^r - X_0^*\|^2] + \int_{\mathbb{R}^{2n}} \mathbb{E}_{y_0} \left[e^{-2\lambda T(t)} \Xi^r(T(t), Y) \right] \pi_0(dy_0) \\ &\quad + \int_{\mathbb{R}^{2n}} \mathbb{E}_{y_0} \left[e^{-(2\lambda + \boldsymbol{\omega})T(t)} \Xi_{\mathcal{U}}^r(T(t), Y; \boldsymbol{\omega}) \right] \pi_0(dy_0), \quad \forall t \in [\hat{t}, T], \end{aligned} \quad (80)$$

and

$$\begin{aligned} \|\mathcal{V}(Y_t)\|_{L_p} &\leq \alpha_2 e^{-2\lambda\hat{t}} \|X_0^r - X_0^*\|_{L_{2p}}^2 \\ &\quad + \left(\int_{\mathbb{R}^{2n}} \left(\left\| e^{-2\lambda T(t)} \Xi^r(T(t), Y) \right\|_{L_p}^{y_0} + \left\| e^{-(2\lambda + \boldsymbol{\omega})T(t)} \Xi_{\mathcal{U}}^r(T(t), Y; \boldsymbol{\omega}) \right\|_{L_p}^{y_0} \right)^p \pi_0(dy_0) \right)^{\frac{1}{p}}, \end{aligned} \quad (81)$$

for all $(t, \mathbf{p}) \in [\hat{t}, T] \times \{2, \dots, p^*\}$. Next, for the particular $\mathbf{p} \in \{1, \dots, p^*\}$ used in the definitions and conditions in Section 3.2, we have the following from Lemma 4.3 and Assumption 3:

$$\sup_{\nu \in [\hat{t}, T]} \|X_\nu^r - X_\nu^*\|_{L_{2p}} \leq \varrho_r(\mathbf{p}), \quad \sup_{\nu \in [\hat{t}, T]} \|X_\nu^r - X_\nu^*\|_{L_{2p}} + \sup_{\nu \in [\hat{t}, T]} \|X_\nu^*\|_{L_{2p}} \leq \varrho_r(\mathbf{p}) + \Delta_*. \quad (82)$$

Using (80), (81), and (82), we following the analysis in the proof of Lemma 4.2 to obtain the following that is analogous to (68)

$$\begin{aligned} &\|X_t^r - X_t^*\|_{L_{2p}}^2 \\ &\leq e^{-2\lambda\hat{t}} \|X_0^r - X_0^*\|_{L_{2p}}^2 + \frac{\Delta_g^\perp \Delta_\mu^\perp}{\lambda} \varrho_r(\mathbf{p})^2 + \Gamma_r(\varrho_r(\mathbf{p}), \mathbf{p}, \boldsymbol{\omega}) + \frac{1}{|2\lambda - \boldsymbol{\omega}|} \Theta_r(\varrho_r(\mathbf{p}), \mathbf{p}, \boldsymbol{\omega}) \doteq \hat{\varrho}_r(\mathbf{p}, \boldsymbol{\omega}, \hat{t})^2, \end{aligned} \quad (83)$$

for all $t \in [\hat{t}, T]$. Using the filter bandwidth condition in (21a) with $\alpha(X_0^r, X_0^*)_{2p} = \|X_0^r - X_0^*\|_{L_{2p}}$, we see that

$$\begin{aligned} \hat{\varrho}_r(\mathbf{p}, \boldsymbol{\omega}, \hat{t})^2 &= e^{-2\lambda\hat{t}} \|X_0^r - X_0^*\|_{L_{2p}}^2 + \frac{\Delta_g^\perp \Delta_\mu^\perp}{\lambda} \varrho_r(\mathbf{p})^2 + \Gamma_r(\varrho_r(\mathbf{p}), \mathbf{p}, \boldsymbol{\omega}) + \frac{1}{|2\lambda - \boldsymbol{\omega}|} \Theta_r(\varrho_r(\mathbf{p}), \mathbf{p}, \boldsymbol{\omega}) \\ &< \varrho_r(\mathbf{p})^2 - (1 - e^{-2\lambda\hat{t}}) \|X_0^r - X_0^*\|_{L_{2p}}^2 \end{aligned}$$

$$< \varrho_r(\mathbf{p})^2,$$

for all $\hat{t} \in (0, T)$. Coupled to the fact in (20) that $\varrho_r(\mathbf{p}) > \|X_0^r - X_0^*\|_{L_{2\mathbf{p}}}$, we conclude that

$$\hat{\varrho}_r(\mathbf{p}, \boldsymbol{\omega}, \hat{t}) \in (0, \varrho_r(\mathbf{p})), \quad \forall \hat{t} \in (0, T),$$

which, along with (83) concludes the proof. \square

We now state the main result of the section.

Lemma 4.4 *Let the assumptions in Sec. 2.2 hold and let $\varrho_r(\mathbf{p}) \in \mathbb{R}_{>0}$, $\mathbf{p} \in \{1, \dots, \mathbf{p}^*\}$, be as defined in (16a) with $\alpha(X_0^r, X_0^*)_{2\mathbf{p}} = W_{2\mathbf{p}}(\xi_0, \xi_0^*)$, where \mathbf{p}^* is defined in Assumption 3. If the filter bandwidth condition in (21a) holds with $\alpha(X_0^r, X_0^*)_{2\mathbf{p}} = W_{2\mathbf{p}}(\xi_0, \xi_0^*)$, then*

$$W_{2\mathbf{p}}(X_t^r, X_t^*) < \varrho_r(\mathbf{p}), \quad \forall t \in [0, T], \quad (\text{Distributional UB}) \quad (84a)$$

$$W_{2\mathbf{p}}(X_t^r, X_t^*) \leq \hat{\varrho}_r(\mathbf{p}, \boldsymbol{\omega}, \hat{t}) < \varrho_r(\mathbf{p}), \quad \forall t \in [\hat{t}, T], \quad (\text{Distributional UUB}) \quad (84b)$$

where $\hat{t} \in (0, T)$ is any deterministic constant, and

$$\hat{\varrho}_r(\mathbf{p}, \boldsymbol{\omega}, \hat{t}) = \left(e^{-2\lambda\hat{t}} W_{2\mathbf{p}}(\xi_0, \xi_0^*)^2 + \frac{\Delta_g^\perp \Delta_\mu^\perp}{\lambda} \varrho_r(\mathbf{p})^2 + \Gamma_r(\varrho_r(\mathbf{p}), \mathbf{p}, \boldsymbol{\omega}) + \frac{1}{|2\lambda - \boldsymbol{\omega}|} \Theta_r(\varrho_r(\mathbf{p}), \mathbf{p}, \boldsymbol{\omega}) \right)^{\frac{1}{2}}, \quad (85)$$

and where the pertinent definitions are provided in Sec 3.2.

Furthermore, if the hypotheses stated above hold for $\alpha(X_0^r, X_0^*)_{2\mathbf{p}} = \|X_0^r - X_0^*\|_{L_{2\mathbf{p}}}$, then for any $\delta \in (0, 1)$ such that $\log \sqrt{1/\delta} = \mathbf{p}$,

$$\mathbb{P}(\|X_t^r - X_t^*\| \geq \kappa_r(\delta)) < \delta, \quad \forall t \in [0, T], \quad (\text{Pathwise UB}) \quad (86a)$$

$$\mathbb{P}(\|X_t^r - X_t^*\| \geq \hat{\kappa}_r(\delta, \boldsymbol{\omega}, \hat{t})) < \delta, \quad \forall t \in [\hat{t}, T], \quad (\text{Pathwise UUB}) \quad (86b)$$

where $\kappa_r(\delta) = e \cdot \varrho_r(\log \sqrt{1/\delta})$ and $\hat{\kappa}_r(\delta, \boldsymbol{\omega}, \hat{t}) = e \cdot \hat{\varrho}_r(\log \sqrt{1/\delta}, \boldsymbol{\omega}, \hat{t})$.

Proof. Recall that we proved Lemma 4.2 for moments with respect to the law $\mathbb{Y}_N(\cdot, t) : \mathcal{B}(\mathbb{R}^{2n}) \rightarrow [0, 1]$ that resulted under the arbitrary initial coupling $Y_0 = Y_{N,0} \sim \pi_0$ (see (48)). In order to prove (84), we will therefore appeal to the proof of Lemma 4.2 but under the law that results from the joint process initialized under the *optimal* coupling. Let us denote by π_0^* the optimal coupling between the probability measures ξ_0 and ξ_0^* that are the respective distributions of the initial states x_0 and x_0^* [29, Chp. 1]. From Lemma 4.3, the unique strong solution Y_t of the joint process exists, and is adapted to the filtration generated by Y_0 and \bar{W}_t [15, Thm. 5.2.1] (initial vector Y_0 is independent of the Brownian motion \bar{W}_t). Let us denote by $\mathbb{Y}_{\pi_0^*}(\cdot, t)$ the law of the process Y_t under $Y_0 \sim \pi_0^*$. Hence, $\mathbb{Y}_{\pi_0^*}(\cdot, t)$ is a probability measure on \mathbb{R}^{2n} given as

$$\mathbb{Y}_{\pi_0^*}(B, t) = \mathbb{P}(Y_t \in B) = \int_{\mathbb{R}^{2n}} \mathbb{Y}(y_0, 0; B, t) \pi_0^*(dy_0), \quad \forall B \in \mathcal{B}(\mathbb{R}^{2n}), \quad (87)$$

where $\mathbb{Y}(y_0, 0; B, t)$ denotes the transition probability for the process Y_t , initialized from y_0 [48, Sec. 3.1]. Then, for $\mathcal{V}(Y_t) = \|X_t^r - X_t^*\|^2$ we can define

$$\begin{aligned} \mathbb{E}_{\pi_0^*}[\mathcal{V}(Y_t)^{\mathbf{p}}] &= \int_{\mathbb{R}^{2n}} \mathcal{V}(y)^{\mathbf{p}} \mathbb{Y}_{\pi_0^*}(dy, t) = \int_{\mathbb{R}^{2n}} \int_{\mathbb{R}^{2n}} \mathcal{V}(y)^{\mathbf{p}} \mathbb{Y}(y_0, 0; dy, t) \pi_0^*(dy_0) \\ &= \int_{\mathbb{R}^{2n}} \mathbb{E}[\mathcal{V}(Y_t)^{\mathbf{p}} | Y_0 = y_0] \pi_0^*(dy_0) = \int_{\mathbb{R}^{2n}} \mathbb{E}_{y_0}[\mathcal{V}(Y_t)^{\mathbf{p}}] \pi_0^*(dy_0), \end{aligned} \quad (88)$$

for $t \in [0, T]$. Note that if we were to replace the optimal initial coupling π_0^* with an arbitrary coupling π_0 , then (88) yields (49) and (50). Since π_0^* is the optimal coupling between the initial measures ξ_0 and ξ_0^* , it follows from the definition of the Wasserstein distance that

$$\int_{\mathbb{R}^{2n}} \mathcal{V}(y_0)^{\mathbf{p}} \pi_0^*(dy_0) = \int_{\mathbb{R}^{2n}} \|x_0 - x_0^*\|^{2\mathbf{p}} \pi_0^*(dx_0, dx_0^*) = W_{2\mathbf{p}}(\xi_0, \xi_0^*)^{2\mathbf{p}}.$$

We can thus follow the analysis in Lemma 4.2 for the law $\mathbb{Y}_{\pi_0^*}(\cdot, t)$ of the joint process Y_t up to the stopping time T and obtain

$$\|X_t^r - X_t^*\|_{L_{2p}}^{\pi_0^*} = \mathbb{E}_{\pi_0^*} \left[\|X_t^r - X_t^*\|^{2p} \right]^{\frac{1}{2p}} < \varrho_r(p), \quad \forall t \in [0, T], \quad (89)$$

where $\varrho_r(p)$ is defined in (16a) with $\alpha(X_0^r, X_0^*)_{2p} = W_{2p}(\xi_0, \xi_0^*)$. We now follow the approach in [109] and, using (87), define the measure $\pi_t = \mathbb{Y}_{\pi_0^*}(\cdot, t)$. It then follows from (88)

$$\mathbb{E}_{\pi_0^*} \left[\|X_t^r - X_t^*\|^{2p} \right] = \mathbb{E}_{\pi_0^*} [\mathcal{V}(Y_t)^p] = \int_{\mathbb{R}^{2n}} \mathcal{V}(y)^p \mathbb{Y}_{\pi_0^*}(dy, t) = \int_{\mathbb{R}^{2n}} \mathcal{V}(y)^p \pi_t(dy).$$

Since π_t is not necessarily the optimal coupling between the state distributions \mathbb{X}_t^r and \mathbb{X}_t^* , we reason that

$$W_{2p}(\mathbb{X}_t^r, \mathbb{X}_t^*)^{2p} = \int_{\mathbb{R}^{2n}} \mathcal{V}(y)^p \pi_t^*(dy) \leq \int_{\mathbb{R}^{2n}} \mathcal{V}(y)^p \pi_t(dy) = \mathbb{E}_{\pi_0^*} \left[\|X_t^r - X_t^*\|^{2p} \right], \quad \forall t \in [0, T],$$

where π_t^* denotes the optimal coupling between the state distributions \mathbb{X}_t^r and \mathbb{X}_t^* . Substituting the above into (89) then establishes (84a). The bound in (84b) follows *mutatis mutandis* from the analysis above and the proof of Proposition 4.2.

Next, for any $\kappa \in \mathbb{R}_{>0}$, we can use the Markov's inequality [108, Cor. 3.1.2] and Lemma 4.3 to obtain

$$\mathbb{P}(\|X_t^r - X_t^*\| \geq \kappa) \leq \frac{\mathbb{E} \left[\|X_t^r - X_t^*\|^{2p} \right]}{\kappa^{2p}} < \left(\frac{\varrho_r(p)}{\kappa} \right)^{2p}, \quad \forall (t, p) \in [0, T] \times [1, p^*].$$

By setting $p = \log \sqrt{\frac{1}{\delta}}$ and $\kappa = e \cdot \varrho_r(p)$, we obtain

$$\mathbb{P}(\|X_t^r - X_t^*\| \geq \kappa_r(\delta)) < \delta, \quad \forall t \in [0, T],$$

thus establishing (86a). The proof of (86b) follows *mutatis mutandis* by using the uniform ultimate bound from Proposition 4.2. □

4.2 Performance Analysis: True (Uncertain) Process

We now consider the true (uncertain) system in (4a) operating under the \mathcal{L}_1 -DRAC feedback law $\mathcal{F}_{\mathcal{L}_1}$ defined in (13), Sec. 3.1.

Definition 7 (True (Uncertain) \mathcal{L}_1 -DRAC Closed-loop Process) We say that X_t , $t \in [0, T]$, for any $T \in (0, \infty)$, is the **true (uncertain) \mathcal{L}_1 -DRAC process**, if X_t is a unique strong solution to the true (uncertain) Itô SDE in (4a) under the \mathcal{L}_1 -DRAC feedback law $\mathcal{F}_{\mathcal{L}_1}$ in (13):

$$dX_t = F_\mu(t, X_t, U_{\mathcal{L}_1, t}) dt + F_\sigma(t, X_t) dW_t, \quad X_0 = x_0 \sim \xi_0 \text{ (}\mathbb{P}\text{-a.s.)}, \quad (90)$$

where

$$U_{\mathcal{L}_1, t} \doteq \mathcal{F}_{\mathcal{L}_1}(X)(t). \quad (91)$$

The \mathcal{L}_1 -DRAC feedback operator $\mathcal{F}_{\mathcal{L}_1} : \mathcal{C}([0, T] : \mathbb{R}^n) \rightarrow \mathcal{C}([0, T] : \mathbb{R}^m)$ is defined in (13), which we restate next. Using the definition in (13)

$$\mathcal{F}_{\mathcal{L}_1}(X) = \mathcal{F}_\omega \circ \mathcal{F}_{T_s} \circ \mathcal{F}_{\lambda_s}(X), \quad (92)$$

we write

$$U_{\mathcal{L}_1} = \mathcal{F}_{\mathcal{L}_1}(X) = \mathcal{F}_\omega(\hat{\Lambda}^\parallel), \quad \hat{\Lambda}^\parallel = \mathcal{F}_{T_s}^\parallel(\hat{\Lambda}), \quad \hat{\Lambda} = \mathcal{F}_{T_s}(\hat{X}, X), \quad \hat{X} = \mathcal{F}_{\lambda_s}(X). \quad (93)$$

Using (15a), we see that the input $U_{\mathcal{L}_1, t}$ is defined as the output of the **low-pass filter**:

$$U_{\mathcal{L}_1, t} = \mathcal{F}_\omega(\hat{\Lambda}^\parallel)(t) = -\omega \int_0^t e^{-\omega(t-\nu)} \hat{\Lambda}^\parallel(\nu) d\nu, \quad (94)$$

where $\omega \in \mathbb{R}_{>0}$ is the **filter bandwidth**. The **adaptive estimates** $\hat{\Lambda}^\parallel$ and $\hat{\Lambda}$ are obtained via the **adaptation law** operator $\mathcal{F}_{T_s}(\mathcal{F}_{T_s}^\parallel)$ in (15b) as follows:

$$\begin{aligned}\hat{\Lambda}^\parallel(t) &= \mathcal{F}_{T_s}^\parallel(\hat{\Lambda})(t) = \sum_{i=0}^{\lfloor \frac{t}{T_s} \rfloor} \Theta_{ad}(iT_s) \hat{\Lambda}(t) \mathbb{1}_{[iT_s, (i+1)T_s)}(t), \\ \hat{\Lambda}(t) &= \mathcal{F}_{T_s}(\hat{X}_t, X)(t) \\ &= 0_n \mathbb{1}_{[0, T_s)}(t) + \lambda_s (1 - e^{\lambda_s T_s})^{-1} \sum_{i=1}^{\lfloor \frac{t}{T_s} \rfloor} \tilde{X}_{iT_s} \mathbb{1}_{[iT_s, (i+1)T_s)}(t), \quad \tilde{X}_{iT_s} \doteq \hat{X}_{iT_s} - X_{iT_s},\end{aligned}\tag{95}$$

where $T_s \in \mathbb{R}_{>0}$ is the **sampling period** and $\Theta_{ad}(t) = [\mathbb{I}_m \quad 0_{m, n-m}] \bar{g}(t)^{-1} \in \mathbb{R}^{m \times n}$, with $\bar{g}(t) = [g(t) \quad g(t)^\perp] \in \mathbb{R}^{n \times n}$ for g^\perp defined in Assumption 4. The parameter $\lambda_s \in \mathbb{R}_{>0}$ contributes to the **prediction process** \hat{X}_t as a parameter to the operator \mathcal{F}_{λ_s} in (15c) which induces the **process predictor** as follows:

$$\begin{aligned}\hat{X}_t &= \mathcal{F}_{\lambda_s}(X)(t) \\ \Rightarrow \hat{X}_t &= x_0 + \int_0^t \left(-\lambda_s \mathbb{I}_n \tilde{X}_\nu + f(\nu, X_\nu) + g(\nu) \mathcal{F}_{\mathcal{L}_1}(X)(\nu) + \hat{\Lambda}(\nu) \right) d\nu, \quad \tilde{X}_t \doteq \hat{X}_t - X_t.\end{aligned}\tag{96}$$

We collectively refer to $\{\omega, T_s, \lambda_s\}$ as the control parameters.

Finally, in Definition 2, \mathbb{X}_t denotes the **true law** induced by the process X_t on $\mathcal{B}(\mathbb{R}^n)$, i.e., $X_t \sim \mathbb{X}_t$.

Our goal in this section is to establish the performance of the \mathcal{L}_1 -DRAC closed-loop true (uncertain) process in (90) relative to the reference process in Definition 4. For this purpose, we provide the following:

Definition 8 (True (Uncertain)-Reference Error Process) Consider the true process X_t and the reference process $X_t^r \sim \mathbb{X}_t^r$ in Definitions 7 and 4, respectively. We say that $Z_t, t \in [0, T]$, for any $T \in (0, \infty)$, is the **true (uncertain)-reference joint process**, if it is a unique strong solution of the following **true (uncertain)-reference joint Itô SDE** on $(\Omega, \mathcal{F}, \mathbb{W}_{0,t}, \mathbb{P})$:

$$dZ_t = J_\mu(t, Z_t) dt + J_\sigma(t, Z_t) dW_t, \quad t \in [0, T], \quad Z_0 = 0, \quad Z_t \sim \mathbb{Z}_t,\tag{97}$$

where

$$\begin{aligned}Z_t &\doteq X_t - X_t^r \in \mathbb{R}^n, \quad J_\mu(t, Z_t) \doteq F_\mu(t, X_t, U_{\mathcal{L}_1, t}) - F_\mu(t, X_t^r, U_t^r) \in \mathbb{R}^n, \\ J_\sigma(t, Z_t) &\doteq F_\sigma(t, X_t) - F_\sigma(t, X_t^r) \in \mathbb{R}^{n \times d}.\end{aligned}$$

Remark 4.2 The definition of the joint process above between the true (uncertain) process and the reference process is **not** a direct analogue to the joint known (nominal)-reference process in Definition 5. The reason is that the true (uncertain) process and the reference process are driven by the identical Brownian motion W_t and share the same initial condition $x_0 \sim \xi_0$. This is in contrast to the joint known (nominal)-reference process in Definition 5, where the known (nominal) process and the reference process are driven by **independent Brownian motions** W_t^* and W_t and have **distinct initial data** $x_0 \sim \xi_0$ and $x_0^* \sim \xi_0^*$, respectively. Recall from our discussion after Definition 4 that the reference process is **not-realizable** and represents the best achievable performance, and thus we designed the reference process to be driven by the identical Brownian motion and share the same initial condition as the true (uncertain) process.

Similar to the truncated joint known (nominal)-reference process in Definition 6, we require the following:

Definition 9 (Truncated Joint True (Uncertain)-Reference Process) Let

$$\tilde{U}_N \doteq \{a \in \mathbb{R}^n : \|a\| < N\} \subset \subset \mathbb{R}^n, \quad \forall N \in \mathbb{R}_{>0}.\tag{98}$$

Next, we define the **truncated joint true (uncertain)-reference Itô SDE** as

$$dZ_{N,t} = J_{N,\mu}(t, Z_{N,t}) dt + J_{N,\sigma}(t, Z_{N,t}) dW_t, \quad Z_{N,0} = Z_0 = 0,\tag{99}$$

where, the process and the drift and diffusion vector fields are defined as

$$Z_{N,t} \doteq X_{N,t} - X_{N,t}^r \in \mathbb{R}^n, \quad J_{N,\mu}(t, a) (J_{N,\sigma}(t, a)) = \begin{cases} J_\mu(t, a) (J_\sigma(t, a)), & \|a\| \leq N \\ 0_n (0_{n,d}), & \|a\| \geq 2N \end{cases},$$

for all $(a, t) \in \mathbb{R}^n \times [0, T]$, where $J_{N,\mu}(t, a)$ and $J_{N,\sigma}(t, a)$ are any **uniformly Lipschitz continuous maps** for all $a \in \mathbb{R}^n$ and $t \in \mathbb{R}_{\geq 0}$. Similar to Z_t , we refer to $Z_{N,t} \in \mathbb{R}^n$ as the **truncated joint known(nominal)-reference process** if it is a unique strong solution of (99).

We begin by establishing the uniqueness and existence of strong solutions for the truncated joint true (uncertain)-reference process.

Proposition 4.3 (Well-Posedness of (99)) *If Assumptions 1 and 4 hold, then for any $N \in \mathbb{R}_{>0}$, $Z_{N,t}$ is a unique strong solution of (99), $\forall t \in [0, T]$, for any $T \in (0, \infty)$ and is a strong Markov process $\forall t \in \mathbb{R}_{\geq 0}$. Furthermore, define⁶*

$$\tau_N \doteq T \wedge \inf \left\{ t \in [0, T] : Z_{N,t} \notin \tilde{U}_N \right\}, \quad (100)$$

where the open and bounded set \tilde{U}_N is defined in (98), for an arbitrary $N \in \mathbb{R}_{>0}$. Then, $Z_{N,t}$ uniquely solves (97), in the strong sense, for all $t \in [0, \tau_N]$.

Proof. See Appendix D. □

Remark 4.3 Recall that τ_N in (32), Proposition 4.1, denotes the first exit time of the process $Y_{N,t}$ from the set $U_N \subset \mathbb{R}^{2n}$. Similarly, with an abuse of notation and from this point onward, τ_N in (100) denotes the first exit time of the process $Z_{N,t}$ from the set $\tilde{U}_N \subset \mathbb{R}^n$.

Using the well-posedness of the truncated process, we can now define the following transition probability:

$$\mathbb{Z}_N(Z_{N,s}, s; B, t) \doteq \mathbb{P}(Z_{N,t} \in B \mid Z_{N,s}), \quad \forall 0 < s \leq t \leq T, \quad B \in \mathcal{B}(\mathbb{R}^n). \quad (101)$$

With an abuse of notation, let us also denote the law of the process $Z_{N,t}$ by \mathbb{Z}_N . Thus, $\mathbb{Z}_N(\cdot, t)$ is a probability measure on \mathbb{R}^n given as

$$\mathbb{Z}_N(B, t) = \mathbb{P}(Z_{N,t} \in B) = \int_{\mathbb{R}^n} \mathbb{Z}_N(z_0, 0; B, t) \mathbb{P}(dz_0), \quad \forall B \in \mathcal{B}(\mathbb{R}^n),$$

where $\mathbb{P}(z_0)$ denotes the distribution of the initial condition $Z_{N,0}$. Hence, for any Borel measurable function $a : \mathbb{R}^n \rightarrow \mathbb{R}$, we have

$$\begin{aligned} \mathbb{E}[a(Z_{N,t})] &= \int_{\mathbb{R}^n} a(z) \mathbb{Z}_N(dz, t) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} a(z) \mathbb{Z}_N(z_0, 0; dz, t) \mathbb{P}(dz_0) \\ &= \int_{\mathbb{R}^n} \mathbb{E}[a(Z_{N,t}) \mid Z_{N,0} = z_0] \mathbb{P}(dz_0). \end{aligned}$$

However, since $\mathbb{P}(Z_{N,0} = 0) = 1$, the previous expression can be written as

$$\begin{aligned} \mathbb{E}[a(Z_{N,t})] &= \int_{\mathbb{R}^n} \mathbb{E}[a(Z_{N,t}) \mid Z_{N,0} = z_0] \delta_0(dz_0) = \mathbb{E}[a(Z_{N,t}) \mid Z_{N,0} = 0] \\ &= \int_{\mathbb{R}^n} a(z) \mathbb{Z}_N(0, 0; dz, t), \end{aligned} \quad (102)$$

where δ_0 is the Dirac measure supported on $0 \in \mathbb{R}^n$. Hence, the following analysis uses the total expectation as it is identical to the expectation conditioned on the zero initial condition.

Next, we study the *uniform in time* bounds on the moments of the truncated joint true (uncertain)-reference process. In order to streamline the derivations in this section, we define the following:

$$\mathring{f}(\nu, Z_{N,\nu}) = f(\nu, X_{N,\nu}) - f(\nu, X_{N,\nu}^r), \quad (103a)$$

$$\mathring{\Lambda}_{\mu}^{\{\cdot, \parallel, \perp\}}(\nu, Z_{N,\nu}) = \Lambda_{\mu}^{\{\cdot, \parallel, \perp\}}(\nu, X_{N,\nu}) - \Lambda_{\mu}^{\{\cdot, \parallel, \perp\}}(\nu, X_{N,\nu}^r), \quad (103b)$$

$$\mathring{F}_{\sigma}^{\{\cdot, \parallel, \perp\}}(\nu, Z_{N,\nu}) = F_{\sigma}^{\{\cdot, \parallel, \perp\}}(\nu, X_{N,\nu}) - F_{\sigma}^{\{\cdot, \parallel, \perp\}}(\nu, X_{N,\nu}^r). \quad (103c)$$

$$\mathring{\mathcal{F}}_r(Z_N)(\nu) = \mathcal{F}_r(X_N)(\nu) - \mathcal{F}_r(X_N^r)(\nu). \quad (103d)$$

Lemma 4.5 *Let the assumptions in Sec. 2.2 hold. For an arbitrary $N \in \mathbb{R}_{>0}$, let the stopping time τ_N be as in (32) for the truncated joint process in Definition 6. For any constant $t^* \in \mathbb{R}_{>0}$ define⁷*

$$\tau^* = t^* \wedge \tau_N, \quad \tau(t) = t \wedge \tau^*. \quad (104)$$

⁶The stopping time τ_N takes on the value of either the finite horizon T , or the first exit time of $Z_{N,t}$ from set \tilde{U}_N .

⁷The stopping time τ^* is either the constant t^* , the finite horizon T , or the first exit time of $Z_{N,t}$ from the set \tilde{U}_N . The variable $\tau(t)$ is identical to t up to τ^* , and zero afterwards.

Then, with $\mathcal{V}(Z_{N,t}) \doteq \|Z_{N,t}\|^2 = \|X_{N,t} - X_{N,t}^r\|^2$, the truncated joint process $Z_{N,t}$ in Definition 9 satisfies:

$$\begin{aligned} \mathbb{E}[\mathcal{V}(Z_{N,\tau(t)})] &\leq \mathbb{E}\left[e^{-2\lambda\tau(t)}\Xi(\tau(t), Z_N)\right] + \mathbb{E}\left[e^{-(2\lambda+\omega)\tau(t)}\Xi_{\mathcal{U}}(\tau(t), Z_N)\right] \\ &\quad + \mathbb{E}\left[e^{-2\lambda\tau(t)}\tilde{\Xi}_{\mathcal{U}}(\tau(t), Z_N)\right], \quad \forall t \in \mathbb{R}_{\geq 0}, \end{aligned} \quad (105)$$

and

$$\begin{aligned} \|\mathcal{V}(Z_{N,\tau(t)})\|_{L_p} &\leq \left\|e^{-2\lambda\tau(t)}\Xi(\tau(t), Z_N)\right\|_{L_p} + \left\|e^{-(2\lambda+\omega)\tau(t)}\Xi_{\mathcal{U}}(\tau(t), Z_N)\right\|_{L_p} \\ &\quad + \left\|e^{-2\lambda\tau(t)}\tilde{\Xi}_{\mathcal{U}}(\tau(t), Z_N)\right\|_{L_p}, \quad \forall (t, p) \in \mathbb{R}_{\geq 0} \times \mathbb{N}_{\geq 2}, \end{aligned} \quad (106)$$

where, for any Borel measurable function $a : \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbb{E}[a(Z_{N,t})]$ is defined in (102), and $\|a(t)\|_{L_p}$ follows analogously. The above bounds are defined using the following:

$$\Xi(\tau(t), Z_N) = \int_0^{\tau(t)} e^{2\lambda\nu} (\phi_\mu(\nu, Z_{N,\nu}) d\nu + \phi_\sigma(\nu, Z_{N,\nu}) dW_\nu), \quad (107a)$$

$$\Xi_{\mathcal{U}}(\tau(t), Z_N) = \int_0^{\tau(t)} (\mathcal{U}_\mu(\tau(t), \nu, Z_N; \omega) d\nu + \mathcal{U}_\sigma(\tau(t), \nu, Z_N; \omega) dW_\nu), \quad (107b)$$

$$\tilde{\Xi}_{\mathcal{U}}(\tau(t), Z_N) = \int_0^{\tau(t)} e^{2\lambda\nu} \tilde{\mathcal{U}}(\nu, Z_{N,\nu}) d\nu, \quad (107c)$$

where

$$\phi_\mu(\nu, Z_{N,\nu}) = \partial\mathcal{V}(Z_{N,\nu})^\top g(\nu)^\perp \mathring{\Lambda}_\mu^\perp(\nu, Z_{N,\nu}) + \left\|\mathring{F}_\sigma^\perp(\nu, Z_{N,\nu})\right\|_F^2, \quad (108a)$$

$$\phi_\sigma(\nu, Z_{N,\nu}) = \partial\mathcal{V}(Z_{N,\nu})^\top g(\nu)^\perp \mathring{F}_\sigma^\perp(\nu, Z_{N,\nu}), \quad (108b)$$

$$\tilde{\mathcal{U}}(\nu, Z_{N,\nu}) = \partial\mathcal{V}(Z_{N,\nu})^\top g(\nu) (\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r)(X_N)(\nu), \quad (108c)$$

and

$$\mathcal{U}_\mu(\tau(t), \nu, Z_N; \omega) = \psi(\tau(t), \nu, Z_N) \mathring{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}), \quad \mathcal{U}_\sigma(\tau(t), \nu, Z_N; \omega) = \psi(\tau(t), \nu, Z_N) \mathring{F}_\sigma^\parallel(\nu, Z_{N,\nu}). \quad (109)$$

In the above, along with $\partial\mathcal{V}(Z_{N,\nu}) = 2(X_{N,\nu} - X_{N,\nu}^r)$ and the definitions in (103), we have set

$$\begin{aligned} \psi(\tau(t), \nu, Z_N) &= \frac{\omega}{2\lambda - \omega} \left(e^{\omega(\tau(t)+\nu)} \mathcal{P}(\tau(t), \nu) - e^{(2\lambda\tau(t)+\omega\nu)} \partial\mathcal{V}(Z_{N,\tau(t)})^\top g(\tau(t)) \right) \\ &\quad + \frac{2\lambda}{2\lambda - \omega} e^{(\omega\tau(t)+2\lambda\nu)} \partial\mathcal{V}(Z_{N,\nu})^\top g(\nu) \in \mathbb{R}^{1 \times m}, \end{aligned} \quad (110)$$

and

$$\mathcal{P}(\tau(t), \nu) = \int_\nu^{\tau(t)} e^{(2\lambda-\omega)\beta} d_\beta \left[\partial\mathcal{V}(Z_{N,\beta})^\top g(\beta) \right] \in \mathbb{R}^{1 \times m}, \quad 0 \leq \nu \leq \tau(t), \quad (111)$$

where $d_\beta[\cdot]$ denotes the stochastic differential with respect to β .

Proof. Since $Z_{N,0} = 0$, we have that $\mathbb{P}(Z_{N,0} \in \tilde{U}_N) = 1$. We therefore apply the Itô lemma [91, Thm. 1.6.4] to $e^{2\lambda\tau(t)}\mathcal{V}(Z_{N,t})$, using the dynamics in (99), to obtain

$$\begin{aligned} e^{2\lambda\tau(t)}\mathcal{V}(Z_{N,\tau(t)}) &= 2\lambda \int_0^{\tau(t)} e^{2\lambda\nu} \mathcal{V}(Z_{N,\nu}) d\nu + \int_0^{\tau(t)} e^{2\lambda\nu} \nabla\mathcal{V}(Z_{N,\nu})^\top J_{N,\sigma}(\nu, Z_{N,\nu}) dW_\nu \\ &\quad + \int_0^{\tau(t)} e^{2\lambda\nu} \left(\nabla\mathcal{V}(Z_{N,\nu})^\top J_{N,\mu}(\nu, Z_{N,\nu}) + \frac{1}{2} \text{Tr}[K_{N,\sigma}(\nu, Z_{N,\nu}) \nabla^2\mathcal{V}(Z_{N,\nu})] \right) d\nu, \end{aligned}$$

for all $t \in \mathbb{R}_{\geq 0}$, where $K_{N,\sigma}(\nu, Y_{N,\nu}) = J_{N,\sigma}(\nu, Y_{N,\nu}) J_{N,\sigma}(\nu, Y_{N,\nu})^\top \in \mathbb{S}^n$. Next, from Proposition 4.3, $Z_{N,t}$ is also unique strong solution of (97), for all $t \in [0, \tau^*]$ because $[0, \tau^*] \subseteq [0, \tau_N]$. Therefore, we may replace $J_{N,\mu}$, $J_{N,\sigma}$, and $K_{N,\sigma}$ with J_μ , J_σ , and K_σ , respectively, in the above inequality and obtain

$$e^{2\lambda\tau(t)}\mathcal{V}(Z_{N,\tau(t)}) = 2\lambda \int_0^{\tau(t)} e^{2\lambda\nu} \mathcal{V}(Z_{N,\nu}) d\nu + \int_0^{\tau(t)} e^{2\lambda\nu} \nabla \mathcal{V}(Z_{N,\nu})^\top J_\sigma(\nu, Z_{N,\nu}) dW_\nu \\ + \int_0^{\tau(t)} e^{2\lambda\nu} \left(\nabla \mathcal{V}(Z_{N,\nu})^\top J_\mu(\nu, Z_{N,\nu}) + \frac{1}{2} \text{Tr}[K_\sigma(\nu, Z_{N,\nu}) \nabla^2 \mathcal{V}(Z_{N,\nu})] \right) d\nu,$$

for all $t \in \mathbb{R}_{\geq 0}$, where $K_\sigma(\nu, Z_{N,\nu}) = J_\sigma(\nu, Z_{N,\nu}) J_\sigma(\nu, Z_{N,\nu})^\top \in \mathbb{S}^n$. Substituting the expressions in (D.24), Proposition D.4, for the last two terms on the right hand side leads to

$$e^{2\lambda\tau(t)}\mathcal{V}(Z_{N,\tau(t)}) \leq \int_0^{\tau(t)} e^{2\lambda\nu} \left([\phi_\mu(\nu, Z_{N,\nu}) + \tilde{\mathcal{U}}(\nu, Z_{N,\nu})] d\nu + \phi_\sigma(\nu, Z_{N,\nu}) dW_\nu \right) \\ + \int_0^{\tau(t)} e^{2\lambda\nu} \left([\phi_{\mu^\parallel}(\nu, Z_{N,\nu}) + \phi_U(\nu, Z_{N,\nu})] d\nu + \phi_{\sigma^\parallel}(\nu, Z_{N,\nu}) dW_\nu \right), \quad (112)$$

for all $t \in \mathbb{R}_{\geq 0}$, where

$$\phi_{\mu^\parallel}(\nu, Z_{N,\nu}) = \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \mathring{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}), \quad \phi_{\sigma^\parallel}(\nu, Z_{N,\nu}) = \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \mathring{F}_\sigma^\parallel(\nu, Z_{N,\nu}), \\ \phi_U(\nu, Z_{N,\nu}) = \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \mathring{F}_r(Z_N)(\nu).$$

Next, we use Proposition D.5 to obtain the following expression

$$\int_0^{\tau(t)} e^{2\lambda\nu} \left([\phi_U(\nu, Z_{N,\nu}) + \phi_{\mu^\parallel}(\nu, Z_{N,\nu})] d\nu + \phi_{\sigma^\parallel}(\nu, Z_{N,\nu}) dW_\nu \right) \\ = \int_0^{\tau(t)} e^{2\lambda\nu} \left(\phi_{\mu^\parallel}(\nu, Z_{N,\nu}) + \phi_{U_\mu}(\nu, Z_{N,\nu}; \omega) \right) d\nu + \int_0^{\tau(t)} e^{2\lambda\nu} \left(\phi_{\sigma^\parallel}(\nu, Z_{N,\nu}) + \phi_{U_\sigma}(\nu, Z_{N,\nu}; \omega) \right) dW_\nu \\ + \int_0^{\tau(t)} \left(\hat{\mathcal{U}}_\mu(\tau(t), \nu, Z_N; \omega) d\nu + \hat{\mathcal{U}}_\sigma(\tau(t), \nu, Z_N; \omega) dW_\nu \right), \quad \forall t \in \mathbb{R}_{\geq 0}, \quad (113)$$

where

$$\hat{\mathcal{U}}_\mu(\tau(t), \nu, Z_N; \omega) = e^{-\omega\tau(t)} \frac{\omega}{2\lambda - \omega} \left(e^{\omega\tau(t)} \mathcal{P}(\tau(t), \nu) - e^{2\lambda\tau(t)} \partial \mathcal{V}(Z_{N,\tau(t)})^\top g(\tau(t)) \right) e^{\omega\nu} \mathring{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}), \\ \hat{\mathcal{U}}_\sigma(\tau(t), \nu, Z_N; \omega) = e^{-\omega\tau(t)} \frac{\omega}{2\lambda - \omega} \left(e^{\omega\tau(t)} \mathcal{P}(\tau(t), \nu) - e^{2\lambda\tau(t)} \partial \mathcal{V}(Z_{N,\tau(t)})^\top g(\tau(t)) \right) e^{\omega\nu} \mathring{F}_\sigma^\parallel(\nu, Z_{N,\nu}), \\ \phi_{U_\mu}(\nu, Z_{N,\nu}; \omega) = \frac{\omega}{2\lambda - \omega} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \mathring{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}), \\ \phi_{U_\sigma}(\nu, Z_{N,\nu}; \omega) = \frac{\omega}{2\lambda - \omega} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \mathring{F}_\sigma^\parallel(\nu, Z_{N,\nu}).$$

Using the definitions of ϕ_{μ^\parallel} and ϕ_{σ^\parallel} in (112), we obtain

$$\phi_{\mu^\parallel}(\nu, Z_{N,\nu}) + \phi_{U_\mu}(\nu, Z_{N,\nu}; \omega) = \frac{2\lambda}{2\lambda - \omega} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \mathring{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}), \\ \phi_{\sigma^\parallel}(\nu, Z_{N,\nu}) + \phi_{U_\sigma}(\nu, Z_{N,\nu}; \omega) = \frac{2\lambda}{2\lambda - \omega} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \mathring{F}_\sigma^\parallel(\nu, Z_{N,\nu}).$$

Substituting into (113) thus produces

$$\int_0^{\tau(t)} e^{2\lambda\nu} \left([\phi_U(\nu, Z_{N,\nu}) + \phi_{\mu^\parallel}(\nu, Z_{N,\nu})] d\nu + \phi_{\sigma^\parallel}(\nu, Z_{N,\nu}) dW_\nu \right) \\ = e^{-\omega\tau(t)} \int_0^{\tau(t)} \left(\mathcal{U}_\mu(\tau(t), \nu, Z_N; \omega) d\nu + \mathcal{U}_\sigma(\tau(t), \nu, Z_N; \omega) dW_\nu \right), \quad \forall t \in \mathbb{R}_{\geq 0},$$

where \mathcal{U}_μ and \mathcal{U}_σ are defined in (109). Substituting the above expression for the last integral on the right hand side of (112) and using the definitions of Ξ , Ξ_μ , and Ξ_σ in (107) yields the following bound:

$$e^{2\lambda\tau(t)}\mathcal{V}(Z_{N,\tau(t)}) \leq \Xi(\tau(t), Z_N) + e^{-\omega\tau(t)} \Xi_\mu(\tau(t), Z_N) + \Xi_\sigma(\tau(t), Z_N), \quad \forall t \in \mathbb{R}_{\geq 0},$$

which implies

$$\mathcal{V}(Z_{N,\tau(t)}) \leq e^{-2\lambda\tau(t)} \Xi(\tau(t), Z_N) + e^{-(2\lambda+\omega)\tau(t)} \Xi_{\mathcal{U}}(\tau(t), Z_N) + e^{-2\lambda\tau(t)} \tilde{\Xi}_{\mathcal{U}}(\tau(t), Z_N), \quad \forall t \in \mathbb{R}_{\geq 0}. \quad (114)$$

Then, using the linearity of the expectation operator, we obtain

$$\begin{aligned} \mathbb{E}[\mathcal{V}(Z_{N,\tau(t)})] &\leq \mathbb{E}\left[e^{-2\lambda\tau(t)} \Xi(\tau(t), Z_N)\right] + \mathbb{E}\left[e^{-(2\lambda+\omega)\tau(t)} \Xi_{\mathcal{U}}(\tau(t), Z_N)\right] \\ &\quad + \mathbb{E}\left[e^{-2\lambda\tau(t)} \tilde{\Xi}_{\mathcal{U}}(\tau(t), Z_N)\right], \quad \forall t \in \mathbb{R}_{\geq 0}, \end{aligned}$$

which establishes (105). Finally, applying Minkowski's inequality to (114) for $p \in \mathbb{N}_{\geq 2}$ yields the bound in (106), thus concluding the proof. \square

We next derive bounds on the conditional moments of the truncated joint true (uncertain)-reference process up to the stopping time τ_N .

Lemma 4.6 *Let the assumptions in Sec. 2.2 hold and let $\varrho_a(p) \in \mathbb{R}_{>0}$, $p \in \{1, \dots, p^*\}$, be as defined in (16b), where p^* is defined in Assumption 3. Further suppose that the filter bandwidth and sampling period conditions in (21) and (24), respectively, hold. Then the truncated joint process $Z_{N,t} = X_{N,t} - X_{N,t}^r$ in Definition 9, satisfies*

$$\|X_{N,t} - X_{N,t}^r\|_{L_{2p}} < \varrho_a(p), \quad \forall t \in [0, \tau_N], \quad (115)$$

where the stopping time τ_N is as in (100) for arbitrary $N \in \mathbb{R}_{>0}$.

Proof. We prove (115) by contradiction. The t -continuity of the strong solutions $X_{N,t}$ and $X_{N,t}^r$ on $[0, \tau_N]$ implies that $\|X_{N,t} - X_{N,t}^r\|_{L_{2p}}$ is t -continuous as well. We have from Definitions 8 - 9 that

$$\|Z_{N,0}\| = \|X_{N,0} - X_{N,0}^r\| = 0.$$

Thus, since $\varrho_a(p) > 0$, the deterministic nature and continuity of $t \mapsto \|X_{N,t} - X_{N,t}^r\|_{L_{2p}}$ allow us to formulate the negation of (115) as

$$\neg(115) : \exists \text{ (deterministic) } t^* \in [0, \tau_N] \text{ such that } \begin{cases} \|X_{N,t} - X_{N,t}^r\|_{L_{2p}} < \varrho_a(p), & \forall t \in [0, t^*), \\ \|X_{N,t^*} - X_{N,t^*}^r\|_{L_{2p}} = \varrho_a(p). \end{cases} \quad (116)$$

Now, $\tau(t) = (t \wedge t^*) \in [0, t^*]$, for all $t \in \mathbb{R}_{\geq 0}$ because $\tau^* = t^*$ in (104) since we are assuming that $t^* < \tau_N$. Thus, we may write the bounds (105) and (106) in Lemma 4.5 as

$$\begin{aligned} \mathbb{E}[\mathcal{V}(Z_{N,t})] &\leq \mathbb{E}\left[e^{-2\lambda\tau(t)} \Xi(\tau(t), Z_N)\right] + \mathbb{E}\left[e^{-(2\lambda+\omega)\tau(t)} \Xi_{\mathcal{U}}(\tau(t), Z_N)\right] \\ &\quad + \mathbb{E}\left[e^{-2\lambda\tau(t)} \tilde{\Xi}_{\mathcal{U}}(\tau(t), Z_N)\right], \quad \forall t \in [0, t^*], \end{aligned} \quad (117)$$

and

$$\begin{aligned} \|\mathcal{V}(Z_{N,t})\|_{L_p} &\leq \left\|e^{-2\lambda\tau(t)} \Xi(\tau(t), Z_N)\right\|_{L_p} + \left\|e^{-(2\lambda+\omega)\tau(t)} \Xi_{\mathcal{U}}(\tau(t), Z_N)\right\|_{L_p} \\ &\quad + \left\|e^{-2\lambda\tau(t)} \tilde{\Xi}_{\mathcal{U}}(\tau(t), Z_N)\right\|_{L_p}, \end{aligned} \quad (118)$$

for all $(t, p) \in [0, t^*] \times \{2, \dots, p^*\}$. Now, let us define

$$\tilde{\mathcal{E}}'(t, 2p) \doteq \|X_{N,t} - X_{N,t}^r\|_{L_{2p}}, \quad (119a)$$

$$\tilde{\mathcal{E}}(2p) \doteq \sup_{\nu \in [0, t^*]} \|X_{N,\nu} - X_{N,\nu}^r\|_{L_{2p}}, \quad \mathcal{E}(2p) \doteq \sup_{\nu \in [0, t^*]} \|X_{N,\nu}\|_{L_{2p}}, \quad \mathcal{E}_r(2p) \doteq \sup_{\nu \in [0, t^*]} \|X_{N,\nu}^r\|_{L_{2p}}, \quad (119b)$$

for $p \in \{1, \dots, p^*\}$. Then, using the definition $\mathcal{V}(Z_{N,t}) \doteq \|Z_{N,t}\|^2 = \|X_{N,t} - X_{N,t}^r\|^2$ in Lemma 4.5, along with the probability measure $\mathbb{Z}_N(0, 0; dz, t)$ defined in (102), we see that

$$\|\mathcal{V}(Z_{N,t})\|_{L_p} = \left(\int_{\mathbb{R}^n} \|Z_{N,t}\|^{2p} \mathbb{Z}_N(0, 0; dz, t) \right)^{\frac{2}{2p}} = \left(\|Z_{N,t}\|_{L_{2p}} \right)^2 = \left(\|X_{N,t} - X_{N,t}^r\|_{L_{2p}} \right)^2 = \left(\tilde{\mathcal{E}}'(t, 2p) \right)^2.$$

Hence, we can write (117) and (118) as

$$\begin{aligned} \left(\tilde{\mathcal{E}}'(t, 2) \right)^2 &\leq \mathbb{E} \left[e^{-2\lambda\tau(t)} \Xi(\tau(t), Z_N) \right] + \mathbb{E} \left[e^{-(2\lambda + \omega)\tau(t)} \Xi_{\mathcal{U}}(\tau(t), Z_N) \right] \\ &\quad + \mathbb{E} \left[e^{-2\lambda\tau(t)} \tilde{\Xi}_{\mathcal{U}}(\tau(t), Z_N) \right], \quad \forall t \in [0, t^*], \end{aligned}$$

and

$$\left(\tilde{\mathcal{E}}'(t, 2\mathbf{p}) \right)^2 \leq \left\| e^{-2\lambda\tau(t)} \Xi(\tau(t), Z_N) \right\|_{L_p} + \left\| e^{-(2\lambda + \omega)\tau(t)} \Xi_{\mathcal{U}}(\tau(t), Z_N) \right\|_{L_p} + \left\| e^{-2\lambda\tau(t)} \tilde{\Xi}_{\mathcal{U}}(\tau(t), Z_N) \right\|_{L_p},$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \{2, \dots, \mathbf{p}^*\}$. Since $\tau^* = t^*$ by assumption, we can substitute the bounds derived in Lemma D.5 into the above inequalities and get

$$\left(\tilde{\mathcal{E}}'(t, 2) \right)^2 \leq \Phi(1, \omega, \mathbf{T}_s), \quad \left(\tilde{\mathcal{E}}'(t, 2\mathbf{p}) \right)^2 \leq \Phi(\mathbf{p}, \omega, \mathbf{T}_s), \quad \forall (t, \mathbf{p}) \in [0, t^*] \times \{2, \dots, \mathbf{p}^*\},$$

where we have defined the following for $\mathbf{p} \in \{1, \dots, \mathbf{p}^*\}$:

$$\begin{aligned} \Phi(\mathbf{p}, \omega, \mathbf{T}_s) &= \Delta_{\odot}(\mathbf{p}, \omega) \left(\tilde{\mathcal{E}}(2\mathbf{p}) \right)^{\frac{1}{2}} + \Delta_{\odot}(\mathbf{p}, \omega) \tilde{\mathcal{E}}(2\mathbf{p}) \\ &\quad + \left(\tilde{\Delta}_{\otimes}(\mathbf{p}, \omega) (\mathcal{E}(2\mathbf{p}) + \mathcal{E}_r(2\mathbf{p})) + \Delta_{\otimes}(\mathbf{p}, \omega) \tilde{\mathcal{E}}(2\mathbf{p}) \right) \left(\tilde{\mathcal{E}}(2\mathbf{p}) \right)^{\frac{1}{2}} \\ &\quad + \left(\tilde{\Delta}_{\otimes}(\omega) (\mathcal{E}(2\mathbf{p}) + \mathcal{E}_r(2\mathbf{p})) + \Delta_{\otimes}(\omega) \tilde{\mathcal{E}}(2\mathbf{p}) \right) \tilde{\mathcal{E}}(2\mathbf{p}) + \sum_{i=1}^3 \Upsilon_i \left(\mathcal{E}(2\mathbf{p}), \tilde{\mathcal{E}}(2\mathbf{p}); \mathbf{p}, \omega, \mathbf{T}_s \right). \end{aligned} \quad (120)$$

The two bounds above can be combined into the following single bound over $\mathbf{p} \in \{1, \dots, \mathbf{p}^*\}$:

$$\left(\tilde{\mathcal{E}}'(t, 2\mathbf{p}) \right)^2 \leq \Phi(\mathbf{p}, \omega, \mathbf{T}_s), \quad \forall (t, \mathbf{p}) \in [0, t^*] \times \{1, \dots, \mathbf{p}^*\}. \quad (121)$$

Next, as in the proof of Lemma 4.2, the continuity of $\nu \mapsto \|X_{N,\nu} - X_{N,\nu}^r\|_{L_{2\mathbf{p}}}$ over the closed interval $[0, t^*]$ implies the existence of some $t' \in [0, t^*]$, such that

$$\tilde{\mathcal{E}}(2\mathbf{p}) \doteq \sup_{\nu \in [0, t^*]} \|X_{N,\nu} - X_{N,\nu}^r\|_{L_{2\mathbf{p}}} = \|X_{N,t'} - X_{N,t'}^r\|_{L_{2\mathbf{p}}}.$$

Hence, the contradiction statement in (116) leads us to conclude that

$$\tilde{\mathcal{E}}(2\mathbf{p}) \doteq \sup_{\nu \in [0, t^*]} \|X_{N,\nu} - X_{N,\nu}^r\|_{L_{2\mathbf{p}}} \leq \varrho_a(\mathbf{p}), \quad (122)$$

for $\mathbf{p} \in \{1, \dots, \mathbf{p}^*\}$. Similarly, the definitions of $\mathcal{E}(2\mathbf{p})$ and $\mathcal{E}_r(2\mathbf{p})$ in (119b), along with the contradiction statement in (116), Assumption 3, and the result in Lemma 4.3 imply

$$\begin{aligned} \mathcal{E}_r(2\mathbf{p}) &\doteq \sup_{\nu \in [0, t^*]} \|X_{N,\nu}^r\|_{L_{2\mathbf{p}}} = \sup_{\nu \in [0, t^*]} \|X_{N,\nu}^r\|_{L_{2\mathbf{p}}} \leq \sup_{\nu \in [0, t^*]} \|X_{N,\nu}^r - X_{N,\nu}^*\|_{L_{2\mathbf{p}}} + \sup_{\nu \in [0, t^*]} \|X_{N,\nu}^*\|_{L_{2\mathbf{p}}} \\ &< \varrho_r(\mathbf{p}) + \Delta_*, \end{aligned} \quad (123a)$$

$$\mathcal{E}(2\mathbf{p}) \doteq \sup_{\nu \in [0, t^*]} \|X_{N,\nu}\|_{L_{2\mathbf{p}}} \leq \sup_{\nu \in [0, t^*]} \|X_{N,\nu} - X_{N,\nu}^r\|_{L_{2\mathbf{p}}} + \sup_{\nu \in [0, t^*]} \|X_{N,\nu}^r\|_{L_{2\mathbf{p}}} < \varrho_a(\mathbf{p}) + \varrho_r(\mathbf{p}) + \Delta_*. \quad (123b)$$

Substituting (122) - (123) into (120), and substituting the resulting bound into (121) gives:

$$\begin{aligned} \left(\tilde{\mathcal{E}}'(t, 2\mathbf{p}) \right)^2 &< \Delta_{\odot}(\mathbf{p}, \omega) (\varrho_a(\mathbf{p}))^{\frac{1}{2}} + \Delta_{\odot}(\mathbf{p}, \omega) \varrho_a(\mathbf{p}) + \left(\tilde{\Delta}_{\otimes}(\mathbf{p}, \omega) \varrho''(\mathbf{p}) + \Delta_{\otimes}(\mathbf{p}, \omega) \varrho_a(\mathbf{p}) \right) (\varrho_a(\mathbf{p}))^{\frac{1}{2}} \\ &\quad + \left(\tilde{\Delta}_{\otimes}(\omega) \varrho''(\mathbf{p}) + \Delta_{\otimes}(\omega) \varrho_a(\mathbf{p}) \right) \varrho_a(\mathbf{p}) + \sum_{i=1}^3 \Upsilon_i (\varrho'(\mathbf{p}), \varrho_a(\mathbf{p}); \mathbf{p}, \omega, \mathbf{T}_s), \end{aligned} \quad (124)$$

for all $t \in [0, t^*]$, where

$$\varrho'(\mathbf{p}) = \varrho_a(\mathbf{p}) + \varrho_r(\mathbf{p}) + \Delta_*, \quad \varrho''(\mathbf{p}) = \varrho_a(\mathbf{p}) + 2(\varrho_r(\mathbf{p}) + \Delta_*).$$

Furthermore, we have used the following due to the definitions in (25), Section 3.2.2, and the monotonicity property (A.21) in Proposition A.1:

$$\Upsilon_i \left(\mathcal{E}(2\mathbf{p}), \tilde{\mathcal{E}}(2\mathbf{p}); \mathbf{p}, \boldsymbol{\omega}, T_s \right) \leq \Upsilon_i \left(\boldsymbol{\varrho}'(\mathbf{p}), \boldsymbol{\varrho}_a(\mathbf{p}); \mathbf{p}, \boldsymbol{\omega}, T_s \right), \quad \forall i \in \{1, 2, 3\}.$$

Now, using the definition of $\tilde{\mathcal{E}}(\cdot, 2\mathbf{p})$ in (119a) and the contradiction statement in (116), one sees that $\tilde{\mathcal{E}}'(t^*, 2\mathbf{p}) = \boldsymbol{\varrho}_a(\mathbf{p})$ at the temporal instance $t = t^*$. Hence, at the temporal instance $t = t^*$, the inequality in (124) reduces to

$$\begin{aligned} \boldsymbol{\varrho}_a(\mathbf{p})^2 &< \Delta_{\odot}(\mathbf{p}, \boldsymbol{\omega}) (\boldsymbol{\varrho}_a(\mathbf{p}))^{\frac{1}{2}} + \Delta_{\odot}(\mathbf{p}, \boldsymbol{\omega}) \boldsymbol{\varrho}_a(\mathbf{p}) + \left(\check{\Delta}_{\otimes}(\mathbf{p}, \boldsymbol{\omega}) \boldsymbol{\varrho}''(\mathbf{p}) + \Delta_{\otimes}(\mathbf{p}, \boldsymbol{\omega}) \boldsymbol{\varrho}_a(\mathbf{p}) \right) (\boldsymbol{\varrho}_a(\mathbf{p}))^{\frac{1}{2}} \\ &\quad + \left(\check{\Delta}_{\otimes}(\boldsymbol{\omega}) \boldsymbol{\varrho}''(\mathbf{p}) + \Delta_{\otimes}(\boldsymbol{\omega}) \boldsymbol{\varrho}_a(\mathbf{p}) \right) \boldsymbol{\varrho}_a(\mathbf{p}) + \sum_{i=1}^3 \Upsilon_i \left(\boldsymbol{\varrho}'(\mathbf{p}), \boldsymbol{\varrho}_a(\mathbf{p}); \mathbf{p}, \boldsymbol{\omega}, T_s \right). \end{aligned}$$

Substituting the definition of constants $\Delta_i(\mathbf{p}, \boldsymbol{\omega})$, $i \in \{\odot, \ominus, \otimes, \oplus\}$ from Lemma D.5 into the above inequality and re-arranging the terms leads to

$$\begin{aligned} \left(1 - \frac{\Delta_g^{\perp} L_{\mu}^{\perp}}{\lambda} \right) \boldsymbol{\varrho}_a(\mathbf{p})^2 &< \Gamma_a(\boldsymbol{\varrho}_a(\mathbf{p}), \mathbf{p}, \boldsymbol{\omega}) + \frac{1}{|2\lambda - \boldsymbol{\omega}|} \Theta_a(\boldsymbol{\varrho}_a(\mathbf{p}), \boldsymbol{\varrho}_r(\mathbf{p}), \mathbf{p}, \boldsymbol{\omega}) \\ &\quad + \sum_{i=1}^3 \Upsilon_{a_i}(\boldsymbol{\varrho}'(\mathbf{p}), \boldsymbol{\varrho}_a(\mathbf{p}); \mathbf{p}, \boldsymbol{\omega}, T_s), \end{aligned}$$

which further implies that

$$\begin{aligned} \sum_{i=1}^3 \Upsilon_{a_i}(\boldsymbol{\varrho}'(\mathbf{p}), \boldsymbol{\varrho}_a(\mathbf{p}); \mathbf{p}, \boldsymbol{\omega}, T_s) &> \left(1 - \frac{\Delta_g^{\perp} L_{\mu}^{\perp}}{\lambda} \right) \boldsymbol{\varrho}_a(\mathbf{p})^2 - \Gamma_a(\boldsymbol{\varrho}_a(\mathbf{p}), \mathbf{p}, \boldsymbol{\omega}) \\ &\quad - \frac{1}{|2\lambda - \boldsymbol{\omega}|} \Theta_a(\boldsymbol{\varrho}_a(\mathbf{p}), \boldsymbol{\varrho}_r(\mathbf{p}), \mathbf{p}, \boldsymbol{\omega}). \quad (125) \end{aligned}$$

However, one sees from (24) that the choice of T_s ensures that

$$\begin{aligned} \sum_{i=1}^3 \Upsilon_{a_i}(\boldsymbol{\varrho}'(\mathbf{p}), \boldsymbol{\varrho}_a(\mathbf{p}); \mathbf{p}, \boldsymbol{\omega}, T_s) &\leq \left(1 - \frac{\Delta_g^{\perp} L_{\mu}^{\perp}}{\lambda} \right) \boldsymbol{\varrho}_a(\mathbf{p})^2 - \Gamma_a(\boldsymbol{\varrho}_a(\mathbf{p}), \mathbf{p}, \boldsymbol{\omega}) \\ &\quad - \frac{1}{|2\lambda - \boldsymbol{\omega}|} \Theta_a(\boldsymbol{\varrho}_a(\mathbf{p}), \boldsymbol{\varrho}_r(\mathbf{p}), \mathbf{p}, \boldsymbol{\omega}), \end{aligned}$$

which contradicts the conclusion above in (125) that follows from the postulation in (116). Therefore, the desired result in (115) is proved since its negation (116) leads to a contradiction. \square

Next, we extend the strong solutions of the truncated process $Z_{N,t}$ to the original joint process Z_t in Definition 8.

Lemma 4.7 *If the hypotheses of Lemmas 4.3 and 4.6 hold, then the joint process in Definition 8 admits a unique strong solution $Z_t = X_t - X_t^r$, for all $T \in (0, \infty)$. Furthermore,*

$$\|X_t - X_t^r\|_{L_{2p}} < \boldsymbol{\varrho}_a(\mathbf{p}), \quad \forall t \in [0, T], \quad (126)$$

for any $T \in (0, \infty)$.

Proof. We omit the details of the proof since it follows identically to Lemma 4.3 with $Y_N = (X_N^r, X_N^*) \in \mathbb{R}^{2n}$ and $U_N \subset \mathbb{R}^{2n}$ (see (30)), replaced by $Z_{N,t} = X_{N,t} - X_{N,t}^r \in \mathbb{R}^n$ and $\tilde{U}_N \subset \mathbb{R}^n$ (see (98)), respectively. \square

The next result follows from Lemma 4.4 *mutatis mutandis*.

Lemma 4.8 *Let the hypotheses of Lemma 4.6 hold. Then, if in addition, the hypotheses of Lemma 4.3 holds for $\alpha(X_0^r, X_0^*)_{2p} = W_{2p}(\xi_0, \xi_0^*)$, then*

$$W_{2p}(\mathbb{X}_t, \mathbb{X}_t^r) < \boldsymbol{\varrho}_a(\mathbf{p}), \quad \forall t \in [0, T]. \quad (127)$$

If instead the hypotheses of Lemma 4.3 holds for $\alpha(X_0^r, X_0^*)_{2p} = \|X_0^r - X_0^*\|_{L_{2p}}$, then for any $\delta \in (0, 1)$ such that $\log \sqrt{1/\delta} = p$,

$$\mathbb{P}(\|X_t - X_t^r\| \geq \kappa_a(\delta)) < \delta, \quad \forall t \in [0, T], \quad (128)$$

where $\kappa_a(\delta) = e \cdot \varrho_a(\log \sqrt{1/\delta})$.

We can now state the main result of the manuscript which follows directly from Lemmas 4.4 and 4.8.

Theorem 4.1 Suppose that the assumptions in Sec. 2.2 hold, and let $\varrho_r(p), \varrho_a(p) \in \mathbb{R}_{>0}$, $p \in \{1, \dots, p^*\}$, be as defined in (16) with $\alpha(X_0^r, X_0^*)_{2p} = W_{2p}(\xi_0, \xi_0^*)$. Additionally, let the filter bandwidth and sampling period conditions in (21) and (24), respectively, hold. Then,

$$W_{2p}(\mathbb{X}_t, \mathbb{X}_t^*) < \varrho(p), \quad \forall t \in [0, T], \quad (\text{Distributional UB}) \quad (129a)$$

$$W_{2p}(\mathbb{X}_t^r, \mathbb{X}_t^*) \leq \hat{\varrho}(p, \omega, \hat{t}) < \varrho(p), \quad \forall t \in [\hat{t}, T], \quad (\text{Distributional UUB}) \quad (129b)$$

where $\hat{t} \in (0, T)$ is any deterministic constant, and as given in (17), $\varrho(p) = \varrho_r(p) + \varrho_a(p)$. Moreover,

$$\hat{\varrho}(p, \omega, \hat{t}) = \hat{\varrho}_r(p, \omega, \hat{t}) + \varrho_a(p),$$

where $\hat{\varrho}_r(p, \omega, \hat{t})$ is defined in statement of Lemma 4.4.

If instead the above hypotheses are satisfied for $\alpha(X_0^r, X_0^*)_{2p} = \|X_0^r - X_0^*\|_{L_{2p}}$, then for any $\delta \in (0, 1)$ such that $\log \sqrt{1/\delta} = p$,

$$\mathbb{P}(\|X_t - X_t^*\| \geq \kappa(\delta)) < \delta, \quad \forall t \in [0, T], \quad (\text{Pathwise UB}) \quad (130a)$$

$$\mathbb{P}(\|X_t - X_t^*\| \geq \hat{\kappa}(\delta, \omega, \hat{t})) < \delta, \quad \forall t \in [\hat{t}, T], \quad (\text{Pathwise UUB}) \quad (130b)$$

where $\kappa(\delta) = e \cdot \varrho(\log \sqrt{1/\delta})$ and $\hat{\kappa}(\delta, \omega, \hat{t}) = e \cdot \hat{\varrho}(\log \sqrt{1/\delta}, \omega, \hat{t})$.

Proof. Since the Wasserstein metrics satisfy the axioms of distance [29], we get

$$W_{2p}(\mathbb{X}_t, \mathbb{X}_t^*) \leq W_{2p}(\mathbb{X}_t, \mathbb{X}_t^r) + W_{2p}(\mathbb{X}_t^r, \mathbb{X}_t^*), \quad \forall t \in [0, T].$$

The bounds in (129) then follow from (84) and (127) in Lemmas 4.4 and 4.8, respectively.

The bounds in (130) follow similarly from (86) and (128) in Lemmas 4.4 and 4.8, respectively, and the following due to the Minkowski's inequality:

$$\|X_t - X_t^*\|_{L_{2p}} \leq \|X_t - X_t^r\|_{L_{2p}} + \|X_t^r - X_t^*\|_{L_{2p}}, \quad \forall t \in [0, T].$$

□

4.3 Discussion

A few critical points merit a discussion.

Tighter High-probability Guarantees: Recall the definition $\varrho(p) = \varrho_r(p) + \varrho_a(p)$ from (17), where $\varrho_r(p)$ and $\varrho_a(p)$ are defined in (16). The terms $\varrho_r(p)$ and $\varrho_a(p)$ depend on the moment $p \in \{1, \dots, p^*\}$, through the constants presented in Appendix A.1. It is easy to establish that $\{\varrho_r(p), \varrho_a(p)\} \in \mathcal{O}(p^2)$, as $p \rightarrow \infty$. It then follows that the pathwise uniform bound $\kappa(\delta)$ in (130a), Thm. 4.1, satisfies $\kappa(\delta) \in \mathcal{O}(\log^2 \sqrt{1/\delta})$, as $\delta \downarrow 0$. It can be similarly shown that the pathwise uniform ultimate bound $\hat{\kappa}(\delta, \omega, \hat{t})$ in (130b) is also of the same order in $\delta \in (0, 1)$.

The order 2 *polylog* (polylogarithmic) dependence of the bounds κ and $\hat{\kappa}$ on the probability of violation $\delta \in (0, 1)$ implies much tighter bounds for small δ when compared to bounds resulting from using up to only the first two moments. For example, if the \mathcal{L}_1 -DRAC controller could only guarantee bounds up to $p = 1$ (moments up to $2p$), then $\kappa \in \mathcal{O}(\sqrt{1/\delta})$, as $\delta \downarrow 0$, which leads to substantially large bounds for small δ . As an example, for $\delta = 10^{-10}$, $\log^2 \sqrt{1/\delta} = 132.55$, while $\sqrt{1/\delta} = 10^5$.

Generalization of the Deterministic Design and Performance: As discussed in Remark 3.2, the filter bandwidth conditions (21) are $\mathcal{O}(1/\sqrt{\omega})$, as $\omega \rightarrow \infty$. For the purposes of this discussion, we may interpret the order of $1/\omega$ on

which the bandwidth functions depend as the ‘strength’ of the controller’s ability. In the absence of the diffusion vector fields, however, one sees that

$$\left\{ \frac{1}{|2\lambda - \omega|} \Theta_r(\underline{q}_r(p), p, \omega), \frac{1}{|2\lambda - \omega|} \Theta_a(\underline{q}_a(p), \underline{q}_r(p), p, \omega) \right\} \in \mathcal{O}\left(\frac{1}{\omega}\right),$$

as $\omega \rightarrow \infty$, which represents a ‘stronger’ ability of the controller. Indeed, the $\mathcal{O}(1/\omega)$ dependence of the bandwidth conditions is consistent with our previous results on \mathcal{L}_1 adaptive controller design for deterministic systems, see e.g. [104, 110]. It can be similarly shown that the functions Υ_{a_i} , $i \in \{1, 2, 3\}$, that define the sampling period condition in (24) satisfy

$$\sum_{i=1}^3 \Upsilon_{a_i}(\underline{q}'(p), \underline{q}_a(p); p, \omega, T_s) \in \mathcal{O}\left(\sqrt{T_s}\right), \text{ as } T_s \downarrow 0.$$

Instead, in the absence of the diffusion terms, the function above belongs to $\mathcal{O}(T_s)$, as $T_s \downarrow 0$, which is similar to piecewise constant adaptation laws utilized in the \mathcal{L}_1 adaptive controller for deterministic systems, see e.g. [111, 112]. Finally, in the absence of diffusion terms, and unmatched uncertainties, i.e., deterministic system subject to only matched uncertainties, we recover definitions of the bounds $\underline{q}_r(p)$ and $\underline{q}_a(p)$ that are consistent with our previous results on deterministic systems in [104].

These observations lead to the conclusion that the \mathcal{L}_1 -DRAC control generalizes \mathcal{L}_1 adaptive control design for deterministic systems, in both design requirements and performance guarantees.

5 Conclusion

We provide \mathcal{L}_1 -DRAC methodology that allows the synthesis of controllers that can guarantee distributionally robust control on uncertain nonlinear processes. The \mathcal{L}_1 -DRAC is based on the \mathcal{L}_1 adaptive control architecture. The robustness guarantees ensured by the \mathcal{L}_1 -DRAC feedback are in the form of Wasserstein ambiguity sets centered on nominal state distributions with *a priori* known radius. The class of uncertain nonlinear systems for which the \mathcal{L}_1 -DRAC controller is applicable consists of systems with unbounded epistemic uncertainties in both drift and diffusion: a significantly larger class of systems than currently considered for similar synthesis problems in the literature.

6 Acknowledgements

The authors are thankful to Evangelos Theodorou from GaTech for many useful discussions in the early stages of this development. The authors gratefully acknowledge funding support by AFOSR Grant FA9550-21-1-0411, the National Aeronautics and Space Administration (NASA) under Grants 80NSSC22M0070 and 80NSSC20M0229, and by the National Science Foundation (NSF) under Grants CMMI 2135925 and IIS 2331878.

References

- [1] P. A. Ioannou and J. Sun, *Robust adaptive control*. PTR Prentice-Hall Upper Saddle River, NJ, 1996, vol. 1.
- [2] E. Lavretsky and K. A. Wise, *Robust and adaptive control: With aerospace applications*, 2 ed. Springer, 2024.
- [3] S. Kuutti, R. Bowden, Y. Jin, P. Barber, and S. Fallah, “A survey of deep learning applications to autonomous vehicle control,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 2, pp. 712–733, 2020.
- [4] G. Shi, X. Shi, M. O’Connell, R. Yu, K. Azizzadenesheli, A. Anandkumar, Y. Yue, and S.-J. Chung, “Neural lander: Stable drone landing control using learned dynamics,” in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 9784–9790.
- [5] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, “Deep reinforcement learning: A brief survey,” *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 26–38, 2017.
- [6] I. Lenz, R. A. Knepper, and A. Saxena, “DeepMPC: Learning deep latent features for model predictive control,” in *Robotics: Science and Systems*, vol. 10. Rome, Italy, 2015, p. 25.
- [7] D. Bertsekas, *Reinforcement learning and optimal control*. Athena Scientific, 2019.
- [8] J. Moos, K. Hansel, H. Abdulsamad, S. Stark, D. Clever, and J. Peters, “Robust reinforcement learning: A review of foundations and recent advances,” *Machine Learning and Knowledge Extraction*, vol. 4, no. 1, pp. 276–315, 2022.

- [9] I. M. Mitchell, A. M. Bayen, and C. J. Tomlin, “A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games,” *IEEE Transactions on automatic control*, vol. 50, no. 7, pp. 947–957, 2005.
- [10] S. L. Herbert, M. Chen, S. Han, S. Bansal, J. F. Fisac, and C. J. Tomlin, “FaSTrack: A modular framework for fast and guaranteed safe motion planning,” in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 1517–1522.
- [11] S. Singh, B. Landry, A. Majumdar, J.-J. Slotine, and M. Pavone, “Robust feedback motion planning via contraction theory,” *The International Journal of Robotics Research*, 2019.
- [12] V. S. Borkar, *Stochastic approximation: A dynamical systems viewpoint*. Springer, 2009, vol. 48.
- [13] X. Mao, *Exponential stability of stochastic differential equations*. Marcel Dekker, 1994.
- [14] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic differential equations*. Springer Berlin, Heidelberg, 1992.
- [15] B. Oksendal, *Stochastic differential equations: An introduction with applications*. Springer Science & Business Media, 2013.
- [16] V. N. Vapnik, “An overview of statistical learning theory,” *IEEE transactions on neural networks*, vol. 10, no. 5, pp. 988–999, 1999.
- [17] M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of machine learning*. MIT press, 2018.
- [18] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, “Score-based generative modeling through stochastic differential equations,” in *International Conference on Learning Representations (ICLR)*, 2021.
- [19] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*. Springer, 2006, vol. 4, no. 4.
- [20] N. Boffi, S. Tu, N. Matni, J.-J. Slotine, and V. Sindhvani, “Learning stability certificates from data,” in *Conference on Robot Learning*. PMLR, 2021, pp. 1341–1350.
- [21] J. Quinero-Candela, M. Sugiyama, A. Schwaighofer, and N. D. Lawrence, *Dataset shift in machine learning*. Mit Press, 2008.
- [22] M. Sugiyama and M. Kawanabe, *Machine learning in non-stationary environments: Introduction to covariate shift adaptation*. MIT press, 2012.
- [23] O. Wiles, S. Goyal, F. Stimberg, S.-A. Rebuffi, I. Ktena, K. D. Dvijotham, and A. T. Cemgil, “A fine-grained analysis on distribution shift,” in *International Conference on Learning Representations*, 2021.
- [24] A. Filos, P. Tigkas, R. McAllister, N. Rhinehart, S. Levine, and Y. Gal, “Can autonomous vehicles identify, recover from, and adapt to distribution shifts?” in *International Conference on Machine Learning*. PMLR, 2020, pp. 3145–3153.
- [25] X. B. Peng, M. Andrychowicz, W. Zaremba, and P. Abbeel, “Sim-to-real transfer of robotic control with dynamics randomization,” in *2018 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2018, pp. 3803–3810.
- [26] Y. Du, O. Watkins, T. Darrell, P. Abbeel, and D. Pathak, “Auto-tuned sim-to-real transfer,” in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 1290–1296.
- [27] D. Ho, K. Rao, Z. Xu, E. Jang, M. Khansari, and Y. Bai, “RetinaGAN: An object-aware approach to sim-to-real transfer,” in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 10920–10926.
- [28] D. Pfommer, T. Zhang, S. Tu, and N. Matni, “Tasil: Taylor series imitation learning,” *Advances in Neural Information Processing Systems*, vol. 35, pp. 20 162–20 174, 2022.
- [29] C. Villani, *Optimal transport: Old and new*. Springer, 2009, vol. 338.
- [30] N. Hovakimyan and C. Cao, \mathcal{L}_1 adaptive control theory: Guaranteed robustness with fast adaptation. SIAM, 2010.
- [31] I. Gregory, E. Xargay, C. Cao, and N. Hovakimyan, “Flight test of an \mathcal{L}_1 adaptive controller on the NASA AirSTAR flight test vehicle,” in *Proceedings of AIAA Guidance, Navigation, and Control Conference*, Toronto, Ontario, Canada, August 2010, AIAA 2010-8015.
- [32] K. A. Ackerman, E. Xargay, R. Choe, N. Hovakimyan, M. C. Cotting, R. B. Jeffrey, M. P. Blackstun, T. P. Fulkerson, T. R. Lau, and S. S. Stephens, “Evaluation of an \mathcal{L}_1 adaptive flight control law on Calspan’s variable-stability Learjet,” *AIAA Journal of Guidance, Control, and Dynamics*, vol. 40, no. 4, pp. 1051–1060, 2017.

- [33] I. Kaminer, A. Pascoal, E. Xargay, N. Hovakimyan, C. Cao, and V. Dobrokhodov, “Path following for small unmanned aerial vehicles using \mathcal{L}_1 adaptive augmentation of commercial autopilots,” *AIAA Journal of Guidance, Control, and Dynamics*, vol. 33, no. 2, pp. 550–564, 2010.
- [34] H. Jafarnejadsani, D. Sun, H. Lee, and N. Hovakimyan, “Optimized \mathcal{L}_1 adaptive controller for trajectory tracking of an indoor quadrotor,” *AIAA Journal of Guidance, Control, and Dynamics*, vol. 40, no. 6, pp. 1415–1427, 2017.
- [35] G. C. Pflug, “Stochastic minimization with constant step-size: Asymptotic laws,” *SIAM Journal on Control and Optimization*, vol. 24, no. 4, pp. 655–666, 1986.
- [36] A. Benveniste and G. Ruget, “A measure of the tracking capability of recursive stochastic algorithms with constant gains,” *IEEE Transactions on Automatic Control*, vol. 27, no. 3, pp. 639–649, 1982.
- [37] B. Kumar, V. Borkar, and A. Shetty, “Non-asymptotic error bounds for constant stepsize stochastic approximation for tracking mobile agents,” *Mathematics of Control, Signals, and Systems*, vol. 31, pp. 589–614, 2019.
- [38] Q.-C. Pham, N. Tabareau, and J.-J. Slotine, “A contraction theory approach to stochastic incremental stability,” *IEEE Transactions on Automatic Control*, vol. 54, no. 4, pp. 816–820, 2009.
- [39] E. Mazumdar, T. Westenbroek, M. I. Jordan, and S. S. Sastry, “High confidence sets for trajectories of stochastic time-varying nonlinear systems,” in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 4275–4280.
- [40] C. Lakshminarayanan and C. Szepesvari, “Linear stochastic approximation: How far does constant step-size and iterate averaging go?” in *International Conference on Artificial Intelligence and Statistics*. PMLR, 2018, pp. 1347–1355.
- [41] E. D. Sontag, “A Lyapunov-like characterization of asymptotic controllability,” *SIAM journal on control and optimization*, vol. 21, no. 3, pp. 462–471, 1983.
- [42] H. K. Khalil, *Nonlinear systems*. London, UK: Pearson, 2002.
- [43] H. Deng, M. Krstic, and R. J. Williams, “Stabilization of stochastic nonlinear systems driven by noise of unknown covariance,” *IEEE Transactions on automatic control*, vol. 46, no. 8, pp. 1237–1253, 2001.
- [44] S.-J. Liu and J.-F. Zhang, “Output-feedback control of a class of stochastic nonlinear systems with linearly bounded unmeasurable states,” *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal*, vol. 18, no. 6, pp. 665–687, 2008.
- [45] H. Deng and M. Krstić, “Stochastic nonlinear stabilization—I: A backstepping design,” *Systems & Control Letters*, vol. 32, no. 3, pp. 143–150, 1997.
- [46] H. Deng and M. Krstic, “Output-feedback stochastic nonlinear stabilization,” *IEEE Transactions on Automatic Control*, vol. 44, no. 2, pp. 328–333, 1999.
- [47] H. J. Kushner, *Stochastic stability and control*. Academic Press, 1967.
- [48] R. Khasminskii, *Stochastic stability of differential equations*. Springer Science & Business Media, 2011, vol. 66.
- [49] W. Lohmiller and J.-J. E. Slotine, “On contraction analysis for non-linear systems,” *Automatica*, vol. 34, no. 6, pp. 683–696, 1998.
- [50] D. Angeli, “A Lyapunov approach to incremental stability properties,” *IEEE Transactions on Automatic Control*, vol. 47, no. 3, pp. 410–421, 2002.
- [51] S. Singh, A. Majumdar, J.-J. Slotine, and M. Pavone, “Robust online motion planning via contraction theory and convex optimization,” in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 5883–5890.
- [52] I. R. Manchester and J.-J. E. Slotine, “Control contraction metrics: Convex and intrinsic criteria for nonlinear feedback design,” *IEEE Transactions on Automatic Control*, vol. 62, no. 6, pp. 3046–3053, 2017.
- [53] A. P. Dani, S.-J. Chung, and S. Hutchinson, “Observer design for stochastic nonlinear systems via contraction-based incremental stability,” *IEEE Transactions on Automatic Control*, vol. 60, no. 3, pp. 700–714, 2014.
- [54] H. Tsukamoto, S.-J. Chung, and J.-J. E. Slotine, “Contraction theory for nonlinear stability analysis and learning-based control: A tutorial overview,” *Annual Reviews in Control*, vol. 52, pp. 135–169, 2021.
- [55] H. Tsukamoto and S.-J. Chung, “Robust controller design for stochastic nonlinear systems via convex optimization,” *IEEE Transactions on Automatic Control*, vol. 66, no. 10, pp. 4731–4746, 2020.

- [56] H. Tsukamoto, S.-J. Chung, and J.-J. E. Slotine, “Neural stochastic contraction metrics for learning-based control and estimation,” *IEEE Control Systems Letters*, vol. 5, no. 5, pp. 1825–1830, 2020.
- [57] H. Min, S. Xu, B. Zhang, and Q. Ma, “Globally adaptive control for stochastic nonlinear time-delay systems with perturbations and its application,” *Automatica*, vol. 102, pp. 105–110, 2019.
- [58] H.-B. Ji and H.-S. Xi, “Adaptive output-feedback tracking of stochastic nonlinear systems,” *IEEE Transactions on Automatic Control*, vol. 51, no. 2, pp. 355–360, 2006.
- [59] H. Tsukamoto, S.-J. Chung, and J.-J. Slotine, “Learning-based adaptive control using contraction theory,” in *2021 60th IEEE Conference on Decision and Control (CDC)*. IEEE, 2021, pp. 2533–2538.
- [60] I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*. Springer Science & Business Media, 1991, vol. 113.
- [61] S. Richards, N. Azizan, J.-J. Slotine, and M. Pavone, “Adaptive-control-oriented meta-learning for nonlinear systems,” in *Robotics science and systems*, 2021.
- [62] H. Sun, Z. Li, N. Hovakimyan, and T. Başar, “Performance prediction of \mathcal{L}_1 adaptive control in linear stochastic systems,” in *52nd IEEE Conference on Decision and Control*. IEEE, 2013, pp. 1–6.
- [63] K. Kang, P. Gradu, J. J. Choi, M. Janner, C. Tomlin, and S. Levine, “Lyapunov density models: Constraining distribution shift in learning-based control,” in *International Conference on Machine Learning*. PMLR, 2022, pp. 10 708–10 733.
- [64] H. Seo, D. Lee, C. Y. Son, I. Jang, C. J. Tomlin, and H. J. Kim, “Real-time robust receding horizon planning using Hamilton–Jacobi reachability analysis,” *IEEE Transactions on Robotics*, vol. 39, no. 1, pp. 90–109, 2022.
- [65] C. Knuth, G. Chou, J. Reese, and J. Moore, “Statistical safety and robustness guarantees for feedback motion planning of unknown underactuated stochastic systems,” in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 12 700–12 706.
- [66] D. Applebaum, *Lévy processes and stochastic calculus*. Cambridge university press, 2009.
- [67] L. Hewing and M. N. Zeilinger, “Stochastic model predictive control for linear systems using probabilistic reachable sets,” in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 5182–5188.
- [68] —, “Scenario-based probabilistic reachable sets for recursively feasible stochastic model predictive control,” *IEEE Control Systems Letters*, vol. 4, no. 2, pp. 450–455, 2019.
- [69] M. Fiacchini and T. Alamo, “Probabilistic reachable and invariant sets for linear systems with correlated disturbance,” *Automatica*, vol. 132, p. 109808, 2021.
- [70] A. P. Vinod and M. M. Oishi, “Stochastic reachability of a target tube: Theory and computation,” *Automatica*, vol. 125, p. 109458, 2021.
- [71] K. Sivaramakrishnan, V. Sivaramakrishnan, and M. M. Oishi, “Stochastic reachability of discrete-time stochastic systems via probability measures,” *arXiv preprint arXiv:2304.00598*, 2023.
- [72] Z. Liu, S. Jafarpour, and Y. Chen, “Safety verification of stochastic systems: A set-erosion approach,” *IEEE Control Systems Letters*, 2024.
- [73] S. Jafarpour, Z. Liu, and Y. Chen, “Probabilistic reachability analysis of stochastic control systems,” *IEEE Transactions on Automatic Control*, 2025.
- [74] F. Lin, X. Fang, and Z. Gao, “Distributionally robust optimization: A review on theory and applications,” *Numerical Algebra, Control and Optimization*, vol. 12, no. 1, pp. 159–212, 2022.
- [75] H. Rahimian and S. Mehrotra, “Distributionally robust optimization: A review,” *arXiv preprint arXiv:1908.05659*, 2019.
- [76] P. Mohajerin Esfahani and D. Kuhn, “Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations,” *Mathematical Programming*, vol. 171, no. 1-2, pp. 115–166, 2018.
- [77] J. Coulson, J. Lygeros, and F. Dörfler, “Regularized and distributionally robust data-enabled predictive control,” in *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 2696–2701.
- [78] A. Dixit, M. Ahmadi, and J. W. Burdick, “Distributionally robust model predictive control with total variation distance,” *IEEE Control Systems Letters*, vol. 6, pp. 3325–3330, 2022.
- [79] V. Renganathan, I. Shames, and T. H. Summers, “Towards integrated perception and motion planning with distributionally robust risk constraints,” *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 15 530–15 536, 2020.

- [80] Z. Zhong, E. A. del Rio-Chanona, and P. Petsagkourakis, “Distributionally robust MPC for nonlinear systems,” *IFAC-PapersOnLine*, vol. 55, no. 7, pp. 606–613, 2022.
- [81] I. Yang, “Wasserstein distributionally robust stochastic control: A data-driven approach,” *IEEE Transactions on Automatic Control*, vol. 66, no. 8, pp. 3863–3870, 2020.
- [82] A. Hakobyan and I. Yang, “Distributionally robust differential dynamic programming with Wasserstein distance,” *IEEE Control Systems Letters*, 2023.
- [83] —, “Wasserstein distributionally robust control of partially observable linear stochastic systems,” *IEEE Transactions on Automatic Control (to appear)*, *arXiv preprint arXiv:2212.04644*, 2022.
- [84] M. Schuurmans and P. Patrinos, “Learning-based distributionally robust model predictive control of Markovian switching systems with guaranteed stability and recursive feasibility,” in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 4287–4292.
- [85] A. Farid, S. Veer, and A. Majumdar, “Task-driven out-of-distribution detection with statistical guarantees for robot learning,” in *Conference on Robot Learning*. PMLR, 2022, pp. 970–980.
- [86] A. Hakobyan and I. Yang, “Toward improving the distributional robustness of risk-aware controllers in learning-enabled environments,” in *2021 60th IEEE Conference on Decision and Control (CDC)*. IEEE, 2021, pp. 6024–6031.
- [87] A. Z. Ren and A. Majumdar, “Distributionally robust policy learning via adversarial environment generation,” *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 1379–1386, 2022.
- [88] A. Hakobyan and I. Yang, “Wasserstein distributionally robust motion control for collision avoidance using conditional value-at-risk,” *IEEE Transactions on Robotics*, vol. 38, no. 2, pp. 939–957, 2021.
- [89] R. Durrett, *Stochastic calculus: A practical introduction*. CRC press, 1996, vol. 6.
- [90] H. Kuo, *Introduction to Stochastic Integration*. Springer New York, 2006.
- [91] X. Mao, *Stochastic differential equations and applications*. Elsevier, 2007.
- [92] B. E. Fristedt and L. F. Gray, *A modern approach to probability theory*. Springer Science & Business Media, 2013.
- [93] R. B. Ash and M. F. Gardner, *Topics in Stochastic Processes: Probability and Mathematical Statistics: A Series of Monographs and Textbooks*. Academic press, 2014, vol. 27.
- [94] A. V. Balakrishnan, *Stochastic Differential Systems I: Filtering and Control-A Function Space Approach*. Springer Science & Business Media, 2012, vol. 84.
- [95] F. Bullo, *Contraction Theory for Dynamical Systems*, 1.2 ed. Kindle Direct Publishing, 2024. [Online]. Available: <https://fbullo.github.io/ctds>
- [96] A. Davydov, S. Jafarpour, and F. Bullo, “Non-Euclidean contraction theory for robust nonlinear stability,” *IEEE Transactions on Automatic Control*, vol. 67, no. 12, pp. 6667–6681, 2022.
- [97] E. D. Sontag, “Contractive systems with inputs,” in *Perspectives in Mathematical System Theory, Control, and Signal Processing*. Springer, 2010, pp. 217–228.
- [98] Z. Aminzare and E. D. Sontag, “Contraction methods for nonlinear systems: A brief introduction and some open problems,” in *53rd IEEE Conference on Decision and Control*. IEEE, 2014, pp. 3835–3847.
- [99] D. Angeli, “Further results on incremental input-to-state stability,” *IEEE Transactions on Automatic Control*, vol. 54, no. 6, pp. 1386–1391, 2009.
- [100] F. Forni and R. Sepulchre, “A differential Lyapunov framework for contraction analysis,” *IEEE transactions on automatic control*, vol. 59, no. 3, pp. 614–628, 2013.
- [101] D. Wu and G.-R. Duan, “Further geometric and Lyapunov characterizations of incrementally stable systems on Finsler manifolds,” *IEEE Transactions on Automatic Control*, vol. 67, no. 10, pp. 5614–5621, 2021.
- [102] J. W. Simpson-Porco and F. Bullo, “Contraction theory on Riemannian manifolds,” *Systems & Control Letters*, vol. 65, pp. 74–80, 2014.
- [103] N. Hovakimyan and C. Cao, *\mathcal{L}_1 adaptive control theory: Guaranteed robustness with fast adaptation*. SIAM, 2010.
- [104] A. Lakshmanan, A. Gahlawat, and N. Hovakimyan, “Safe feedback motion planning: A contraction theory and \mathcal{L}_1 -adaptive control based approach,” in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 1578–1583.

- [105] L. C. Evans, *Partial differential equations*. American Mathematical Society, 2022, vol. 19.
- [106] I. I. Gikhman and A. V. Skorokhod, *Introduction to the theory of random processes*. W.B. Saunders Co., Philadelphia, 1969.
- [107] L. C. Evans, *An introduction to stochastic differential equations*. American Mathematical Soc., 2012, vol. 82.
- [108] K. B. Athreya and S. N. Lahiri, *Measure theory and probability theory*. Springer, 2006, vol. 19.
- [109] J. Bouvrie and J.-J. Slotine, “Wasserstein contraction of stochastic nonlinear systems,” *arXiv preprint arXiv:1902.08567*, 2019.
- [110] X. Wang and N. Hovakimyan, “ \mathcal{L}_1 adaptive controller for nonlinear time-varying reference systems,” *Systems & Control Letters*, vol. 61, no. 4, pp. 455–463, 2012.
- [111] P. Zhao, S. Snyder, N. Hovakimyan, and C. Cao, “Robust adaptive control of linear parameter-varying systems with unmatched uncertainties,” *Journal of Guidance, Control, and Dynamics*, vol. 47, no. 10, pp. 2085–2102, 2024.
- [112] P. Zhao, Y. Mao, C. Tao, N. Hovakimyan, and X. Wang, “Adaptive robust quadratic programs using control Lyapunov and barrier functions,” in *59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 3353–3358.
- [113] F. E. Burk, *A garden of integrals*. American Mathematical Soc., 2007, vol. 31.
- [114] G. Hardy, J. Littlewood, and G. Pólya, *Inequalities*. Cambridge University Press, 1952.

Appendices

A Constants

In this section we collect cumbersome constants that we use throughout the manuscript. The subsequent definitions use the constants $\Delta_{(\cdot)}$ defined in Assumptions 1-8 in Sec. 2.2. We first define a few constants that are used globally in the manuscript:

$$p(p) = \left(p \frac{p-1}{2}\right)^{\frac{1}{2}}, \quad p'(p) = \left(p \frac{2p-1}{2}\right)^{\frac{1}{2}}, \quad p''(p) = (p(4p-1))^{\frac{1}{2}}, \quad (\text{A.1a})$$

$$\mathbb{1}_{\text{Lip}}(f) \doteq \begin{cases} 1, & \text{if Assumption 8 holds} \\ 0, & \text{otherwise} \end{cases}. \quad (\text{A.1b})$$

Next, we provide the definitions of constants and maps bespoke to the analysis of the reference process in Sec. 4.1 and Appendix C.

A.1 Reference Process Analysis

We begin with the definition of constants $\hat{\Delta}_1^r$, $\hat{\Delta}_2^r$, $\hat{\Delta}_3^r$, and $\hat{\Delta}_4^r$:

$$\hat{\Delta}_1^r(p) = \Delta_g \left(\frac{1}{\sqrt{\lambda}} [\Delta_f (2 + \Delta_*) (1 - \mathbb{1}_{\text{Lip}}(f)) + \Delta_\mu] + p(p) (2\Delta_p + \Delta_\sigma) \right), \quad (\text{A.2a})$$

$$\hat{\Delta}_2^r(p) = \Delta_g p(p) \Delta_\sigma, \quad \hat{\Delta}_3^r = \frac{1}{\sqrt{\lambda}} \Delta_g (\Delta_f (1 - \mathbb{1}_{\text{Lip}}(f)) + \Delta_\mu), \quad \hat{\Delta}_4^r = \frac{1}{\sqrt{\lambda}} (\Delta_g L_f \mathbb{1}_{\text{Lip}}(f) + \Delta_{\hat{g}}), \quad (\text{A.2b})$$

where $\mathbb{1}_{\text{Lip}}(f)$ is defined above in (A.1b).

Next, we define the constants Δ_{\circ}^r , Δ_{\circ}^r , Δ_{\circ}^r , Δ_{\circ}^r , Δ_{\circ}^r . In addition to the constants $\Delta_{(\cdot)}$ and λ defined in Assumptions 1-5 in Sec. 2.2, the following also use the constants in (A.2) above. We first begin with $\Delta_{\circ_i}^r$, $i \in \{1, \dots, 4\}$, that are defined as follows:

$$\Delta_{\circ_1}^r = \Delta_p^2 + (\Delta_p + \Delta_\sigma)^2, \quad \Delta_{\circ_2}^r(p) = \frac{\Delta_\mu^\parallel}{\sqrt{\lambda}} \left(\hat{\Delta}_1^r(p) + \frac{\Delta_g^2 \Delta_\mu^\parallel}{\sqrt{\lambda}} \right), \quad (\text{A.3a})$$

$$\Delta_{\circ_3}^r(p) = \frac{p'(p)}{\sqrt{\lambda}} (\Delta_p^\parallel + \Delta_\sigma^\parallel) \left(\hat{\Delta}_1^r(p) + \frac{2\Delta_g^2 \Delta_\mu^\parallel}{\sqrt{\lambda}} \right), \quad (\text{A.3b})$$

$$\Delta_{\odot_4}^r(\mathbf{p}) = (\Delta_p^\parallel + \Delta_\sigma^\parallel) \frac{\Delta_g}{\lambda} (\mathbf{p}'(\mathbf{p})^2 \Delta_g (\Delta_p^\parallel + \Delta_\sigma^\parallel) + \sqrt{m} (2\Delta_p + \Delta_\sigma)). \quad (\text{A.3c})$$

Next, we define the constants $\Delta_{\odot_i}^r, i \in \{1, \dots, 4\}$, as follows:

$$\Delta_{\odot_1}^r = 2\Delta_\sigma (\Delta_p + \Delta_\sigma), \quad \Delta_{\odot_2}^r(\mathbf{p}) = \frac{\Delta_\mu^\parallel}{\sqrt{\lambda}} \hat{\Delta}_2^r(\mathbf{p}), \quad (\text{A.4a})$$

$$\Delta_{\odot_3}^r(\mathbf{p}) = \frac{\mathbf{p}'(\mathbf{p})}{\sqrt{\lambda}} \left((\Delta_p^\parallel + \Delta_\sigma^\parallel) \hat{\Delta}_2^r(\mathbf{p}) + \Delta_\sigma^\parallel \left(\hat{\Delta}_1^r(\mathbf{p}) + \frac{2\Delta_g^2 \Delta_\mu^\parallel}{\sqrt{\lambda}} \right) \right), \quad (\text{A.4b})$$

$$\Delta_{\odot_4}^r(\mathbf{p}) = \frac{\Delta_g}{\lambda} \left((\Delta_p^\parallel + \Delta_\sigma^\parallel) (2\mathbf{p}'(\mathbf{p})^2 \Delta_g \Delta_\sigma^\parallel + \sqrt{m} \Delta_\sigma) + \sqrt{m} \Delta_\sigma^\parallel (2\Delta_p + \Delta_\sigma) \right). \quad (\text{A.4c})$$

The constants $\Delta_{\odot_i}^r, i \in \{1, \dots, 8\}$, are defined as:

$$\Delta_{\odot_1}^r = \Delta_\sigma^2, \quad \Delta_{\odot_2}^r = 2\Delta_g^\perp \Delta_\mu^\perp, \quad \Delta_{\odot_3}^r(\mathbf{p}) = 2\mathbf{p}(\mathbf{p}) (\Delta_g^\perp (\Delta_p^\perp + \Delta_\sigma^\perp) + \Delta_p), \quad (\text{A.5a})$$

$$\Delta_{\odot_4}^r(\mathbf{p}) = \frac{\Delta_\mu^\parallel}{\sqrt{\lambda}} \left(\hat{\Delta}_1^r(\mathbf{p}) + \hat{\Delta}_3^r + \frac{2\Delta_g^2 \Delta_\mu^\parallel}{\sqrt{\lambda}} \right), \quad \Delta_{\odot_5}^r(\mathbf{p}) = \Delta_\mu^\parallel \left(\frac{\hat{\Delta}_4^r}{\sqrt{\lambda}} + 4\Delta_g \right) + 2\sqrt{\lambda} \Delta_g \mathbf{p}(\mathbf{p}) (\Delta_p^\parallel + \Delta_\sigma^\parallel), \quad (\text{A.5b})$$

$$\Delta_{\odot_6}^r(\mathbf{p}) = \frac{\mathbf{p}'(\mathbf{p})}{\sqrt{\lambda}} \left((\Delta_p^\parallel + \Delta_\sigma^\parallel) \left(\hat{\Delta}_3^r + \frac{2\Delta_g^2 \Delta_\mu^\parallel}{\sqrt{\lambda}} \right) + \Delta_\sigma^\parallel \hat{\Delta}_2^r(\mathbf{p}) \right), \quad (\text{A.5c})$$

$$\Delta_{\odot_7}^r(\mathbf{p}) = \mathbf{p}'(\mathbf{p}) (\Delta_p^\parallel + \Delta_\sigma^\parallel) \left(\frac{\hat{\Delta}_4^r}{\sqrt{\lambda}} + 2\Delta_g \right), \quad \Delta_{\odot_8}^r(\mathbf{p}) = \Delta_\sigma^\parallel \frac{\Delta_g}{\lambda} (\mathbf{p}'(\mathbf{p})^2 \Delta_g \Delta_\sigma^\parallel + \sqrt{m} \Delta_\sigma), \quad (\text{A.5d})$$

while the constants $\Delta_{\otimes_i}^r, i \in \{1, \dots, 5\}$, are defined as follows:

$$\Delta_{\otimes_1}^r(\mathbf{p}) = 2\mathbf{p}(\mathbf{p}) \Delta_g^\perp \Delta_\sigma^\perp, \quad \Delta_{\otimes_2}^r(\mathbf{p}) = \Delta_\mu^\parallel \frac{\hat{\Delta}_2^r(\mathbf{p})}{\sqrt{\lambda}}, \quad \Delta_{\otimes_3}^r(\mathbf{p}) = 2\mathbf{p}(\mathbf{p}) \sqrt{\lambda} \Delta_g \Delta_\sigma^\parallel, \quad (\text{A.6a})$$

$$\{\Delta_{\otimes_4}^r(\mathbf{p}), \Delta_{\otimes_5}^r(\mathbf{p})\} = \mathbf{p}'(\mathbf{p}) \Delta_\sigma^\parallel \odot \left\{ \frac{1}{\sqrt{\lambda}} \left(\hat{\Delta}_3^r + \frac{2\Delta_g^2 \Delta_\mu^\parallel}{\sqrt{\lambda}} \right), \left(\frac{\hat{\Delta}_4^r}{\sqrt{\lambda}} + 2\Delta_g \right) \right\}. \quad (\text{A.6b})$$

Finally, we define the constants $\Delta_{\oplus_i}^r, i \in \{1, \dots, 3\}$, as follows:

$$\Delta_{\oplus_1}^r = 2\Delta_g^\perp \Delta_\mu^\perp, \quad \{\Delta_{\oplus_2}^r, \Delta_{\oplus_3}^r\} = \Delta_\mu^\parallel \odot \left\{ \frac{1}{\sqrt{\lambda}} \left(\hat{\Delta}_3^r + \frac{\Delta_g^2 \Delta_\mu^\parallel}{\sqrt{\lambda}} \right), \frac{\hat{\Delta}_4^r}{\sqrt{\lambda}} + 4\Delta_g \right\}. \quad (\text{A.7a})$$

A.2 True Process Analysis

Similar to the definitions in Sec. A.1 above, we collect constants used in the analysis of the true process, starting with $\hat{\Delta}_i, i \in \{1, \dots, 5\}$:

$$\hat{\Delta}_1(\mathbf{p}) = \frac{2}{\sqrt{\lambda}} \Delta_g \Delta_f (1 - \mathbb{1}_{\text{Lip}}(f)), \quad (\text{A.8a})$$

$$\hat{\Delta}_2(\mathbf{p}) = \Delta_g \mathbf{p}(\mathbf{p}) (L_p + L_\sigma), \quad \hat{\Delta}_3 = \frac{1}{\sqrt{\lambda}} \Delta_g \Delta_f (1 - \mathbb{1}_{\text{Lip}}(f)), \quad (\text{A.8b})$$

$$\hat{\Delta}_4 = \frac{1}{\sqrt{\lambda}} (\Delta_g (L_\mu + L_f \mathbb{1}_{\text{Lip}}(f)) + \Delta_{\dot{g}}), \quad \hat{\Delta}_5 = \sqrt{m} \Delta_g (L_p + L_\sigma). \quad (\text{A.8c})$$

Next, we define Δ_{\odot_i} and $\Delta_{\odot_i}, i \in \{1, \dots, 4\}$ as follows:

$$\Delta_{\odot_1}(\mathbf{p}) = \frac{1}{\sqrt{\lambda}} \mathbf{p}'(\mathbf{p}) (L_p^\parallel + L_\sigma^\parallel) \hat{\Delta}_1(\mathbf{p}), \quad (\text{A.9})$$

and

$$\Delta_{\odot_1} = (L_p + L_\sigma)^2, \quad \Delta_{\odot_2}(\mathbf{p}) = \frac{1}{\sqrt{\lambda}} L_\mu^\parallel \hat{\Delta}_1(\mathbf{p}), \quad \Delta_{\odot_3}(\mathbf{p}) = \frac{1}{\sqrt{\lambda}} \mathbf{p}'(\mathbf{p}) (L_p^\parallel + L_\sigma^\parallel) \hat{\Delta}_2(\mathbf{p}), \quad (\text{A.10a})$$

$$\Delta_{\odot_4}(\mathbf{p}) = \frac{1}{\lambda} (L_p^\parallel + L_\sigma^\parallel) \left(\hat{\Delta}_6 + \Delta_g^2 \mathbf{p}'(\mathbf{p})^2 (L_p^\parallel + L_\sigma^\parallel) \right). \quad (\text{A.10b})$$

Finally, the constants $\Delta_{\otimes_i}, i \in \{1, \dots, 4\}$, and $\Delta_{\oplus_j}, j \in \{1, 2, 3\}$, are defined as

$$\Delta_{\otimes_1}(\mathbf{p}) = 2\Delta_g^\perp \mathbf{p}(\mathbf{p}) (L_p^\perp + L_\sigma^\perp), \quad \Delta_{\otimes_2}(\mathbf{p}) = 2\sqrt{\lambda} \Delta_g \mathbf{p}(\mathbf{p}) (L_p^\parallel + L_\sigma^\parallel) + L_\mu^\parallel \frac{1}{\sqrt{\lambda}} \hat{\Delta}_2(\mathbf{p}), \quad (\text{A.11a})$$

$$\Delta_{\otimes_3}(\mathbf{p}) = \frac{1}{\sqrt{\lambda}} \mathbf{p}'(\mathbf{p}) (L_p^\parallel + L_\sigma^\parallel) \hat{\Delta}_3, \quad (\text{A.11b})$$

$$\Delta_{\otimes_4}(\mathbf{p}) = \mathbf{p}'(\mathbf{p}) (L_p^\parallel + L_\sigma^\parallel) \left(\frac{1}{\sqrt{\lambda}} \hat{\Delta}_4 + 2\Delta_g \left(1 + \frac{\Delta_g}{\lambda} L_\mu^\parallel \right) \right), \quad (\text{A.11c})$$

and

$$\Delta_{\oplus_1} = 2\Delta_g^\perp L_\mu^\perp, \quad \Delta_{\oplus_2} = \frac{1}{\sqrt{\lambda}} L_\mu^\parallel \hat{\Delta}_3, \quad \Delta_{\oplus_3} = L_\mu^\parallel \left(\frac{1}{\sqrt{\lambda}} \hat{\Delta}_4 + \Delta_g \left(4 + \frac{\Delta_g}{\lambda} L_\mu^\parallel \right) \right). \quad (\text{A.12})$$

A.2.1 Piecewise Constant Adaptation Law

In the following we define constants and maps related to the piecewise constant adaptation law in (15b) (and (95)). Note the explicit dependence of the following entities on the control parameters $\omega, T_s, \lambda_s \in \mathbb{R}_{>0}$ (see (13) - (15)).

We begin with the definitions of $\gamma_i, \gamma'_i, i \in \{1, 2\}$, and γ'' :

$$\gamma_1(T_s) = \left(\lambda_s \frac{e^{\lambda_s T_s} + 1}{e^{\lambda_s T_s} - 1} \right)^{\frac{1}{2}}, \quad \gamma'_1(\omega, T_s) = (1 - e^{-2\lambda T_s}) (1 - e^{-\omega T_s}) \quad (\text{A.13a})$$

$$\gamma_2(\omega, T_s) = \max \left\{ e^{(\omega - \lambda_s) T_s}, 1 \right\} \frac{\lambda_s}{\omega} \frac{e^{\omega T_s} - 1}{e^{\lambda_s T_s} - 1}, \quad (\text{A.13b})$$

$$\gamma'_2(\omega, T_s) = \max \left\{ \left| 1 - e^{(\omega - \lambda_s) T_s} \frac{\lambda_s}{\omega} \frac{e^{\omega T_s} - 1}{e^{\lambda_s T_s} - 1} \right|, \left| 1 - \frac{\lambda_s}{\omega} \frac{e^{\omega T_s} - 1}{e^{\lambda_s T_s} - 1} \right| \right\}, \quad (\text{A.13c})$$

$$\gamma''(\mathbf{p}, \omega, T_s) = \mathbf{p}'(\mathbf{p}) \gamma'_2(\omega, T_s) + \mathbf{p}''(\mathbf{p}) (1 - e^{-2\omega T_s})^{\frac{1}{2}} (2 + \gamma_2(\omega, T_s)). \quad (\text{A.13d})$$

Similar to the above, we define

$$\gamma_\mu(\omega, T_s) = L_\mu^\parallel + L_\mu \gamma_2(\omega, T_s) \Delta_\Theta, \quad \hat{\gamma}_\mu(\omega, T_s) = \hat{L}_\mu^\parallel + \hat{L}_\mu \gamma_2(\omega, T_s) \Delta_\Theta, \quad (\text{A.14a})$$

$$\gamma_\sigma(\omega, T_s) = L_p^\parallel + L_\sigma^\parallel + \gamma_2(\omega, T_s) \Delta_\Theta (L_p + L_\sigma), \quad (\text{A.14b})$$

$$\hat{\gamma}_\sigma(\omega, T_s) = \hat{L}_p^\parallel + \hat{L}_\sigma^\parallel + \gamma_2(\omega, T_s) \Delta_\Theta (\hat{L}_p + \hat{L}_\sigma). \quad (\text{A.14c})$$

Next, the constants $\bar{\Delta}_i, i \in \{1, \dots, 3\}$, are defined as follows:

$$\bar{\Delta}_1(T_s, \mathbf{p}) = 2T_s [\Delta_f + (1 + \Delta_g \Delta_\Theta e^{-\lambda_s T_s}) \Delta_\mu + \mathbf{p}'(\mathbf{p}) \Delta_g \Delta_\Theta e^{-\lambda_s T_s} \gamma_1(T_s) (\Delta_p + \Delta_\sigma)] \\ + 2\sqrt{2} \sqrt{T_s} \mathbf{p}'(\mathbf{p}) (\Delta_p + \Delta_\sigma), \quad (\text{A.15a})$$

$$\bar{\Delta}_2(T_s, \mathbf{p}) = 2\mathbf{p}'(\mathbf{p}) \Delta_\sigma \left(\sqrt{2} \sqrt{T_s} + T_s \Delta_g \Delta_\Theta e^{-\lambda_s T_s} \gamma_1(T_s) \right), \quad (\text{A.15b})$$

$$\bar{\Delta}_3(T_s) = 2T_s (\Delta_f + (1 + \Delta_g \Delta_\Theta e^{-\lambda_s T_s}) \Delta_\mu). \quad (\text{A.15c})$$

The final set of constants Δ_{μ_i} and $\Delta_{\sigma_i}, i \in \{1, \dots, 3\}$, are defined as follows:

$$\Delta_{\mu_1}(\mathbf{p}, \omega, T_s) = T_s \hat{\gamma}_\mu(\omega, T_s) + \bar{\Delta}_1(T_s, \mathbf{p}) \gamma_\mu(\omega, T_s) + \Delta_\mu^\parallel (1 - e^{-\lambda_s T_s}), \quad (\text{A.16a})$$

$$\Delta_{\mu_2}(\mathbf{p}, \omega, T_s) = \bar{\Delta}_2(T_s, \mathbf{p}) \gamma_\mu(\omega, T_s), \quad (\text{A.16b})$$

$$\Delta_{\mu_3}(\mathbf{p}, \omega, T_s) = T_s \gamma_\mu(\omega, T_s) \bar{\Delta}_3(T_s) + \Delta_\mu^\parallel (1 - e^{-\lambda_s T_s}), \quad (\text{A.16c})$$

and

$$\Delta_{\sigma_1}(\mathbf{p}, \omega, T_s) = \left[(L_p^\parallel + L_\sigma^\parallel) (\bar{\Delta}_1(T_s, \mathbf{p}))^{\frac{1}{2}} + T_s (\hat{L}_p^\parallel + \hat{L}_\sigma^\parallel) \right] (1 + \gamma_2(\omega, T_s)) \\ + \gamma_\sigma(\omega, T_s) (\bar{\Delta}_1(T_s, \mathbf{p}))^{\frac{1}{2}} + T_s \hat{\gamma}_\sigma(\omega, T_s), \quad (\text{A.17a})$$

$$\Delta_{\{\sigma_2, \sigma_3\}}(\mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s) = (\bar{\Delta}_{\{2,3\}}(\mathbf{T}_s, \mathbf{p}))^{\frac{1}{2}} [\gamma_{\sigma}(\boldsymbol{\omega}, \mathbf{T}_s) + (1 + \gamma_2(\boldsymbol{\omega}, \mathbf{T}_s)) (L_p^{\parallel} + L_{\sigma}^{\parallel})]. \quad (\text{A.17b})$$

Next, we define the mappings $\Upsilon'_k(\cdot; \mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s) : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$, $k \in \{1, 2, 3\}$, as follows

$$\begin{aligned} \Upsilon'_1(\xi; \mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s) &= \Delta_{\mu_1}(\mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s) + \sqrt{\boldsymbol{\omega}} \mathbf{p}'(\mathbf{p}) \Delta_{\sigma_1}(\mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s) + \sqrt{\boldsymbol{\omega}} \mathbf{p}'(\mathbf{p}) \Delta_{\sigma_2}(\mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s) (\xi)^{\frac{1}{4}} \\ &\quad + (\Delta_{\mu_2}(\mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s) + \sqrt{\boldsymbol{\omega}} \mathbf{p}'(\mathbf{p}) \Delta_{\sigma_3}(\mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s)) (\xi)^{\frac{1}{2}} + \Delta_{\mu_3}(\mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s) \xi, \end{aligned} \quad (\text{A.18a})$$

$$\Upsilon'_2(\xi; \mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s) = \gamma''(\mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s) \left(\Delta_p^{\parallel} + \Delta_{\sigma}^{\parallel} \left(1 + (\xi)^{\frac{1}{2}} \right) \right), \quad (\text{A.18b})$$

$$\Upsilon'_3(\xi; \mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s) = \Delta_{\mu}^{\parallel} (e^{\boldsymbol{\omega} \mathbf{T}_s} - 1) (1 + \xi), \quad (\text{A.18c})$$

for any $\xi \in \mathbb{R}_{\geq 0}$. We similarly define the mappings $\{\tilde{\Upsilon}^-(\cdot; \mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s), \Upsilon^-(\cdot; \mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s)\} : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ and $\mathring{\Upsilon}(\cdot; \mathbf{p}, \boldsymbol{\omega}) : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ as follows

$$\begin{aligned} \tilde{\Upsilon}^-(\xi; \mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s) &= \frac{1}{\lambda} (1 - e^{-2\lambda \mathbf{T}_s}) \Delta_g \\ &\quad \times \left(\sqrt{\boldsymbol{\omega}} \mathbf{p}'(\mathbf{p}) (1 - e^{-2\boldsymbol{\omega} \mathbf{T}_s})^{\frac{1}{2}} \left(\Delta_p^{\parallel} + \Delta_{\sigma}^{\parallel} \left(1 + (\xi)^{\frac{1}{2}} \right) \right) + \Delta_{\mu}^{\parallel} (1 + \xi) \right), \end{aligned} \quad (\text{A.19a})$$

$$\Upsilon^-(\xi; \mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s) = \gamma'_1(\boldsymbol{\omega}, \mathbf{T}_s) \left(2\mathbf{p}'(\mathbf{p}) \left(\Delta_p^{\parallel} + \Delta_{\sigma}^{\parallel} \left(1 + (\xi)^{\frac{1}{2}} \right) \right) + \frac{\Delta_{\mu}^{\parallel}}{\sqrt{\boldsymbol{\omega}}} (1 + \xi) \right), \quad (\text{A.19b})$$

$$\mathring{\Upsilon}(\xi; \mathbf{p}, \boldsymbol{\omega}) = \Delta_g^2 \left(\mathbf{p}'(\mathbf{p}) (L_p^{\parallel} + L_{\sigma}^{\parallel}) (\xi)^{\frac{1}{2}} + \frac{1}{\sqrt{\boldsymbol{\omega}}} L_{\mu}^{\parallel} \xi \right), \quad (\text{A.19c})$$

for any $\xi \in \mathbb{R}_{\geq 0}$. We conclude the section by proving a few key but straightforward properties of the maps defined above.

Proposition A.1 *The mapping $\xi \mapsto \mathring{\Upsilon}(\xi; \mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s)$, $\mathring{\Upsilon} \in \{\Upsilon'_1, \Upsilon'_2, \Upsilon'_3, \tilde{\Upsilon}^-, \Upsilon^-, \mathring{\Upsilon}\}$, is non-decreasing over $\mathbb{R}_{\geq 0}$, i.e.*

$$\mathring{\Upsilon}(\xi_1; \mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s) \leq \mathring{\Upsilon}(\xi_2; \mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s), \quad \forall (\xi_1, \xi_2, \mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s) \in \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathbb{N} \times \mathbb{R}_{>0} \times \mathbb{R}_{>0}, \quad \xi_1 \leq \xi_2. \quad (\text{A.20})$$

Furthermore, the following limits are satisfied for any fixed $\{\xi, \mathbf{p}, \boldsymbol{\omega}\} \in \mathbb{R}_{\geq 0} \times \mathbb{N} \times \mathbb{R}_{>0}$:

$$\lim_{\mathbf{T}_s \downarrow 0} \left\{ \tilde{\Upsilon}^-, \Upsilon^-, \Upsilon'_1, \Upsilon'_2, \Upsilon'_3 \right\}(\xi; \mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s) = 0. \quad (\text{A.21})$$

Proof. The non-decreasing property in (A.20) is trivial to prove. In order to prove (A.21), we begin by considering the terms in (A.13) for any fixed $\lambda_s \in \mathbb{R}_{>0}$. We obtain the following limits in a straightforward manner:

$$\lim_{\mathbf{T}_s \downarrow 0} (\mathbf{T}_s \gamma_1(\mathbf{T}_s)) = \sqrt{\lambda_s} \lim_{\mathbf{T}_s \downarrow 0} (e^{\lambda_s \mathbf{T}_s} + 1)^{\frac{1}{2}} \lim_{\mathbf{T}_s \downarrow 0} \frac{\mathbf{T}_s}{(e^{\lambda_s \mathbf{T}_s} - 1)^{\frac{1}{2}}} = 0, \quad (\text{A.22a})$$

$$\lim_{\mathbf{T}_s \downarrow 0} \gamma'_1(\boldsymbol{\omega}, \mathbf{T}_s) = \lim_{\mathbf{T}_s \downarrow 0} (1 - e^{-2\lambda \mathbf{T}_s}) \lim_{\mathbf{T}_s \downarrow 0} (1 - e^{-\boldsymbol{\omega} \mathbf{T}_s}) = 0, \quad (\text{A.22b})$$

$$\lim_{\mathbf{T}_s \downarrow 0} \gamma_2(\boldsymbol{\omega}, \mathbf{T}_s) = \frac{\lambda_s}{\boldsymbol{\omega}} \left(\lim_{\mathbf{T}_s \downarrow 0} \max \left\{ e^{(\boldsymbol{\omega} - \lambda_s) \mathbf{T}_s}, 1 \right\} \right) \left(\lim_{\mathbf{T}_s \downarrow 0} \frac{e^{\boldsymbol{\omega} \mathbf{T}_s} - 1}{e^{\lambda_s \mathbf{T}_s} - 1} \right) = 1, \quad (\text{A.22c})$$

$$\lim_{\mathbf{T}_s \downarrow 0} \gamma'_2(\boldsymbol{\omega}, \mathbf{T}_s) = \max \left\{ \lim_{\mathbf{T}_s \downarrow 0} \left| 1 - e^{(\boldsymbol{\omega} - \lambda_s) \mathbf{T}_s} \frac{\lambda_s}{\boldsymbol{\omega}} \frac{e^{\boldsymbol{\omega} \mathbf{T}_s} - 1}{e^{\lambda_s \mathbf{T}_s} - 1} \right|, \lim_{\mathbf{T}_s \downarrow 0} \left| 1 - \frac{\lambda_s}{\boldsymbol{\omega}} \frac{e^{\boldsymbol{\omega} \mathbf{T}_s} - 1}{e^{\lambda_s \mathbf{T}_s} - 1} \right| \right\} = 0, \quad (\text{A.22d})$$

$$\lim_{\mathbf{T}_s \downarrow 0} \gamma''(\mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s) = \mathbf{p}'(\mathbf{p}) \lim_{\mathbf{T}_s \downarrow 0} \gamma'_2(\boldsymbol{\omega}, \mathbf{T}_s) + \mathbf{p}''(\mathbf{p}) \lim_{\mathbf{T}_s \downarrow 0} (1 - e^{-2\boldsymbol{\omega} \mathbf{T}_s})^{\frac{1}{2}} \lim_{\mathbf{T}_s \downarrow 0} (2 + \gamma_2(\boldsymbol{\omega}, \mathbf{T}_s)) = 0. \quad (\text{A.22e})$$

Using the limit in (A.22c), we obtain the following for the terms in (A.14):

$$\lim_{\mathbf{T}_s \downarrow 0} \gamma_{\mu}(\boldsymbol{\omega}, \mathbf{T}_s) = L_{\mu}^{\parallel} + L_{\mu} \Delta_{\Theta}, \quad \lim_{\mathbf{T}_s \downarrow 0} \hat{\gamma}_{\mu}(\boldsymbol{\omega}, \mathbf{T}_s) = \hat{L}_{\mu}^{\parallel} + \hat{L}_{\mu} \Delta_{\Theta}, \quad (\text{A.23a})$$

$$\lim_{\mathbf{T}_s \downarrow 0} \gamma_{\sigma}(\boldsymbol{\omega}, \mathbf{T}_s) = L_p^{\parallel} + L_{\sigma}^{\parallel} + \Delta_{\Theta} (L_p + L_{\sigma}), \quad \lim_{\mathbf{T}_s \downarrow 0} \hat{\gamma}_{\sigma}(\boldsymbol{\omega}, \mathbf{T}_s) = \hat{L}_p^{\parallel} + \hat{L}_{\sigma}^{\parallel} + \Delta_{\Theta} (\hat{L}_p + \hat{L}_{\sigma}). \quad (\text{A.23b})$$

Similarly, using the limit in (A.22a), we obtain the following for the terms in (A.15):

$$\lim_{\mathbf{T}_s \downarrow 0} \left\{ \bar{\Delta}_1(\mathbf{T}_s, \mathbf{p}), \bar{\Delta}_2(\mathbf{T}_s, \mathbf{p}), \bar{\Delta}_3(\mathbf{T}_s) \right\} = 0. \quad (\text{A.24})$$

We now use (A.22), (A.23), and (A.24) to conclude the following for the terms defined in (A.16) and (A.17):

$$\lim_{T_s \downarrow 0} \Delta_{\mu_i}(\mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s) = 0, \quad \lim_{T_s \downarrow 0} \Delta_{\sigma_i}(\mathbf{p}, \boldsymbol{\omega}, \mathbf{T}_s) = 0, \quad i \in \{1, 2, 3\}. \quad (\text{A.25})$$

The claim in (A.21) then follows by applying the above limits to the mappings in (A.18) and (A.19). □

B Technical Results

Remark B.1 *The results that we present in this appendix consider functions belonging to $\mathcal{M}_2^{loc}(\mathcal{X} | \mathcal{F}_t)$, for some finite-dimensional space \mathcal{X} and filtration \mathcal{F}_t , $t \in [0, T]$. However, it is straightforward to see that the results hold for the more general space $\mathcal{M}_2([0, T]; \mathcal{X} | \mathcal{F}_t)$. In fact, Appendices C and D invoke the following results since they also hold over $\mathcal{M}_2([0, \tau \wedge T]; \mathcal{X} | \mathcal{F}_t)$, where τ is any stopping-time.*

We begin with the following result that is a multidimensional version of the stochastic integration by parts formula in [91, Thm. 1.6.5].

Lemma B.1 *Consider any $L \in \mathcal{M}_2^{loc}(\mathbb{R}^{l \times l'} | \mathfrak{F}_t)$ and $S \in \mathcal{M}_2^{loc}(\mathbb{R}^{l' \times n_q} | \mathfrak{F}_t)$, for some filtration \mathfrak{F}_t on $(\Omega, \mathcal{F}, \mathbb{P})$. Then, for any n_q -dimensional \mathfrak{F}_t -adapted Brownian motion Q_t , we have that*

$$\int_0^t \int_0^\nu L(\nu) S(\beta) dQ_\beta d\nu = \int_0^t \int_\nu^t L(\beta) S(\nu) d\beta dQ_\nu \in \mathbb{R}^l, \quad (\text{B.1a})$$

$$\int_0^t \int_\nu^t L(\nu) S(\beta) dQ_\beta d\nu = \int_0^t \int_0^\nu L(\beta) S(\nu) d\beta dQ_\nu \in \mathbb{R}^l, \quad (\text{B.1b})$$

for all $t \in [0, T]$.

Proof. We provide a proof for (B.1a) only since (B.1b) can be established *mutatis mutandis*. We begin by defining

$$\hat{L}(\varsigma) \doteq \int_0^\varsigma L(\beta) d\beta \in \mathbb{R}^{l \times l'}, \quad \varsigma \in \mathbb{R}_{\geq 0}. \quad (\text{B.2})$$

As a consequence of the above definition, we have that

$$\int_\nu^t L(\beta) d\beta = \int_0^t L(\beta) d\beta - \int_0^\nu L(\beta) d\beta \doteq \hat{L}(t) - \hat{L}(\nu) \in \mathbb{R}^{l \times l'}, \quad 0 \leq \nu \leq t \leq T. \quad (\text{B.3})$$

Using the expression above, we may write the right hand side of (B.1a) as

$$\int_0^t \int_\nu^t L(\beta) S(\nu) d\beta dQ_\nu = \int_0^t \left(\int_\nu^t L(\beta) d\beta \right) S(\nu) dQ_\nu = \int_0^t \left(\hat{L}(t) - \hat{L}(\nu) \right) S(\nu) dQ_\nu \in \mathbb{R}^l,$$

which can further be re-written as

$$\begin{aligned} & \int_0^t \int_\nu^t L(\beta) S(\nu) d\beta dQ_\nu \\ &= \int_0^t \left[\left(\hat{L}_1(t) - \hat{L}_1(\nu) \right) S(\nu) dQ_\nu \quad \cdots \quad \left(\hat{L}_{l'}(t) - \hat{L}_{l'}(\nu) \right) S(\nu) dQ_\nu \right]^\top \\ &= \sum_{i=1}^{l'} \int_0^t \left[\left(\hat{L}_{1,i}(t) - \hat{L}_{1,i}(\nu) \right) S_i(\nu) dQ_\nu \quad \cdots \quad \left(\hat{L}_{l',i}(t) - \hat{L}_{l',i}(\nu) \right) S_i(\nu) dQ_\nu \right]^\top \in \mathbb{R}^l, \end{aligned} \quad (\text{B.4})$$

where $\hat{L}_{1,\dots,l'} \in \mathbb{R}^{1 \times l'}$ and $S_{1,\dots,l'} \in \mathbb{R}^{1 \times n_q}$ are the rows of $\hat{L}(\cdot) \in \mathbb{R}^{l \times l'}$ defined in (B.2) and $S(\cdot) \in \mathbb{R}^{l' \times n_q}$, respectively. Next, we define the following scalar Itô process:

$$\hat{S}_i(t) = \int_0^t S_i(\varsigma) dQ_\varsigma, \quad d\hat{S}_i(t) = S_i(t) dQ_t, \quad t \in [0, T], \quad \hat{S}_i(0) = 0, \quad i \in \{1, \dots, l'\}, \quad (\text{B.5})$$

using which we derive the following expression:

$$\begin{aligned} \int_0^t \left(\hat{L}_{j,i}(t) - \hat{L}_{j,i}(\nu) \right) S_i(\nu) dQ_\nu &= \int_0^t \left(\hat{L}_{j,i}(t) - \hat{L}_{j,i}(\nu) \right) d\hat{S}_i(\nu) \\ &= \hat{L}_{j,i}(t) \int_0^t d\hat{S}_i(\nu) - \int_0^t \hat{L}_{j,i}(\nu) d\hat{S}_i(\nu) \\ &= \hat{L}_{j,i}(t) \hat{S}_i(t) - \int_0^t \hat{L}_{j,i}(\nu) d\hat{S}_i(\nu), \quad (\text{B.6}) \end{aligned}$$

for $(j, i, t) \in \{1, \dots, l\} \times \{1, \dots, l'\} \times [0, T]$. It is clear that $\hat{L}_{j,i}(t)$ is \mathfrak{F}_t -adapted due its definition in (B.2) and the assumed adaptedness of L . Moreover, we may write $\hat{L}_{j,i}(t) = \hat{L}_{j,i}^+(t) - \hat{L}_{j,i}^-(t)$, where

$$\hat{L}_{j,i}^+(t) = \int_0^t \max\{0, L_{j,i}(\varsigma)\} d\varsigma, \quad \hat{L}_{j,i}^-(t) = \int_0^t \max\{-L_{j,i}(\varsigma), 0\} d\varsigma.$$

Since both $\hat{L}_{j,i}^+(t)$ and $\hat{L}_{j,i}^-(t)$ are increasing processes, we conclude that $\hat{L}_{j,i}(t)$ is a finite variation process [91, Sec. 1.3]. Therefore we may use the integration by parts formula [91, Thm. 1.6.5] to obtain

$$\int_0^t \hat{L}_{j,i}(\nu) d\hat{S}_i(\nu) = \hat{L}_{j,i}(t) \hat{S}_i(t) - \int_0^t \hat{S}_i(\nu) d\hat{L}_{j,i}(\nu),$$

where the last integral is a Lebesgue-Stieltjes integral. Substituting the above equality into (B.6), we obtain

$$\int_0^t \left(\hat{L}_{j,i}(t) - \hat{L}_{j,i}(\nu) \right) S_i(\nu) dQ_\nu = \hat{L}_{j,i}(t) \hat{S}_i(t) - \int_0^t \hat{L}_{j,i}(\nu) d\hat{S}_i(\nu) = \int_0^t \hat{S}_i(\nu) d\hat{L}_{j,i}(\nu), \quad (\text{B.7})$$

for $(j, i, t) \in \{1, \dots, l\} \times \{1, \dots, l'\} \times [0, T]$. Then, once again appealing to the decomposition $\hat{L}_{j,i}(t) = \hat{L}_{j,i}^+(t) - \hat{L}_{j,i}^-(t)$, and using the continuity of L and the definition of \hat{L} in (B.2), we apply the fundamental theorem for Lebesgue-Stieltjes integrals [113, Thm. 7.7.1] to (B.7) and obtain

$$\int_0^t \left(\hat{L}_{j,i}(t) - \hat{L}_{j,i}(\nu) \right) S_i(\nu) dQ_\nu = \int_0^t \hat{S}_i(\nu) d\hat{L}_{j,i}(\nu) = \int_0^t \hat{S}_i(\nu) L_{j,i}(\nu) d\nu,$$

for $(j, i, t) \in \{1, \dots, l\} \times \{1, \dots, l'\} \times [0, T]$. The proof of (B.1a) is then concluded by substituting the above expression into (B.4) to obtain

$$\begin{aligned} &\int_0^t \int_\nu L(\beta) S(\nu) d\beta dQ_\nu \\ &= \sum_{i=1}^{l'} \int_0^t \left[\left(\hat{L}_{1,i}(t) - \hat{L}_{1,i}(\nu) \right) S_i(\nu) dQ_\nu \cdots \left(\hat{L}_{l,i}(t) - \hat{L}_{l,i}(\nu) \right) S_i(\nu) dQ_\nu \right]^\top \\ &= \sum_{i=1}^{l'} \int_0^t \left[\hat{S}_i(\nu) L_{1,i}(\nu) d\nu \cdots \hat{S}_i(\nu) L_{l,i}(\nu) d\nu \right]^\top \\ &\stackrel{(\star)}{=} \sum_{i=1}^{l'} \int_0^t \int_0^\nu \left[L_{1,i}(\nu) S_i(\beta) dQ_\beta d\nu \cdots L_{l,i}(\nu) S_i(\beta) dQ_\beta d\nu \right]^\top = \int_0^t \int_0^\nu L(\nu) S(\beta) dQ_\beta d\nu, \end{aligned}$$

where the equality (\star) is obtained by invoking the definition of the scalar process $\hat{S}_i(t)$ in (B.5). \square

The next result is a straightforward moment bound for the Lebesgue integral of pertinent stochastic processes.

Proposition B.1 Consider a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration \mathfrak{F}_t , and let $S \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_s} | \mathfrak{F}_t)$ satisfy

$$\mathbb{E} \left[\int_0^T \|S(\nu)\|^p d\nu \right] < \infty, \quad \forall p \geq 1.$$

Additionally, for any constants $\theta \in \mathbb{R}_{>0}$ and $\xi \in [0, T]$, define

$$N(t) = \int_\xi^t e^{\theta\nu} \|S(\nu)\| d\nu \in \mathbb{R}, \quad t \in [\xi, T].$$

Then,

$$\|N(t)\|_{L_p} \leq \frac{e^{\theta t} - e^{\theta\xi}}{\theta} \sup_{\nu \in [\xi, t]} \|S(\nu)\|_{L_p}, \quad \forall (t, p) \in [\xi, T] \times \mathbb{N}_{\geq 1}. \quad (\text{B.8})$$

Proof. For $p = 1$, we use the Fubini's theorem to conclude that

$$\mathbb{E}[N(t)] = \int_{\xi}^t e^{\theta\nu} \mathbb{E}[\|S(\nu)\|] d\nu \leq \left(\int_{\xi}^t e^{\theta\nu} d\nu \right) \sup_{\nu \in [\xi, t]} \mathbb{E}[\|S(\nu)\|] = \frac{e^{\theta t} - e^{\theta \xi}}{\theta} \sup_{\nu \in [\xi, t]} \mathbb{E}[\|S(\nu)\|],$$

for all $t \in [\xi, T]$, thus establishing (B.8). For $p \in \mathbb{N}_{\geq 2}$, we obtain the following chain of inequalities:

$$\|N(t)\|_{L_p} = \mathbb{E} \left[\left(\int_{\xi}^t e^{\theta\nu} \|S(\nu)\| d\nu \right)^p \right]^{\frac{1}{p}} \stackrel{(\star)}{\leq} \int_{\xi}^t e^{\theta\nu} \mathbb{E}[\|S(\nu)\|^p]^{\frac{1}{p}} d\nu \leq \left(\int_{\xi}^t e^{\theta\nu} d\nu \right) \sup_{\nu \in [\xi, t]} \mathbb{E}[\|S(\nu)\|^p]^{\frac{1}{p}},$$

for all $t \in [\xi, T]$, where (\star) follows from the Minkowski's integral inequality [114, Thm. 202]. The bound in (B.8), $\forall p \in \mathbb{N}_{\geq 2}$, then follows by solving the integral. \square

The following is a simple corollary of the last result.

Corollary B.1 *Let the hypotheses of Proposition B.1 hold. Then, the process*

$$N(t) = \int_{\xi}^t \|S(\nu)\| d\nu \in \mathbb{R}, \quad t \in [\xi, T],$$

satisfies

$$\|N(t)\|_{L_p} \leq (t - \xi) \sup_{\nu \in [\xi, t]} \|S(\nu)\|_{L_p}, \quad \forall (t, p) \in [\xi, T] \times \mathbb{N}_{\geq 1}. \quad (\text{B.9})$$

Proof. The proof follows by noting that $N(t)$ in this corollary can be obtained from $N(t)$ in Proposition B.1 by setting $\theta = 0$. \square

Next, we apply Proposition B.1 to obtain a moment bound for nested Lebesgue integrals.

Proposition B.2 *Consider a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration \mathfrak{F}_t , and let $S_1 \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_1} | \mathfrak{F}_t)$ and $S_2 \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_2} | \mathfrak{F}_t)$ satisfy*

$$\mathbb{E} \left[\int_0^T \left[\|S_1(\nu)\|^{2p} + \|S_2(\nu)\|^{2p} \right] d\nu \right] < \infty, \quad \forall p \geq 1.$$

Additionally, for any constants $\theta_1 \in \mathbb{R}_{>0}$, $\theta_2 \in \mathbb{R}_{>0}$, and $\xi \in [0, T]$, define

$$N(t) = \int_{\xi}^t e^{(\theta_1 - \theta_2)\nu} \|S_1(\nu)\| \left(\int_0^{\nu} e^{\theta_2\beta} \|S_2(\beta)\| d\beta \right) d\nu \in \mathbb{R}, \quad t \in [\xi, T].$$

Then,

$$\|N(t)\|_{L_p} \leq \frac{e^{\theta_1 t} - e^{\theta_1 \xi}}{\theta_1 \theta_2} \sup_{\nu \in [\xi, t]} \|S_1(\nu)\|_{L_{2p}} \sup_{\nu \in [0, t]} \|S_2(\nu)\|_{L_{2p}}, \quad \forall (t, p) \in [\xi, T] \times \mathbb{N}_{\geq 1}. \quad (\text{B.10})$$

Proof. We begin with the case $p = 1$, for which we apply Fubini's theorem, followed by the Cauchy-Schwarz inequality to obtain

$$\begin{aligned} \mathbb{E}[\|N(t)\|] &= \int_{\xi}^t e^{(\theta_1 - \theta_2)\nu} \mathbb{E} \left[\|S_1(\nu)\| \left(\int_0^{\nu} e^{\theta_2\beta} \|S_2(\beta)\| d\beta \right) \right] d\nu \\ &\leq \int_{\xi}^t e^{(\theta_1 - \theta_2)\nu} \mathbb{E}[\|S_1(\nu)\|^2]^{\frac{1}{2}} \mathbb{E} \left[\left(\int_0^{\nu} e^{\theta_2\beta} \|S_2(\beta)\| d\beta \right)^2 \right]^{\frac{1}{2}} d\nu, \quad \forall t \in [\xi, T], \end{aligned}$$

where we have used the fact that, by definition, $N(t) \geq 0$, $\forall t \geq \xi$. Applying Proposition B.1 to the second inner integral with $\{\theta, \xi, n_s, S\} = \{\theta_2, 0, n_2, S_2\}$ leads to

$$\mathbb{E}[\|N(t)\|] \leq \frac{1}{\theta_2} \int_{\xi}^t \left(e^{\theta_1\nu} \mathbb{E}[\|S_1(\nu)\|^2]^{\frac{1}{2}} \sup_{\beta \in [0, \nu]} \mathbb{E}[\|S_2(\beta)\|^2]^{\frac{1}{2}} \right) d\nu$$

$$\leq \frac{1}{\theta_2} \left(\int_{\xi}^t e^{\theta_1 \nu} d\nu \right) \sup_{\nu \in [\xi, t]} \mathbb{E} \left[\|S_1(\nu)\|^2 \right]^{\frac{1}{2}} \sup_{\nu \in [0, t]} \mathbb{E} \left[\|S_2(\nu)\|^2 \right]^{\frac{1}{2}}, \quad \forall t \in [\xi, T],$$

where in the first inequality we have used the bound $1 - e^{-\theta_2 \nu} \leq 1$, $\nu \in [\xi, t]$. Solving the integral on the right hand side then leads to (B.10) for $p = 1$.

Next, we consider the L_p norm of $N(t)$ for $p \in \mathbb{N}_{\geq 2}$ given by

$$\|N(t)\|_{L_p} = \mathbb{E} \left[\left(\int_{\xi}^t e^{(\theta_1 - \theta_2)\nu} \|S_1(\nu)\| \left(\int_0^{\nu} e^{\theta_2 \beta} \|S_2(\beta)\| d\beta \right) d\nu \right)^p \right]^{\frac{1}{p}}, \quad \forall t \in [\xi, T].$$

From Minkowski's integral inequality [114, Thm. 202], we get

$$\begin{aligned} \|N(t)\|_{L_p} &= \mathbb{E} \left[\left(\int_{\xi}^t e^{(\theta_1 - \theta_2)\nu} \|S_1(\nu)\| \left(\int_0^{\nu} e^{\theta_2 \beta} \|S_2(\beta)\| d\beta \right) d\nu \right)^p \right]^{\frac{1}{p}} \\ &\leq \int_{\xi}^t e^{(\theta_1 - \theta_2)\nu} \mathbb{E} \left[\|S_1(\nu)\|^p \left(\int_0^{\nu} e^{\theta_2 \beta} \|S_2(\beta)\| d\beta \right)^p \right]^{\frac{1}{p}} d\nu, \end{aligned}$$

for all $t \in [\xi, T]$. Subsequently, applying the Cauchy-Schwarz inequality then leads to

$$\|N(t)\|_{L_p} \leq \int_{\xi}^t e^{(\theta_1 - \theta_2)\nu} \mathbb{E} \left[\|S_1(\nu)\|^{2p} \right]^{\frac{1}{2p}} \mathbb{E} \left[\left(\int_0^{\nu} e^{\theta_2 \beta} \|S_2(\beta)\| d\beta \right)^{2p} \right]^{\frac{1}{2p}} d\nu, \quad \forall t \in [\xi, T]. \quad (\text{B.11})$$

Next, using Proposition B.1 for $\{\theta, \xi, n_s, S\} = \{\theta_2, 0, n_2, S_2\}$, we obtain

$$\begin{aligned} \mathbb{E} \left[\left(\int_0^{\nu} e^{\theta_2 \beta} \|S_2(\beta)\| d\beta \right)^{2p} \right]^{\frac{1}{2p}} &\leq \left(\frac{e^{\theta_2 \nu} - 1}{\theta_2} \right) \sup_{\beta \in [0, \nu]} \mathbb{E} \left[\|S_2(\beta)\|^{2p} \right]^{\frac{1}{2p}} \\ &\leq \frac{e^{\theta_2 \nu}}{\theta_2} \sup_{\beta \in [0, \nu]} \mathbb{E} \left[\|S_2(\beta)\|^{2p} \right]^{\frac{1}{2p}}, \quad \forall \nu \in [\xi, t], \end{aligned}$$

where we have bounded $e^{\theta_2 \nu} - 1$ by $e^{\theta_2 \nu}$. Substituting the above bound in (B.11) yields

$$\begin{aligned} \|N(t)\|_{L_p} &\leq \frac{1}{\theta_2} \int_{\xi}^t e^{\theta_1 \nu} \mathbb{E} \left[\|S_1(\nu)\|^{2p} \right]^{\frac{1}{2p}} \sup_{\beta \in [0, \nu]} \mathbb{E} \left[\|S_2(\beta)\|^{2p} \right]^{\frac{1}{2p}} d\nu \\ &\leq \frac{1}{\theta_2} \left(\int_{\xi}^t e^{\theta_1 \nu} d\nu \right)^{\frac{1}{p}} \sup_{\nu \in [\xi, t]} \mathbb{E} \left[\|S_1(\nu)\|^{2p} \right]^{\frac{1}{2p}} \sup_{\nu \in [0, t]} \mathbb{E} \left[\|S_2(\nu)\|^{2p} \right]^{\frac{1}{2p}}, \\ &\leq \frac{e^{\theta_1 t} - e^{\theta_1 \xi}}{\theta_1 \theta_2} \sup_{\nu \in [\xi, t]} \mathbb{E} \left[\|S_1(\nu)\|^{2p} \right]^{\frac{1}{2p}} \sup_{\nu \in [0, t]} \mathbb{E} \left[\|S_2(\nu)\|^{2p} \right]^{\frac{1}{2p}}, \quad \forall t \in [\xi, T], \end{aligned}$$

which establishes (B.10) for all $p \in \mathbb{N}_{\geq 2}$. □

The following result provides moment bound for a pertinent class of Itô integrals.

Lemma B.2 Consider a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration \mathfrak{F}_t , and let $L \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_l \times n_q} | \mathfrak{F}_t)$ satisfy

$$\mathbb{E} \left[\int_0^T \|L(\nu)\|_F^p d\nu \right] < \infty, \quad \forall p \geq 2.$$

Additionally, for any constants $\theta \in \mathbb{R}_{>0}$ and $\xi \in [0, T]$, and an \mathfrak{F}_t -adapted Brownian motion $Q_t \in \mathbb{R}^{n_q}$, define

$$N(t) = \int_{\xi}^t e^{\theta \nu} L(\nu) dQ_{\nu} \in \mathbb{R}^{n_l}, \quad t \in [\xi, T].$$

Then,

$$\|N(t)\|_{L_p} \leq \left(\frac{p-1}{2} \right)^{\frac{1}{2}} \left(\frac{e^{2\theta t} - e^{2\theta \xi}}{2\theta} \right)^{\frac{1}{2}} \sup_{\nu \in [\xi, t]} \|L(\nu)\|_{L_p}, \quad \forall (t, p) \in [\xi, T] \times \mathbb{N}_{\geq 2}. \quad (\text{B.12})$$

Proof. We begin with the case when $p = 2$. Since $\xi \in [0, T]$ is a constant, and hence a stopping time, we use Itô isometry [91, Thm. 1.5.21] to obtain

$$\mathbb{E} [\|N(t)\|^2] = \mathbb{E} \left[\left\| \int_{\xi}^t e^{\theta\nu} L(\nu) dQ_{\nu} \right\|^2 \right] = \mathbb{E} \left[\int_{\xi}^t e^{2\theta\nu} \|L(\nu)\|_F^2 d\nu \right].$$

It then follows from Fubini's theorem that

$$\begin{aligned} \mathbb{E} [\|N(t)\|^2] &\leq \int_{\xi}^t e^{2\theta\nu} \mathbb{E} [\|L(\nu)\|_F^2] d\nu \leq \left(\int_{\xi}^t e^{2\theta\nu} d\nu \right) \sup_{\nu \in [\xi, t]} \mathbb{E} [\|L(\nu)\|_F^2] \\ &= \frac{e^{2\theta t} - e^{2\theta\xi}}{2\theta} \sup_{\nu \in [\xi, t]} \mathbb{E} [\|L(\nu)\|_F^2], \quad \forall t \in [\xi, T], \end{aligned}$$

which establishes (B.12) for $p = 2$.

Next, let $p \in \mathbb{N}_{>2}$. Using Itô's lemma [91, Thm. 1.6.4] in a similar fashion to the proof of [91, Thm. 1.7.1], we get

$$\mathbb{E} [\|N(t)\|^p] = \frac{p}{2} \mathbb{E} \left[\int_{\xi}^t \left(\|N(\nu)\|^{p-2} e^{2\theta\nu} \|L(\nu)\|_F^2 + (p-2) \|N(\nu)\|^{p-4} e^{2\theta\nu} \|N(\nu)^{\top} L(\nu)\|^2 \right) d\nu \right], \quad (\text{B.13})$$

for $t \in [\xi, T]$. Thus, we obtain the following bound using the submultiplicativity of norms:

$$\begin{aligned} \mathbb{E} [\|N(t)\|^p] &\leq p \frac{p-1}{2} \mathbb{E} \left[\int_{\xi}^t \|N(\nu)\|^{p-2} e^{2\theta\nu} \|L(\nu)\|_F^2 d\nu \right] \\ &= p \frac{p-1}{2} \mathbb{E} \left[\int_{\xi}^t e^{2\theta\nu(p-2)/p} \|N(\nu)\|^{p-2} e^{4\theta\nu/p} \|L(\nu)\|_F^2 d\nu \right], \quad \forall t \in [\xi, T], \end{aligned}$$

where we obtain the last expression by writing $e^{2\theta\nu} = e^{2\theta\nu[(p-2)/p+2/p]} = e^{2\theta\nu(p-2)/p} e^{4\theta\nu/p}$. Using Hölder's inequality with conjugates $p/(p-2)$ and $p/2$, followed by Fubini's theorem, leads to

$$\begin{aligned} \mathbb{E} [\|N(t)\|^p] &\leq p \frac{p-1}{2} \mathbb{E} \left[\int_{\xi}^t e^{2\theta\nu} \|N(\nu)\|^p d\nu \right]^{\frac{p-2}{p}} \mathbb{E} \left[\int_{\xi}^t e^{2\theta\nu} \|L(\nu)\|_F^p d\nu \right]^{\frac{2}{p}} \\ &= p \frac{p-1}{2} \left(\int_{\xi}^t e^{2\theta\nu} \mathbb{E} [\|N(\nu)\|^p] d\nu \right)^{\frac{p-2}{p}} \left(\int_{\xi}^t e^{2\theta\nu} \mathbb{E} [\|L(\nu)\|_F^p] d\nu \right)^{\frac{2}{p}}, \quad \forall t \in [\xi, T]. \end{aligned}$$

Note from (B.13) that $\mathbb{E} [\|N(t)\|^p]$ is non-decreasing in t . It then follows that

$$\begin{aligned} \mathbb{E} [\|N(t)\|^p] &\leq p \frac{p-1}{2} \left(\int_{\xi}^t e^{2\theta\nu} d\nu \right)^{\frac{p-2}{p}} \mathbb{E} [\|N(t)\|^p]^{\frac{p-2}{p}} \left(\int_{\xi}^t e^{2\theta\nu} \mathbb{E} [\|L(\nu)\|_F^p] d\nu \right)^{\frac{2}{p}} \\ &\leq p \frac{p-1}{2} \left(\int_{\xi}^t e^{2\theta\nu} d\nu \right)^{\frac{p-2}{p}} \mathbb{E} [\|N(t)\|^p]^{\frac{p-2}{p}} \left(\int_{\xi}^t e^{2\theta\nu} d\nu \right)^{\frac{2}{p}} \sup_{\nu \in [\xi, t]} \mathbb{E} [\|L(\nu)\|_F^p]^{\frac{2}{p}} \\ &= p \frac{p-1}{2} \mathbb{E} [\|N(t)\|^p]^{\frac{p-2}{p}} \left(\int_{\xi}^t e^{2\theta\nu} d\nu \right) \sup_{\nu \in [\xi, t]} \mathbb{E} [\|L(\nu)\|_F^p]^{\frac{2}{p}}, \quad \forall t \in [\xi, T], \end{aligned}$$

and thus

$$\begin{aligned} \mathbb{E} [\|N(t)\|^p]^{\frac{2}{p}} &\leq p \frac{p-1}{2} \left(\int_{\xi}^t e^{2\theta\nu} d\nu \right) \sup_{\nu \in [\xi, t]} \mathbb{E} [\|L(\nu)\|_F^p]^{\frac{2}{p}} \\ &= p \frac{p-1}{2} \frac{e^{2\theta t} - e^{2\theta\xi}}{2\theta} \sup_{\nu \in [\xi, t]} \mathbb{E} [\|L(\nu)\|_F^p]^{\frac{2}{p}}, \quad \forall t \in [\xi, T], \end{aligned}$$

which proves (B.12), $\forall p \in \mathbb{N}_{>2}$. □

Next, we provide a bound on a particular Lebesgue integral with an Itô integral in the integrand.

Lemma B.3 Consider a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration \mathfrak{F}_t , and let $S \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_s} | \mathfrak{F}_t)$ and $L \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_l \times n_q} | \mathfrak{F}_t)$ satisfy

$$\mathbb{E} \left[\int_0^T \left(\|S(\nu)\|^{2p} + \|L(\nu)\|_F^{2p} \right) d\nu \right] < \infty, \quad \forall p \geq 1.$$

Additionally, for any $\theta_1, \theta_2 \in \mathbb{R}_{>0}$, $\xi \in \mathbb{R}_{\geq 0}$, and an \mathfrak{F}_t -adapted Brownian motion $Q_t \in \mathbb{R}^{n_q}$ define

$$N(t) = \int_{\xi}^t e^{(\theta_1 - \theta_2)\nu} \|S(\nu)\| \left\| \int_0^{\nu} e^{\theta_2\beta} \mathbb{1}_{[\zeta_1, \zeta_2]}(\beta) L(\beta) dQ_{\beta} \right\| d\nu \in \mathbb{R}, \quad t \in [\xi, T],$$

where and $0 \leq \zeta_1 \leq \zeta_2 \leq t$ are any deterministic variables. Then,

$$\|N(t)\|_{L_p} \leq \left(p \frac{2p-1}{2} \right)^{\frac{1}{2}} \frac{e^{\theta_1 t} - e^{\theta_1(\xi \vee \zeta_1)}}{\theta_1 \sqrt{\theta_2}} \left(1 - e^{-2\theta_2(\zeta_2 - \zeta_1)} \right)^{\frac{1}{2}} \sup_{\nu \in [\xi, t]} \|S(\nu)\|_{L_{2p}} \sup_{\nu \in [\zeta_1, \zeta_2]} \|L(\nu)\|_{L_{2p}}, \quad (\text{B.14})$$

for all $(t, p) \in [\xi, T] \times \mathbb{N}_{\geq 1}$.

Proof. We provide the proof only for the case when $p \in \mathbb{N}_{\geq 2}$, since the proof for $p = 1$ is established *mutatis mutandis*. The presence of the indicator function in the inner integrand allows us to write

$$\begin{aligned} N(t) &= \int_{\xi}^t e^{(\theta_1 - \theta_2)\nu} \|S(\nu)\| \left\| \int_0^{\nu} e^{\theta_2\beta} \mathbb{1}_{[\zeta_1, \zeta_2]}(\beta) L(\beta) dQ_{\beta} \right\| d\nu \\ &= \int_{\xi \vee \zeta_1}^t e^{(\theta_1 - \theta_2)\nu} \|S(\nu)\| \left\| \int_{\zeta_1}^{\nu \wedge \zeta_2} e^{\theta_2\beta} L(\beta) dQ_{\beta} \right\| d\nu, \quad t \in [\xi, T]. \end{aligned}$$

Now, consider the L_p norm of $N(t)$. By following the same line of reasoning that led to (B.11) in the proof of Proposition B.2, we obtain:

$$\|N(t)\|_{L_p} \leq \int_{\xi \vee \zeta_1}^t e^{(\theta_1 - \theta_2)\nu} \mathbb{E} \left[\|S(\nu)\|^{2p} \right]^{\frac{1}{2p}} \mathbb{E} \left[\left\| \int_{\zeta_1}^{\nu \wedge \zeta_2} e^{\theta_2\beta} L(\beta) dQ_{\beta} \right\|^{2p} \right]^{\frac{1}{2p}} d\nu, \quad t \in [\xi, T].$$

Next, using Lemma B.2 for $\{\theta, \xi, t\} = \{\theta_2, \zeta_1, \nu \wedge \zeta_2\}$ leads to

$$\begin{aligned} \mathbb{E} \left[\left\| \int_{\zeta_1}^{\nu \wedge \zeta_2} e^{\theta_2\beta} L(\beta) dQ_{\beta} \right\|^{2p} \right]^{\frac{1}{2p}} &\leq \left(p \frac{2p-1}{2\theta_2} \right)^{\frac{1}{2}} \left(e^{2\theta_2(\nu \wedge \zeta_2)} - e^{2\theta_2\zeta_1} \right)^{\frac{1}{2}} \sup_{\beta \in [\zeta_1, \nu \wedge \zeta_2]} \mathbb{E} \left[\|L(\beta)\|^{2p} \right]^{\frac{1}{2p}} \\ &\leq \left(p \frac{2p-1}{2\theta_2} \right)^{\frac{1}{2}} \left(e^{2\theta_2(\nu \wedge \zeta_2)} - e^{2\theta_2\zeta_1} \right)^{\frac{1}{2}} \sup_{\beta \in [\zeta_1, \zeta_2]} \mathbb{E} \left[\|L(\beta)\|^{2p} \right]^{\frac{1}{2p}}, \end{aligned}$$

for all $\nu \in [\xi, t]$, where the last inequality follows from the fact that $[\zeta_1, \nu \wedge \zeta_2] \subseteq [\zeta_1, \zeta_2]$, $\forall \nu \in [\xi, t]$. Therefore,

$$\begin{aligned} \|N(t)\|_{L_p} &\leq \int_{\xi \vee \zeta_1}^t e^{(\theta_1 - \theta_2)\nu} \mathbb{E} \left[\|S(\nu)\|^{2p} \right]^{\frac{1}{2p}} \mathbb{E} \left[\left\| \int_{\zeta_1}^{\nu \wedge \zeta_2} e^{\theta_2\beta} L(\beta) dQ_{\beta} \right\|^{2p} \right]^{\frac{1}{2p}} d\nu \\ &\leq \left(p \frac{2p-1}{2\theta_2} \right)^{\frac{1}{2}} \left(\int_{\xi \vee \zeta_1}^t e^{(\theta_1 - \theta_2)\nu} \left(e^{2\theta_2(\nu \wedge \zeta_2)} - e^{2\theta_2\zeta_1} \right)^{\frac{1}{2}} d\nu \right) \\ &\quad \times \sup_{\nu \in [\xi, t]} \mathbb{E} \left[\|S(\nu)\|^{2p} \right]^{\frac{1}{2p}} \sup_{\nu \in [\zeta_1, \zeta_2]} \mathbb{E} \left[\|L(\nu)\|^{2p} \right]^{\frac{1}{2p}}, \quad (\text{B.15}) \end{aligned}$$

for all $t \in [\xi, T]$, where we have used the fact that $[\xi \vee \zeta_1, t] \subseteq [\xi, t]$. The definition $(\nu \wedge \zeta_2) \doteq \min(\nu, \zeta_2)$ implies that $(\nu \wedge \zeta_2) \leq \nu$ and $(\nu \wedge \zeta_2) \leq \zeta_2$. Hence,

$$\left(e^{2\theta_2(\nu \wedge \zeta_2)} - e^{2\theta_2\zeta_1} \right)^{\frac{1}{2}} = e^{2\theta_2(\nu \wedge \zeta_2)} \left(1 - e^{-2\theta_2((\nu \wedge \zeta_2) - \zeta_1)} \right)^{\frac{1}{2}} \leq e^{2\theta_2\nu} \left(1 - e^{-2\theta_2(\zeta_2 - \zeta_1)} \right)^{\frac{1}{2}}.$$

This bound allows us to further develop the following:

$$\int_{\xi \vee \zeta_1}^t e^{(\theta_1 - \theta_2)\nu} \left(e^{2\theta_2(\nu \wedge \zeta_2)} - e^{2\theta_2 \zeta_1} \right)^{\frac{1}{2}} d\nu \leq \left(\int_{\xi \vee \zeta_1}^t e^{\theta_1 \nu} d\nu \right) \left(1 - e^{-2\theta_2(\zeta_2 - \zeta_1)} \right)^{\frac{1}{2}}.$$

Substituting the above bound in (B.15) yields

$$\begin{aligned} \|N(t)\|_{L_p} &\leq \left(p \frac{2p-1}{2\theta_2} \right)^{\frac{1}{2}} \left(\int_{\xi \vee \zeta_1}^t e^{\theta_1 \nu} d\nu \right) \left(1 - e^{-2\theta_2(\zeta_2 - \zeta_1)} \right)^{\frac{1}{2}} \\ &\quad \times \sup_{\nu \in [\xi, t]} \mathbb{E} \left[\|S(\nu)\|^{2p} \right]^{\frac{1}{2p}} \sup_{\nu \in [\zeta_1, \zeta_2]} \mathbb{E} \left[\|L(\nu)\|^{2p} \right]^{\frac{1}{2p}}, \quad (\text{B.16}) \end{aligned}$$

for all $t \in [\xi, T]$. Solving the integral on the right hand side establishes (B.14) for $p \in \mathbb{N}_{\geq 2}$, thus completing the proof. \square

The following is a corollary of the last result.

Corollary B.2 Consider a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration \mathfrak{F}_t , and let $S \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_s} | \mathfrak{F}_t)$ and $L \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_l \times n_q} | \mathfrak{F}_t)$ satisfy

$$\mathbb{E} \left[\int_0^T \left(\|S(\nu)\|^{2p} + \|L(\nu)\|_F^{4p} \right) d\nu \right] < \infty, \quad \forall p \geq 1.$$

Additionally, for any $\theta_1, \theta_2 \in \mathbb{R}_{>0}$, $\xi \in \mathbb{R}_{\geq 0}$, $\mathbf{T}_s \in \mathbb{R}_{>0}$, and an \mathfrak{F}_t -adapted Brownian motion $Q_t \in \mathbb{R}^{n_q}$ define

$$N_{\{a,b\}}(t) = \int_{\mathbf{T}_s}^t e^{(\theta_1 - \theta_2)\nu} \|S(\nu)\| \|R_{\{a,b\}}(\nu)\| d\nu \in \mathbb{R}, \quad t \in [\mathbf{T}_s, T],$$

where

$$R_a(\nu) = L(\nu) \varsigma_1(\nu) \int_{(i^*(\nu)-1)\mathbf{T}_s}^{i^*(\nu)\mathbf{T}_s} e^{\theta_2 \beta \varsigma_2(\beta)} dQ_\beta, \quad R_b(\nu) = L(\nu) \int_{i^*(\nu)\mathbf{T}_s}^{\nu} e^{\theta_2 \beta} dQ_\beta,$$

and where $i^*(\nu) \doteq \max \{i \in \mathbb{N}_{\geq 1} : i\mathbf{T}_s \leq \nu\} = \lfloor \nu/\mathbf{T}_s \rfloor$. Furthermore, $\varsigma_{\{1,2\}} \in \mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R})$ are any deterministic functions. Then, for all $(t, p) \in [\mathbf{T}_s, T] \times \mathbb{N}_{\geq 1}$:

$$\begin{aligned} \|N_a(t)\|_{L_p} &\leq (p(4p-1))^{\frac{1}{2}} \frac{e^{\theta_1 t} - e^{\theta_1 \mathbf{T}_s}}{\theta_1 \sqrt{\theta_2}} (1 - e^{-2\theta_2 \mathbf{T}_s})^{\frac{1}{2}} \\ &\quad \times \sup_{\nu \in [\mathbf{T}_s, t]} \left(\|S(\nu)\|_{L_{2p}} \|L(\nu)\|_{L_{4p}} |\varsigma_1(\nu)| \sup_{\beta \in [(i^*(\nu)-1)\mathbf{T}_s, i^*(\nu)\mathbf{T}_s]} |\varsigma_2(\beta)| \right), \quad (\text{B.17}) \end{aligned}$$

and

$$\|N_b(t)\|_{L_p} \leq (p(4p-1))^{\frac{1}{2}} \frac{e^{\theta_1 t} - e^{\theta_1 \mathbf{T}_s}}{\theta_1 \sqrt{\theta_2}} (1 - e^{-2\theta_2 \mathbf{T}_s})^{\frac{1}{2}} \sup_{\nu \in [\mathbf{T}_s, t]} \|S(\nu)\|_{L_{2p}} \sup_{\nu \in [\mathbf{T}_s, t]} \|L(\nu)\|_{L_{4p}}. \quad (\text{B.18})$$

Proof. We provide the proof only for the case when $p \in \mathbb{N}_{\geq 2}$. As in the proof of Lemma B.3, we have

$$\|N_{\{a,b\}}(t)\|_{L_p} \leq \int_{\mathbf{T}_s}^t e^{(\theta_1 - \theta_2)\nu} \mathbb{E} \left[\|S(\nu)\|^{2p} \right]^{\frac{1}{2p}} \mathbb{E} \left[\|R_{\{a,b\}}(\nu)\|^{2p} \right]^{\frac{1}{2p}} d\nu, \quad (\text{B.19})$$

for all $(t, p) \in [\mathbf{T}_s, T] \times \mathbb{N}_{\geq 2}$. Now, it follows from the Cauchy-Schwarz inequality that

$$\begin{aligned} \|R_a(\nu)\|_{L_{2p}} &= \left\| L(\nu) \varsigma_1(\nu) \int_{(i^*(\nu)-1)\mathbf{T}_s}^{i^*(\nu)\mathbf{T}_s} e^{\theta_2 \beta \varsigma_2(\beta)} dQ_\beta \right\|_{L_{2p}} \\ &\leq \|L(\nu)\|_{L_{4p}} |\varsigma_1(\nu)| \left\| \int_{(i^*(\nu)-1)\mathbf{T}_s}^{i^*(\nu)\mathbf{T}_s} e^{\theta_2 \beta \varsigma_2(\beta)} dQ_\beta \right\|_{L_{4p}}, \quad \forall \nu \in [\mathbf{T}_s, t], \end{aligned}$$

where we have used the fact that ς_1 is a deterministic function. Using Lemma B.2 to bound the Itô integral term, we get

$$\begin{aligned} \|R_a(\nu)\|_{L_{2p}} &\leq (p(4p-1))^{\frac{1}{2}} e^{\theta_2 i^*(\nu) T_s} \left(\frac{1 - e^{-2\theta_2 T_s}}{\theta_2} \right)^{\frac{1}{2}} \|L(\nu)\|_{L_{4p}} |\varsigma_1(\nu)| \sup_{\beta \in [(i^*(\nu)-1)T_s, i^*(\nu)T_s]} |\varsigma_2(\beta)| \\ &\leq (p(4p-1))^{\frac{1}{2}} e^{\theta_2 \nu} \left(\frac{1 - e^{-2\theta_2 T_s}}{\theta_2} \right)^{\frac{1}{2}} \|L(\nu)\|_{L_{4p}} |\varsigma_1(\nu)| \sup_{\beta \in [(i^*(\nu)-1)T_s, i^*(\nu)T_s]} |\varsigma_2(\beta)|, \end{aligned} \quad (\text{B.20})$$

for all $(t, p) \in [T_s, T] \times \mathbb{N}_{\geq 2}$, where we have used the fact that ς_2 is a deterministic function. Further, the second inequality follows from $i^*(\nu) T_s \leq \nu, \forall \nu$. Similarly, using the Cauchy-Schwarz inequality and Lemma B.2:

$$\begin{aligned} \|R_b(\nu)\|_{L_{2p}} &\leq (p(4p-1))^{\frac{1}{2}} e^{\theta_2 \nu} \left(\frac{1 - e^{2\theta_2(i^*(\nu)T_s - \nu)}}{\theta_2} \right)^{\frac{1}{2}} \|L(\nu)\|_{L_{4p}} \\ &\leq (p(4p-1))^{\frac{1}{2}} e^{\theta_2 \nu} \left(\frac{1 - e^{-2\theta_2 T_s}}{\theta_2} \right)^{\frac{1}{2}} \|L(\nu)\|_{L_{4p}}, \quad \forall (t, p) \in [T_s, T] \times \mathbb{N}_{\geq 2}, \end{aligned} \quad (\text{B.21})$$

where we obtain the second inequality by using the fact that $\nu - i^*(\nu) T_s \leq T_s, \forall \nu$. Substituting (B.20) into (B.19) yields

$$\begin{aligned} \|N_a(t)\|_{L_p} &\leq (p(4p-1))^{\frac{1}{2}} \left(\int_{T_s}^t e^{\theta_1 \nu} d\nu \right) \left(\frac{1 - e^{-2\theta_2 T_s}}{\theta_2} \right)^{\frac{1}{2}} \\ &\quad \times \sup_{\nu \in [T_s, t]} \left[\|S(\nu)\|_{L_{2p}} \|L(\nu)\|_{L_{4p}} |\varsigma_1(\nu)| \sup_{\beta \in [(i^*(\nu)-1)T_s, i^*(\nu)T_s]} |\varsigma_2(\beta)| \right], \end{aligned}$$

for all $(t, p) \in [T_s, T] \times \mathbb{N}_{\geq 2}$. Solving the integral on the right hand side then leads to (B.17). Substituting (B.21) into (B.19) and performing the same manipulations as above, we obtain (B.18). \square

Next, we provide a converse result to Lemma B.3 in which we derive a bound on a particular Itô integral with a Lebesgue integral in the integrand.

Lemma B.4 Given any $S \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_o} | \mathfrak{F}_t)$, $L \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_o \times n_q} | \mathfrak{F}_t)$, \mathfrak{F}_t -adapted Brownian motion $Q_t \in \mathbb{R}^{n_q}$, and strictly positive constants $\theta_1, \theta_2 \in \mathbb{R}_{>0}$, define

$$N(t) = \int_0^t e^{(\theta_1 - \theta_2)\nu} \left(\int_0^\nu e^{\theta_2 \beta} S(\beta) d\beta \right)^\top L(\nu) dQ_\nu \in \mathbb{R}, \quad t \in [0, T]. \quad (\text{B.22})$$

If

$$\mathbb{E} \left[\int_0^T \left(\|S(\nu)\|^{2p} + \|L(\nu)\|_F^{2p} \right) d\nu \right] < \infty, \quad \forall p \geq 2,$$

then

$$\|N(t)\|_{L_p} \leq (p(p-1))^{\frac{1}{2}} \frac{e^{\theta_1 t}}{2\sqrt{\theta_1 \theta_2}} \sup_{\nu \in [0, t]} \|S(\nu)\|_{L_{2p}} \sup_{\nu \in [0, t]} \|L(\nu)\|_{L_{2p}}, \quad \forall (t, p) \in [0, T] \times \mathbb{N}_{\geq 2}. \quad (\text{B.23})$$

Proof. The proof for $p = 2$ follows by using [91, Thm. 1.5.21] and following the proof for $p > 2$ *mutatis mutandis*, which we present below. We begin by noting that the inner Lebesgue integral is absolute continuous on $[0, T]$ by fundamental theorem for the Lebesgue integral [113, Thm. 6.4.1], and $S \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_o} | \mathfrak{F}_t)$ implies that

$$\int_0^t e^{\theta_2 \beta} S(\beta) d\beta \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_o} | \mathfrak{F}_t).$$

It follows then from [90, Thm. 5.5.2] that $N(t)$ is a continuous local martingale with respect to the filtration \mathfrak{F}_t since $L \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_o \times n_q} | \mathfrak{F}_t)$. Then, applying Itô's lemma [91, Thm. 1.6.4] and using [91, Thm. 1.5.21] produces

$$|N(t)|^p = \frac{p}{2} \int_0^t e^{2(\theta_1 - \theta_2)\nu} |N(\nu)|^{p-2} \left\| L(\nu)^\top \left(\int_0^\nu e^{\theta_2 \beta} S(\beta) d\beta \right) \right\|^2 d\nu \\ + \frac{p}{2} (p-2) \int_0^t e^{2(\theta_1 - \theta_2)\nu} |N(\nu)|^{p-4} \left\| N(\nu) L(\nu)^\top \left(\int_0^\nu e^{\theta_2 \beta} S(\beta) d\beta \right) \right\|^2 d\nu, \quad (\text{B.24})$$

and therefore

$$|N(t)|^p \leq \frac{p}{2} (p-1) \int_0^t e^{2(\theta_1 - \theta_2)\nu} |N(\nu)|^{p-2} \|L(\nu)\|_F^2 \left\| \int_0^\nu e^{\theta_2 \beta} S(\beta) d\beta \right\|^2 d\nu, \quad \forall t \in [0, T].$$

Using this bound, we obtain

$$\mathbb{E} [|N(t)|^p] = \frac{p}{2} (p-1) \mathbb{E} \left[\int_0^t e^{2(\theta_1 - \theta_2)\nu} |N(\nu)|^{p-2} \|L(\nu)\|_F^2 \left\| \int_0^\nu e^{\theta_2 \beta} S(\beta) d\beta \right\|^2 d\nu \right] \\ \stackrel{(i)}{=} \frac{p}{2} (p-1) \mathbb{E} \left[\int_0^t \left(e^{2\theta_1(p-2)\nu/p} |N(\nu)|^{p-2} \right) \left(e^{2(2\theta_1/p - \theta_2)\nu} \|L(\nu)\|_F^2 \left\| \int_0^\nu e^{\theta_2 \beta} S(\beta) d\beta \right\|^2 \right) d\nu \right] \\ \stackrel{(ii)}{\leq} \frac{p}{2} (p-1) \mathbb{E} \left[\int_0^t e^{2\theta_1 \nu} |N(\nu)|^p d\nu \right]^{\frac{p-2}{p}} \mathbb{E} \left[\int_0^t e^{(2\theta_1 - \theta_2 p)\nu} \|L(\nu)\|_F^p \left\| \int_0^\nu e^{\theta_2 \beta} S(\beta) d\beta \right\|^p d\nu \right]^{\frac{2}{p}},$$

for all $t \in [0, T]$, where we obtain (i) by writing

$$e^{2(\theta_1 - \theta_2)\nu} = e^{2[\theta_1(\frac{p-2}{p} + \frac{2}{p}) - \theta_2]\nu} = e^{2\theta_1(p-2)\nu/p} e^{2(2\theta_1/p - \theta_2)\nu},$$

and (ii) follows from Hölder's inequality with conjugates $p/(p-2)$ and $p/2$. Applying Fubini's theorem, followed by the Cauchy-Schwarz inequality then leads to

$$\mathbb{E} [|N(t)|^p] \leq \frac{p}{2} (p-1) \left(\int_0^t e^{2\theta_1 \nu} \mathbb{E} [|N(\nu)|^p] d\nu \right)^{\frac{p-2}{p}} \left(\int_0^t e^{(2\theta_1 - \theta_2 p)\nu} \mathbb{E} \left[\|L(\nu)\|_F^p \left\| \int_0^\nu e^{\theta_2 \beta} S(\beta) d\beta \right\|^p \right] d\nu \right)^{\frac{2}{p}} \\ \leq \frac{p}{2} (p-1) \left(\int_0^t e^{2\theta_1 \nu} \mathbb{E} [|N(\nu)|^p] d\nu \right)^{\frac{p-2}{p}} \\ \times \left(\int_0^t e^{(2\theta_1 - \theta_2 p)\nu} \mathbb{E} [\|L(\nu)\|_F^{2p}]^{\frac{1}{2}} \mathbb{E} \left[\left\| \int_0^\nu e^{\theta_2 \beta} S(\beta) d\beta \right\|^{2p} \right]^{\frac{1}{2}} d\nu \right)^{\frac{2}{p}}, \quad (\text{B.25})$$

for all $t \in [0, T]$.

Using (B.24), one sees that $\mathbb{E} [|N(t)|^p]$ is non-decreasing over $t \in [0, T]$, and thus

$$\left(\int_0^t e^{2\theta_1 \nu} \mathbb{E} [|N(\nu)|^p] d\nu \right)^{\frac{p-2}{p}} \leq \left(\int_0^t e^{2\theta_1 \nu} d\nu \right)^{\frac{p-2}{p}} \mathbb{E} [|N(t)|^p]^{\frac{p-2}{p}}, \quad \forall t \in [0, T]. \quad (\text{B.26})$$

Next, we use Proposition B.1 for $\{\theta, \xi, n_s\} = \{\theta_2, 0, n_o\}$ and obtain

$$\mathbb{E} \left[\left\| \int_0^\nu e^{\theta_2 \beta} S(\beta) d\beta \right\|^{2p} \right]^{\frac{1}{2}} \leq \mathbb{E} \left[\left(\int_0^\nu e^{\theta_2 \beta} \|S(\beta)\| d\beta \right)^{2p} \right]^{\frac{1}{2}} \leq \left(\frac{e^{\theta_2 \nu} - 1}{\theta_2} \right)^p \sup_{\beta \in [0, \nu]} \mathbb{E} [\|S(\beta)\|^{2p}]^{\frac{1}{2}} \\ \leq \frac{e^{\theta_2 p \nu}}{\theta_2^p} \sup_{\beta \in [0, \nu]} \mathbb{E} [\|S(\beta)\|^{2p}]^{\frac{1}{2}}, \quad \forall \nu \in [0, t],$$

where we have bounded $e^{\theta_2 \nu} - 1$ by $e^{\theta_2 \nu}$. Therefore,

$$\left(\int_0^t e^{(2\theta_1 - \theta_2 p)\nu} \mathbb{E} [\|L(\nu)\|_F^{2p}]^{\frac{1}{2}} \mathbb{E} \left[\left\| \int_0^\nu e^{\theta_2 \beta} S(\beta) d\beta \right\|^{2p} \right]^{\frac{1}{2}} d\nu \right)^{\frac{2}{p}}$$

$$\begin{aligned} &\leq \left(\frac{1}{\theta_2}\right)^2 \left(\int_0^t e^{2\theta_1\nu} \mathbb{E} \left[\|L(\nu)\|_F^{2p} \right]^{\frac{1}{2}} \sup_{\beta \in [0, \nu]} \mathbb{E} \left[\|S(\beta)\|^{2p} \right]^{\frac{1}{2}} d\nu \right)^{\frac{2}{p}} \\ &\leq \left(\frac{1}{\theta_2}\right)^2 \left(\int_0^t e^{2\theta_1\nu} d\nu \right)^{\frac{2}{p}} \sup_{\nu \in [0, t]} \mathbb{E} \left[\|L(\nu)\|_F^{2p} \right]^{\frac{1}{p}} \sup_{\nu \in [0, t]} \mathbb{E} \left[\|S(\nu)\|^{2p} \right]^{\frac{1}{p}}, \quad (\text{B.27}) \end{aligned}$$

for all $t \in [0, T]$. Substituting (B.26) and (B.27) into (B.25) yields

$$\begin{aligned} \mathbb{E} [|N(t)|^p] &\leq \mathbb{E} [|N(t)|^p]^{\frac{p-2}{p}} p \frac{(p-1)}{2} \frac{1}{\theta_2^2} \left(\int_0^t e^{2\theta_1\nu} d\nu \right) \sup_{\nu \in [0, t]} \mathbb{E} \left[\|L(\nu)\|_F^{2p} \right]^{\frac{1}{p}} \sup_{\nu \in [0, t]} \mathbb{E} \left[\|S(\nu)\|^{2p} \right]^{\frac{1}{p}} \\ &\leq \mathbb{E} [|N(t)|^p]^{\frac{p-2}{p}} p \frac{(p-1)}{4} \frac{e^{2\theta_1 t}}{\theta_1 \theta_2^2} \sup_{\nu \in [0, t]} \mathbb{E} \left[\|L(\nu)\|_F^{2p} \right]^{\frac{1}{p}} \sup_{\nu \in [0, t]} \mathbb{E} \left[\|S(\nu)\|^{2p} \right]^{\frac{1}{p}}, \end{aligned}$$

for all $t \in [0, T]$, where we have bounded $e^{2\theta_1 t} - 1$ by $e^{2\theta_1 t}$. This bound then implies that

$$\mathbb{E} [|N(t)|^p]^{\frac{2}{p}} \leq p \frac{(p-1)}{4} \frac{e^{2\theta_1 t}}{\theta_1 \theta_2^2} \sup_{\nu \in [0, t]} \mathbb{E} \left[\|L(\nu)\|_F^{2p} \right]^{\frac{1}{p}} \sup_{\nu \in [0, t]} \mathbb{E} \left[\|S(\nu)\|^{2p} \right]^{\frac{1}{p}}, \quad \forall t \in [0, T],$$

thus establishing (B.23) for $p > 2$. □

We need the following corollary to the last result for Lemma B.5.

Corollary B.3 Consider a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration \mathfrak{F}_t , and let $L_1, L_2 \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_l \times n_q} | \mathfrak{F}_t)$ satisfy

$$\mathbb{E} \left[\int_0^T \|L_1(\nu) + L_2(\nu)\|_F^{2p} d\nu \right] < \infty, \quad \forall p \geq 2.$$

Additionally, for any strictly positive constants $\theta_1, \theta_2 \in \mathbb{R}_{>0}$ and an \mathfrak{F}_t -adapted Brownian motion $Q_t \in \mathbb{R}^{n_q}$ define

$$\check{N}(t) = \int_0^t e^{(\theta_1 - \theta_2)\nu} \left(\int_0^\nu e^{\theta_2\beta} L_2(\beta) dQ_\beta \right)^\top L_1(\nu) dQ_\nu \in \mathbb{R}.$$

Then,

$$\|\check{N}(t)\|_{L_p} \leq \frac{p\sqrt{(p-1)(2p-1)}}{2\sqrt{2}} \frac{e^{\theta_1 t}}{\sqrt{\theta_1 \theta_2}} \sup_{\nu \in [0, t]} \|L_1(\nu)\|_{L_{2p}} \sup_{\nu \in [0, t]} \|L_2(\nu)\|_{L_{2p}}, \quad \forall (t, p) \in [0, T] \times \mathbb{N}_{\geq 2}. \quad (\text{B.28})$$

Proof. As in the proof of Lemma B.4, we provide a proof for $p > 2$ only since the proof for $p = 2$ follows *mutatis mutandis* by using [91, Thm. 1.5.21]. For $p > 2$, applying Itô's lemma [91, Thm. 1.6.4] and using [91, Thm. 1.5.21] produces

$$\begin{aligned} |\check{N}(t)|^p &= \frac{p}{2} \int_0^t e^{2(\theta_1 - \theta_2)\nu} |\check{N}(\nu)|^{p-2} \left\| L_1(\nu)^\top \left(\int_0^\nu e^{\theta_2\beta} L_2(\beta) dQ_\beta \right) \right\|^2 d\nu \\ &\quad + \frac{p}{2} (p-2) \int_0^t e^{2(\theta_1 - \theta_2)\nu} |\check{N}(\nu)|^{p-4} \left\| \check{N}(\nu) L_1(\nu)^\top \left(\int_0^\nu e^{\theta_2\beta} L_2(\beta) dQ_\beta \right) \right\|^2 d\nu, \quad (\text{B.29}) \end{aligned}$$

and therefore

$$|\check{N}(t)|^p \leq \frac{p}{2} (p-1) \int_0^t e^{2(\theta_1 - \theta_2)\nu} |\check{N}(\nu)|^{p-2} \|L_1(\nu)\|_F^2 \left\| \int_0^\nu e^{\theta_2\beta} L_2(\beta) dQ_\beta \right\|^2 d\nu, \quad \forall t \in [0, T].$$

Following the identical analysis that led to (B.25) in the proof of Lemma B.4, we obtain

$$\mathbb{E} [|\check{N}(t)|^p] \leq \frac{p}{2} (p-1) \left(\int_0^t e^{2\theta_1\nu} \mathbb{E} [|\check{N}(\nu)|^p] d\nu \right)^{\frac{p-2}{p}}$$

$$\times \left(\int_0^t e^{(2\theta_1 - \theta_2 p)\nu} \mathbb{E} \left[\|L_1(\nu)\|_F^{2p} \right]^{\frac{1}{2}} \mathbb{E} \left[\left\| \int_0^\nu e^{\theta_2 \beta} L_2(\beta) dQ_\beta \right\|^{2p} \right]^{\frac{1}{2}} d\nu \right)^{\frac{2}{p}}, \quad (\text{B.30})$$

for all $t \in [0, T]$. Using (B.29), one sees that $\mathbb{E} \left[|\check{N}(t)|^p \right]$ is non-decreasing over $t \in [0, T]$, and thus

$$\left(\int_0^t e^{2\theta_1 \nu} \mathbb{E} \left[|\check{N}(\nu)|^p \right] d\nu \right)^{\frac{p-2}{p}} \leq \left(\int_0^t e^{2\theta_1 \nu} d\nu \right)^{\frac{p-2}{p}} \mathbb{E} \left[|\check{N}(t)|^p \right]^{\frac{p-2}{p}}, \quad \forall t \in [0, T]. \quad (\text{B.31})$$

Next, we use Lemma B.2 for $\{\theta, \xi\} = \{\theta_2, 0\}$ and obtain

$$\mathbb{E} \left[\left\| \int_0^\nu e^{\theta_2 \beta} L_2(\beta) dQ_\beta \right\|^{2p} \right]^{\frac{1}{2}} \leq \left(p \frac{2p-1}{2} \right)^{\frac{p}{2}} \left(\frac{e^{2\theta_2 \nu}}{\theta_2} \right)^{\frac{p}{2}} \sup_{\beta \in [0, \nu]} \mathbb{E} \left[\|L_2(\beta)\|_F^{2p} \right]^{\frac{1}{2}}, \quad \forall \nu \in [0, t],$$

where we have bounded $e^{2\theta_2 \nu} - 1$ by $e^{2\theta_2 \nu}$. Therefore,

$$\begin{aligned} & \left(\int_0^t e^{(2\theta_1 - \theta_2 p)\nu} \mathbb{E} \left[\|L_1(\nu)\|_F^{2p} \right]^{\frac{1}{2}} \mathbb{E} \left[\left\| \int_0^\nu e^{\theta_2 \beta} L_2(\beta) dQ_\beta \right\|^{2p} \right]^{\frac{1}{2}} d\nu \right)^{\frac{2}{p}} \\ & \leq p \frac{2p-1}{2\theta_2} \left(\int_0^t e^{2\theta_1 \nu} \mathbb{E} \left[\|L_1(\nu)\|_F^{2p} \right]^{\frac{1}{2}} \sup_{\beta \in [0, \nu]} \mathbb{E} \left[\|L_2(\beta)\|_F^{2p} \right]^{\frac{1}{2}} d\nu \right)^{\frac{2}{p}} \\ & \leq p \frac{2p-1}{2\theta_2} \left(\int_0^t e^{2\theta_1 \nu} d\nu \right)^{\frac{2}{p}} \sup_{\nu \in [0, t]} \mathbb{E} \left[\|L_1(\nu)\|_F^{2p} \right]^{\frac{1}{p}} \sup_{\nu \in [0, t]} \mathbb{E} \left[\|L_2(\nu)\|_F^{2p} \right]^{\frac{1}{p}}, \quad (\text{B.32}) \end{aligned}$$

for all $t \in [0, T]$. Substituting (B.31) and (B.32) into (B.30) yields

$$\begin{aligned} \mathbb{E} \left[|\check{N}(t)|^p \right] & \leq \mathbb{E} \left[|\check{N}(t)|^p \right]^{\frac{p-2}{p}} \frac{p^2(p-1)(2p-1)}{4\theta_2} \left(\int_0^t e^{2\theta_1 \nu} d\nu \right) \sup_{\nu \in [0, t]} \mathbb{E} \left[\|L_1(\nu)\|_F^{2p} \right]^{\frac{1}{p}} \sup_{\nu \in [0, t]} \mathbb{E} \left[\|L_2(\nu)\|_F^{2p} \right]^{\frac{1}{p}} \\ & \leq \mathbb{E} \left[|\check{N}(t)|^p \right]^{\frac{p-2}{p}} \frac{1}{8} p^2(p-1)(2p-1) \frac{e^{2\theta_1 t}}{\theta_1 \theta_2} \sup_{\nu \in [0, t]} \mathbb{E} \left[\|L_1(\nu)\|_F^{2p} \right]^{\frac{1}{p}} \sup_{\nu \in [0, t]} \mathbb{E} \left[\|L_2(\nu)\|_F^{2p} \right]^{\frac{1}{p}}, \end{aligned}$$

for all $t \in [0, T]$, where we have bounded $e^{2\theta_1 t} - 1$ by $e^{2\theta_1 t}$. This bound then implies that

$$\mathbb{E} \left[|\check{N}(t)|^p \right]^{\frac{2}{p}} \leq \frac{1}{8} p^2(p-1)(2p-1) \frac{e^{2\theta_1 t}}{\theta_1 \theta_2} \sup_{\nu \in [0, t]} \mathbb{E} \left[\|L_1(\nu)\|_F^{2p} \right]^{\frac{1}{p}} \sup_{\nu \in [0, t]} \mathbb{E} \left[\|L_2(\nu)\|_F^{2p} \right]^{\frac{1}{p}}, \quad \forall t \in [0, T],$$

thus establishing (B.28) for $p > 2$. □

The following result provides a bound on a class of nested Itô integrals.

Lemma B.5 Consider a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration \mathfrak{F}_t , and let $L_1, L_2 \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_l \times n_q} | \mathfrak{F}_t)$ satisfy

$$\mathbb{E} \left[\int_0^T \|L_1(\nu) + L_2(\nu)\|_F^{2p} d\nu \right] < \infty, \quad \forall p \geq 2. \quad (\text{B.33})$$

Additionally, for any strictly positive constants $\theta_1, \theta_2 \in \mathbb{R}_{>0}$ and an \mathfrak{F}_t -adapted Brownian motion $Q_t \in \mathbb{R}^{n_q}$ define

$$N(t) = \int_0^t e^{\theta_2 \nu} \left(\int_\nu^t e^{(\theta_1 - \theta_2)\beta} L_1(\beta) dQ_\beta \right)^\top L_2(\nu) dQ_\nu \in \mathbb{R}, \quad t \in [0, T].$$

Then,

$$\mathbb{E} [N(t)] \leq \sqrt{n_l} \frac{e^{\theta_1 t}}{\theta_1} \sup_{\nu \in [0, t]} \mathbb{E} [\|L_1(\nu) L_2(\nu)^\top\|_F], \quad \forall t \in [0, T], \quad (\text{B.34})$$

and

$$\|N(t)\|_{L_p} \leq \frac{1}{\sqrt{2}} \left(p \frac{p-1}{2} \right)^{\frac{1}{2}} \left(p \frac{2p-1}{2} \right)^{\frac{1}{2}} \frac{e^{\theta_1 t}}{\sqrt{\theta_1 \theta_2}} \sup_{\nu \in [0, t]} \|L_1(\nu)\|_{L_{2p}} \sup_{\nu \in [0, t]} \|L_2(\nu)\|_{L_{2p}} + \sqrt{n_l} \frac{e^{\theta_1 t}}{\theta_1} \sup_{\nu \in [0, t]} \|L_1(\nu) L_2(\nu)^\top\|_{L_p}, \quad (\text{B.35})$$

for all $(t, p) \in [0, T] \times \mathbb{N}_{\geq 2}$.

Proof. We begin by setting

$$N_1(t) = \int_0^t e^{(\theta_1 - \theta_2)\nu} L_1(\nu) dQ_\nu \in \mathbb{R}^{n_l}, \quad N_2(t) = \int_0^t e^{\theta_2 \nu} L_2(\nu) dQ_\nu \in \mathbb{R}^{n_l}, \quad t \in [0, T].$$

It then follows from the definition of $N(t)$ that

$$N(t) = \int_0^t e^{\theta_2 \nu} \left(\int_\nu^t e^{(\theta_1 - \theta_2)\beta} L_1(\beta) dQ_\beta \right)^\top L_2(\nu) dQ_\nu = \int_0^t \left(\int_\nu^t dN_1(\beta) \right)^\top dN_2(\nu).$$

Hence,

$$\begin{aligned} N(t) &= \int_0^t \left(\int_0^\nu dN_1(\beta) \right)^\top dN_2(\nu) - \int_0^t \left(\int_0^\nu dN_1(\beta) \right)^\top dN_2(\nu) \\ &= N_1(t)^\top N_2(t) - \int_0^t N_1(\nu)^\top dN_2(\nu), \quad t \in [0, T]. \end{aligned} \quad (\text{B.36})$$

Next, applying Itô's lemma to $N_1(t)^\top N_2(t)$ (or alternatively using the Itô product rule [107, Sec. 4.4.2] applied element wise) produces

$$N_1(t)^\top N_2(t) = \int_0^t N_1(\nu)^\top dN_2(\nu) + \int_0^t N_2(\nu)^\top dN_1(\nu) + \sum_{i=1}^{n_l} \left\langle N_{1,i}, N_{2,i} \right\rangle_t, \quad t \in [0, T],$$

where $N_{1,i}, N_{2,i} \in \mathbb{R}$ denote the i^{th} scalar-valued process, $i \in \{1, \dots, n_l\}$, of $N_1(t)$ and $N_2(t)$, respectively, and $\left\langle N_{1,i}, N_{2,i} \right\rangle_t$ denotes the cross-variation process between $N_{1,i}(t)$ and $N_{2,i}(t)$ [89, Defn. 2.3.9]. Substituting the above expression for $N_1(t)^\top N_2(t)$ into (B.36) produces

$$N(t) = \int_0^t N_2(\nu)^\top dN_1(\nu) + \sum_{i=1}^{n_l} \left\langle N_{1,i}, N_{2,i} \right\rangle_t, \quad t \in [0, T].$$

It then follows from the linearity of the expectation operator and the Minkowski's inequality that

$$\begin{aligned} \mathbb{E} [N(t)] &= \mathbb{E} \left[\int_0^t N_2(\nu)^\top dN_1(\nu) \right] + \mathbb{E} \left[\sum_{i=1}^{n_l} \left\langle N_{1,i}, N_{2,i} \right\rangle_t \right] \\ &\leq \mathbb{E} \left[\int_0^t N_2(\nu)^\top dN_1(\nu) \right] + \mathbb{E} \left[\left| \sum_{i=1}^{n_l} \left\langle N_{1,i}, N_{2,i} \right\rangle_t \right| \right], \quad \forall t \in [0, T], \end{aligned} \quad (\text{B.37})$$

and

$$\mathbb{E} [|N(t)|^p]^{\frac{1}{p}} \leq \mathbb{E} \left[\left| \int_0^t N_2(\nu)^\top dN_1(\nu) \right|^p \right]^{\frac{1}{p}} + \mathbb{E} \left[\left| \sum_{i=1}^{n_l} \left\langle N_{1,i}, N_{2,i} \right\rangle_t \right|^p \right]^{\frac{1}{p}}, \quad \forall (t, p) \in [0, T] \times \mathbb{N}_{\geq 2}. \quad (\text{B.38})$$

Now, let

$$\check{N}(t) \doteq \int_0^t N_2(\nu)^\top dN_1(\nu) = \int_0^t e^{(\theta_1 - \theta_2)\nu} \left(\int_0^\nu e^{\theta_2 \beta} L_2(\beta) dQ_\beta \right)^\top L_1(\nu) dQ_\nu \in \mathbb{R}.$$

The t -continuity of $N_2(t)$ [15, Thm. 3.2.5], along with $L_1, L_2 \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_l \times n_q} | \mathfrak{F}_t)$ imply that $\check{N}(t)$ is a continuous local martingale [90, Thm. 5.5.2]. Hence, it follows from [91, Thm. 1.5.21] that

$$\mathbb{E} [\check{N}(t)] = \mathbb{E} \left[\int_0^t N_2(\nu)^\top dN_1(\nu) \right] = 0, \quad \forall t \in [0, T]. \quad (\text{B.39})$$

Furthermore, from Corollary B.3, we have that

$$\begin{aligned} \mathbb{E} \left[\left| \check{N}(t) \right|^p \right]^{\frac{1}{p}} &= \mathbb{E} \left[\left| \int_0^t N_2(\nu)^\top dN_1(\nu) \right|^p \right]^{\frac{1}{p}} \\ &\leq \frac{p\sqrt{(p-1)(2p-1)}}{2\sqrt{2}} \frac{e^{\theta_1 t}}{\sqrt{\theta_1 \theta_2}} \sup_{\nu \in [0, t]} \|L_1(\nu)\|_{L_{2p}} \sup_{\nu \in [0, t]} \|L_2(\nu)\|_{L_{2p}}, \end{aligned} \quad (\text{B.40})$$

for all $(t, p) \in [0, T] \times \mathbb{N}_{\geq 2}$.

Next, it follows from the definition of the cross-variation process [89, Defn. 2.3.9] that

$$\sum_{i=1}^{n_l} \langle N_{1,i}, N_{2,i} \rangle_t = \frac{1}{4} \sum_{i=1}^{n_l} \left(\langle N_{1,i} + N_{2,i} \rangle_t - \langle N_{1,i} - N_{2,i} \rangle_t \right) = \sum_{i=1}^{n_l} \int_0^t e^{\theta_1 \nu} L_{1,i}(\nu) L_{2,i}(\nu)^\top d\nu,$$

where $L_{1,i}, L_{2,i} \in \mathbb{R}^{1 \times n_q}$ denote the i^{th} rows of $L_1(t)$ and $L_2(t)$, respectively. Hence,

$$\begin{aligned} \left| \sum_{i=1}^{n_l} \langle N_{1,i}, N_{2,i} \rangle_t \right| &\leq \int_0^t e^{\theta_1 \nu} \left| \sum_{i=1}^{n_l} L_{1,i}(\nu) L_{2,i}(\nu)^\top \right| d\nu \leq \sqrt{n_l} \int_0^t e^{\theta_1 \nu} \left(\sum_{i=1}^{n_l} |L_{1,i}(\nu) L_{2,i}(\nu)^\top|^2 \right)^{\frac{1}{2}} d\nu \\ &= \sqrt{n_l} \int_0^t e^{\theta_1 \nu} \|L_1(\nu) L_2(\nu)^\top\|_F d\nu, \end{aligned}$$

for all $t \in [0, T]$. Then, applying Proposition B.1 for $\{\theta, \xi, \|S(t)\|\} = \{\theta_1, 0, \|L_1(t) L_2(t)^\top\|_F\}$ leads to

$$\mathbb{E} \left[\left| \sum_{i=1}^m \langle N_{1,i}, N_{2,i} \rangle_t \right|^p \right]^{\frac{1}{p}} \leq \sqrt{n_l} \frac{e^{\theta_1 t}}{\theta_1} \sup_{\nu \in [0, t]} \|L_1(\nu) L_2(\nu)^\top\|_{L_p}, \quad \forall (t, p) \in [0, T] \times \mathbb{N}_{\geq 1}, \quad (\text{B.41})$$

where we have bounded $e^{\theta_1 t} - 1$ by $e^{\theta_1 t}$. Substituting (B.39), (B.40), and (B.41) into (B.37) and (B.38) produces the desired result. □

The final major result of this section provides a bound on a class of nested Itô and Lebesgue integrals.

Lemma B.6 Consider a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration \mathfrak{F}_t , and let $R \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_1 \times n_2} | \mathfrak{F}_t)$, $S_1 \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_1} | \mathfrak{F}_t)$, $S_2 \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_2} | \mathfrak{F}_t)$, $L_1 \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_1 \times n_q} | \mathfrak{F}_t)$, and $L_2 \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_2 \times n_q} | \mathfrak{F}_t)$ satisfy

$$\sum_{i=1}^2 \mathbb{E} \left[\int_0^T \left(\|R(\nu)\|_F^{2p} + \|S_i(\nu)\|^{2p} + \|L_i(\nu)\|_F^{2p} \right) d\nu \right] < \infty, \quad \forall p \geq 1.$$

Additionally, for any constants $\theta_1, \theta_2 \in \mathbb{R}_{>0}$ and an \mathfrak{F}_t -adapted Brownian motion $Q_t \in \mathbb{R}^{n_q}$ define

$$N(t) = \int_\xi^t e^{2(\theta_1 - \theta_2)\nu} \check{R}_1(\nu)^\top R(\nu) \check{R}_2(\nu) d\nu \in \mathbb{R}, \quad 0 \leq \xi \leq t \leq T, \quad (\text{B.42})$$

where $\xi \in \mathbb{R}_{\geq 0}$ is a constant, and

$$\check{R}_1(t) = \int_0^t e^{\theta_2 \beta} [S_1(\beta) d\beta + L_1(\beta) dQ_\beta] \in \mathbb{R}^{n_1}, \quad \check{R}_2(t) = \int_0^t e^{\theta_2 \beta} [S_2(\beta) d\beta + L_2(\beta) dQ_\beta] \in \mathbb{R}^{n_2}.$$

If there exists a constant $\Delta_R \in \mathbb{R}_{>0}$ such that

$$\|R(t)\|_F \leq \Delta_R, \quad \forall t \in [0, T],$$

then, the following bound holds for all $(t, \mathbf{p}) \in [\xi, T] \times \mathbb{N}_{\geq 1}$:

$$\|N(t)\|_{L_{\mathbf{p}}} \leq \Delta_R \frac{e^{2\theta_1 t} - e^{2\theta_1 \xi}}{2\theta_1} \prod_{i=1}^2 \left(\frac{1 - e^{-\theta_2 t}}{\theta_2} \sup_{\nu \in [0, t]} \|S_i(\nu)\|_{L_{2\mathbf{p}}} + \left(\mathbf{p} \frac{2\mathbf{p} - 1}{2} \right)^{\frac{1}{2}} \left(\frac{1 - e^{-2\theta_2 t}}{\theta_2} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t]} \|L_i(\nu)\|_{L_{2\mathbf{p}}} \right). \quad (\text{B.43})$$

Proof. We begin with the case $\mathbf{p} = 1$ by observing that one has the straightforward bound

$$|N(t)| \leq \Delta_R \int_{\xi}^t e^{2(\theta_1 - \theta_2)\nu} \|\check{R}_1(\nu)\| \|\check{R}_2(\nu)\| d\nu, \quad \forall t \in [\xi, T].$$

Applying Fubini's theorem, followed by the Cauchy-Schwarz inequality produces

$$\mathbb{E} [|N(t)|] \leq \Delta_R \int_{\xi}^t e^{2(\theta_1 - \theta_2)\nu} \|\check{R}_1(\nu)\|_{L_2} \|\check{R}_2(\nu)\|_{L_2} d\nu, \quad \forall t \in [\xi, T]. \quad (\text{B.44})$$

Using the Minkowski's inequality, one can establish that

$$\|\check{R}_i(\nu)\|_{L_2} \leq \left\| \int_0^{\nu} e^{\theta_2 \beta} \|S_i(\beta)\|_{L_2} d\beta \right\|_{L_2} + \left\| \int_0^{\nu} e^{\theta_2 \beta} L_i(\beta) dQ_{\beta} \right\|_{L_2}, \quad \forall (\nu, i) \in [\xi, t] \times \{1, 2\}.$$

Using Proposition B.1 and Lemma B.2, we obtain

$$\|\check{R}_i(\nu)\|_{L_2} \leq e^{\theta_2 \nu} \left(\frac{1 - e^{-\theta_2 \nu}}{\theta_2} \sup_{\beta \in [0, \nu]} \|S_i(\beta)\|_{L_2} + \left(\frac{1 - e^{-2\theta_2 \nu}}{2\theta_2} \right)^{\frac{1}{2}} \sup_{\beta \in [0, \nu]} \|L_i(\beta)\|_{L_2} \right),$$

for all $(\nu, i) \in [\xi, t] \times \{1, 2\}$. Substituting the above inequality into (B.44) leads to

$$\mathbb{E} [|N(t)|] \leq \Delta_R \int_{\xi}^t e^{2\theta_1 \nu} \prod_{i=1}^2 \left(\frac{1 - e^{-\theta_2 \nu}}{\theta_2} \sup_{\beta \in [0, \nu]} \|S_i(\beta)\|_{L_2} + \left(\frac{1 - e^{-2\theta_2 \nu}}{2\theta_2} \right)^{\frac{1}{2}} \sup_{\beta \in [0, \nu]} \|L_i(\beta)\|_{L_2} \right) d\nu,$$

which can be further developed into

$$\mathbb{E} [|N(t)|] \leq \Delta_R \left(\int_{\xi}^t e^{2\theta_1 \nu} d\nu \right) \prod_{i=1}^2 \left(\frac{1 - e^{-\theta_2 t}}{\theta_2} \sup_{\nu \in [0, t]} \|S_i(\nu)\|_{L_2} + \left(\frac{1 - e^{-2\theta_2 t}}{2\theta_2} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t]} \|L_i(\nu)\|_{L_2} \right),$$

for all $t \in [\xi, T]$. Solving the integral on the right hand side of the above inequality produces the desired result for $\mathbf{p} = 1$.

Next, consider $\mathbf{p} \in \mathbb{N}_{\geq 2}$, and define

$$\check{N}(t) = \int_{\xi}^t e^{2(\theta_1 - \theta_2)\nu} \|\check{R}_1(\nu)\| \|\check{R}_2(\nu)\| d\nu \in \mathbb{R}, \quad t \in [\xi, T]. \quad (\text{B.45})$$

Then,

$$\begin{aligned} |N(t)| &\leq \int_{\xi}^t e^{2(\theta_1 - \theta_2)\nu} \|\check{R}_1(\nu)\| \|R(\nu)\|_F \|\check{R}_2(\nu)\| d\nu \leq \Delta_R \int_{\xi}^t e^{2(\theta_1 - \theta_2)\nu} \|\check{R}_1(\nu)\| \|\check{R}_2(\nu)\| d\nu \\ &= \Delta_R \check{N}(t), \end{aligned}$$

and thus

$$\|N(t)\|_{L_{\mathbf{p}}} \leq \Delta_R \|\check{N}(t)\|_{L_{\mathbf{p}}}, \quad \forall (t, \mathbf{p}) \in [\xi, T] \times \mathbb{N}_{\geq 1}. \quad (\text{B.46})$$

By following the same line of reasoning that led to (B.11) in the proof of Proposition B.2, we obtain:

$$\|\check{N}(t)\|_{L_{\mathbf{p}}} \leq \int_{\xi}^t e^{2(\theta_1 - \theta_2)\nu} \mathbb{E} \left[\|\check{R}_1(\nu)\|^{2\mathbf{p}} \right]^{\frac{1}{2\mathbf{p}}} \mathbb{E} \left[\|\check{R}_2(\nu)\|^{2\mathbf{p}} \right]^{\frac{1}{2\mathbf{p}}} d\nu, \quad \forall t \in [\xi, T]. \quad (\text{B.47})$$

Now, it follows from the definition of \check{R}_i , $i \in \{1, 2\}$, that

$$\begin{aligned} \mathbb{E} \left[\left\| \check{R}_i(\nu) \right\|^{2p} \right]^{\frac{1}{2p}} &\leq \mathbb{E} \left[\left(\left\| \int_0^\nu e^{\theta_2 \beta} S_i(\beta) d\beta \right\| + \left\| \int_0^\nu e^{\theta_2 \beta} L_i(\beta) dQ_\beta \right\| \right)^{2p} \right]^{\frac{1}{2p}} \\ &\stackrel{(\star)}{\leq} \mathbb{E} \left[\left\| \int_0^\nu e^{\theta_2 \beta} S_i(\beta) d\beta \right\|^{2p} \right]^{\frac{1}{2p}} + \mathbb{E} \left[\left\| \int_0^\nu e^{\theta_2 \beta} L_i(\beta) dQ_\beta \right\|^{2p} \right]^{\frac{1}{2p}} \\ &\leq \mathbb{E} \left[\left\| \int_0^\nu e^{\theta_2 \beta} \|S_i(\beta)\| d\beta \right\|^{2p} \right]^{\frac{1}{2p}} + \mathbb{E} \left[\left\| \int_0^\nu e^{\theta_2 \beta} \|L_i(\beta)\| dQ_\beta \right\|^{2p} \right]^{\frac{1}{2p}}, \quad \forall (\nu, i) \in [\xi, t] \times \{1, 2\}, \end{aligned}$$

where (\star) is due to the Minkowski's inequality. Using Proposition B.1 and Lemma B.2 then leads to

$$\begin{aligned} &\mathbb{E} \left[\left\| \check{R}_i(\nu) \right\|^{2p} \right]^{\frac{1}{2p}} \\ &\leq e^{\theta_2 \nu} \left(\frac{1 - e^{-\theta_2 \nu}}{\theta_2} \sup_{\beta \in [0, \nu]} \|S_i(\beta)\|_{L_{2p}} + \left(p \frac{2p-1}{2} \right)^{\frac{1}{2}} \left(\frac{1 - e^{-2\theta_2 \nu}}{\theta_2} \right)^{\frac{1}{2}} \sup_{\beta \in [0, \nu]} \|L_i(\beta)\|_{L_{2p}} \right), \end{aligned}$$

for all $(\nu, i) \in [\xi, t] \times \{1, 2\}$. We thus conclude that

$$\begin{aligned} &\mathbb{E} \left[\left\| \check{R}_1(\nu) \right\|^{2p} \right]^{\frac{1}{2p}} \mathbb{E} \left[\left\| \check{R}_2(\nu) \right\|^{2p} \right]^{\frac{1}{2p}} \\ &\leq e^{2\theta_2 \nu} \prod_{i=1}^2 \left(\frac{1 - e^{-\theta_2 \nu}}{\theta_2} \sup_{\beta \in [0, \nu]} \|S_i(\beta)\|_{L_{2p}} + \left(p \frac{2p-1}{2} \right)^{\frac{1}{2}} \left(\frac{1 - e^{-2\theta_2 \nu}}{\theta_2} \right)^{\frac{1}{2}} \sup_{\beta \in [0, \nu]} \|L_i(\beta)\|_{L_{2p}} \right), \end{aligned}$$

for all $\nu \in [\xi, t]$. Substituting the above bound in (B.47) yields

$$\begin{aligned} &\left\| \check{N}(t) \right\|_{L_p} \\ &\leq \int_\xi^t e^{2\theta_1 \nu} \prod_{i=1}^2 \left(\frac{1 - e^{-\theta_2 \nu}}{\theta_2} \sup_{\beta \in [0, \nu]} \|S_i(\beta)\|_{L_{2p}} + \left(p \frac{2p-1}{2} \right)^{\frac{1}{2}} \left(\frac{1 - e^{-2\theta_2 \nu}}{\theta_2} \right)^{\frac{1}{2}} \sup_{\beta \in [0, \nu]} \|L_i(\beta)\|_{L_{2p}} \right) d\nu, \end{aligned}$$

which further implies

$$\begin{aligned} &\left\| \check{N}(t) \right\|_{L_p} \\ &\leq \left(\int_\xi^t e^{2\theta_1 \nu} d\nu \right) \prod_{i=1}^2 \left(\frac{1 - e^{-\theta_2 t}}{\theta_2} \sup_{\nu \in [0, t]} \|S_i(\nu)\|_{L_{2p}} + \left(p \frac{2p-1}{2} \right)^{\frac{1}{2}} \left(\frac{1 - e^{-2\theta_2 t}}{\theta_2} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t]} \|L_i(\nu)\|_{L_{2p}} \right), \end{aligned}$$

for all $t \in [\xi, T]$. Solving the integral on the right hand side and substituting in (B.46) then leads to

$$\begin{aligned} &\|N(t)\|_{L_p} \\ &\leq \Delta_R \frac{e^{2\theta_1 t} - e^{2\theta_1 \xi}}{2\theta_1} \prod_{i=1}^2 \left(\frac{1 - e^{-\theta_2 t}}{\theta_2} \sup_{\nu \in [0, t]} \|S_i(\nu)\|_{L_{2p}} + \left(p \frac{2p-1}{2} \right)^{\frac{1}{2}} \left(\frac{1 - e^{-2\theta_2 t}}{\theta_2} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t]} \|L_i(\nu)\|_{L_{2p}} \right), \end{aligned}$$

for all $t \in [\xi, T]$, thus, establishing (B.43), for $p \in \mathbb{N}_{\geq 2}$. □

We conclude the appendix with a couple of corollaries, starting with a corollary of the last lemma.

Corollary B.4 Consider a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration \mathfrak{F}_t , and let $R \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_1 \times n_2} | \mathfrak{F}_t)$, $S_1 \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_1} | \mathfrak{F}_t)$, $S_2 \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_2} | \mathfrak{F}_t)$, $L_1 \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_1 \times n_q} | \mathfrak{F}_t)$, and $L_2 \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_2 \times n_q} | \mathfrak{F}_t)$ satisfy

$$\sum_{i=1}^2 \mathbb{E} \left[\int_0^T \left(\|R(\nu)\|_F^{2p} + \|S_i(\nu)\|^{2p} + \|L_i(\nu)\|_F^{2p} \right) d\nu \right] < \infty, \quad \forall p \geq 1.$$

Additionally, for any $\theta_1, \theta_2 \in \mathbb{R}_{>0}$, $\xi \in \mathbb{R}_{\geq 0}$, and an \mathfrak{F}_t -adapted Brownian motion $Q_t \in \mathbb{R}^{n_q}$ define

$$N_{\{a,b\}}(t) = \int_{\xi}^t e^{2(\theta_1 - \theta_2)\nu} \check{R}_{\{a,b\}}(\nu)^\top R(\nu) \check{R}_2(\nu) d\nu \in \mathbb{R}, \quad 0 \leq \xi \leq t \leq T,$$

where

$$\begin{aligned} \check{R}_a(\xi) &= \int_0^\xi e^{\theta_2 \beta} S_1(\beta) d\beta \in \mathbb{R}^{n_1}, \quad \check{R}_b(\nu) = \int_0^\nu e^{\theta_2 \beta} \mathbb{1}_{[\zeta_1, \zeta_2]}(\beta) L_1(\beta) dQ_\beta \in \mathbb{R}^{n_1}, \\ \check{R}_2(t) &= \int_0^t e^{\theta_2 \beta} [S_2(\beta) d\beta + L_2(\beta) dQ_\beta] \in \mathbb{R}^{n_2}. \end{aligned}$$

In the above, $0 \leq \zeta_1 \leq \zeta_2 < t$ are any deterministic variables. Then, for all $(t, p) \in [\xi, T] \times \mathbb{N}_{\geq 1}$

$$\begin{aligned} \mathbb{E} [|N_a(t)|^p]^{\frac{1}{p}} &\leq \frac{\Delta_R}{2} \frac{e^{2\theta_1 t} - e^{2\theta_1 \xi}}{\theta_1 \theta_2} (e^{\theta_2 \xi} - 1) \sup_{\nu \in [0, t]} \|S_1(\nu)\|_{L_{2p}} \\ &\quad \times \left(\frac{1 - e^{-\theta_2 t}}{\theta_2} \sup_{\nu \in [0, t]} \|S_2(\nu)\|_{L_{2p}} + \left(p \frac{2p-1}{2} \right)^{\frac{1}{2}} \left(\frac{1 - e^{-2\theta_2 t}}{\theta_2} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t]} \|L_2(\nu)\|_{L_{2p}} \right), \end{aligned} \quad (\text{B.48})$$

and

$$\begin{aligned} \mathbb{E} [|N_b(t)|^p]^{\frac{1}{p}} &\leq \frac{\Delta_R}{2} \left(p \frac{2p-1}{2} \right)^{\frac{1}{2}} \frac{e^{2\theta_1 t} - e^{2\theta_1 (\xi \vee \zeta_1)}}{\theta_1 \sqrt{\theta_2}} \left(1 - e^{-2\theta_2 (\zeta_2 - \zeta_1)} \right)^{\frac{1}{2}} \sup_{\nu \in [\zeta_1, \zeta_2]} \mathbb{E} [\|L_1(\nu)\|^{2p}]^{\frac{1}{2p}} \\ &\quad \times \left(\frac{1 - e^{-\theta_2 t}}{\theta_2} \sup_{\nu \in [0, t]} \|S_2(\nu)\|_{L_{2p}} + \left(p \frac{2p-1}{2} \right)^{\frac{1}{2}} \left(\frac{1 - e^{-2\theta_2 t}}{\theta_2} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t]} \|L_2(\nu)\|_{L_{2p}} \right). \end{aligned} \quad (\text{B.49})$$

Proof. We provide a proof for the case $p \in \mathbb{N}_{\geq 2}$ only and begin by writing

$$N(t) = N_a(t) + N_b(t), \quad t \in [\xi, T],$$

where

$$\begin{aligned} N_{\{a,b\}}(t) &= \int_{\xi}^t e^{2(\theta_1 - \theta_2)\nu} \check{R}_{\{a,b\}}(\nu)^\top R(\nu) \check{R}_2(\nu) d\nu, \\ \check{R}_a(\xi) &= \int_0^\xi e^{\theta_2 \beta} S_1(\beta) d\beta, \quad \check{R}_b(\nu) = \int_0^\nu e^{\theta_2 \beta} \mathbb{1}_{[\zeta_1, \zeta_2]}(\beta) L_1(\beta) dQ_\beta. \end{aligned}$$

Now, following the proof of Lemma B.6 with \check{R}_1 replaced by \check{R}_a leads to

$$\begin{aligned} &\mathbb{E} [|N_a(t)|^p]^{\frac{1}{p}} \\ &\leq \Delta_R \left(\int_{\xi}^t e^{(2\theta_1 - \theta_2)\nu} d\nu \right) \left(\frac{e^{\theta_2 \xi} - 1}{\theta_2} \sup_{\nu \in [0, \xi]} \|S_1(\nu)\|_{L_{2p}} \right) \\ &\quad \times \left(\frac{1 - e^{-\theta_2 t}}{\theta_2} \sup_{\nu \in [0, t]} \|S_2(\nu)\|_{L_{2p}} + \left(p \frac{2p-1}{2} \right)^{\frac{1}{2}} \left(\frac{1 - e^{-2\theta_2 t}}{\theta_2} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t]} \|L_2(\nu)\|_{L_{2p}} \right), \end{aligned}$$

for all $t \in [\xi, T]$. Note that since

$$\int_{\xi}^t e^{(2\theta_1 - \theta_2)\nu} d\nu \leq \int_{\xi}^t e^{2\theta_1 \nu} d\nu = \frac{e^{2\theta_1 t} - e^{2\theta_1 \xi}}{2\theta_1},$$

we can write the previous bound as

$$\mathbb{E} [|N_a(t)|^p]^{\frac{1}{p}} \leq \frac{\Delta_R}{2} \frac{e^{2\theta_1 t} - e^{2\theta_1 \xi}}{\theta_1 \theta_2} (e^{\theta_2 \xi} - 1) \sup_{\nu \in [0, t]} \|S_1(\nu)\|_{L_{2p}}$$

$$\times \left(\frac{1 - e^{-\theta_2 t}}{\theta_2} \sup_{\nu \in [0, t]} \|S_2(\nu)\|_{L_{2p}} + \left(\frac{2p-1}{2} \right)^{\frac{1}{2}} \left(\frac{1 - e^{-2\theta_2 t}}{\theta_2} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t]} \|L_2(\nu)\|_{L_{2p}} \right),$$

for all $t \in [\xi, T]$, where we have also used the bound $\sup_{\nu \in [0, \xi]} \|S_1(\nu)\|_{L_{2p}} \leq \sup_{\nu \in [0, t]} \|S_1(\nu)\|_{L_{2p}}$, thus proving (B.48).

Next, as at the beginning of the proof of Lemma B.3, we can write

$$\begin{aligned} N_b(t) &= \int_{\xi}^t e^{2(\theta_1 - \theta_2)\nu} \left(\int_0^{\nu} e^{\theta_2 \beta} \mathbb{1}_{[\zeta_1, \zeta_2]}(\beta) L_1(\beta) dQ_{\beta} \right)^{\top} R(\nu) \check{R}_2(\nu) d\nu \\ &= \int_{\xi \vee \zeta_1}^t e^{2(\theta_1 - \theta_2)\nu} \left(\int_{\zeta_1}^{\nu \wedge \zeta_2} e^{\theta_2 \beta} L_1(\beta) dQ_{\beta} \right)^{\top} R(\nu) \check{R}_2(\nu) d\nu, \quad \forall t \in [\xi, T]. \end{aligned}$$

Therefore, as in (B.16), and using the bound on $\mathbb{E} \left[\|\check{R}_2(\nu)\|^{2p} \right]^{\frac{1}{2p}}$ in the proof of Lemma B.6, we obtain

$$\begin{aligned} &\mathbb{E} [|N_b(t)|^p]^{\frac{1}{p}} \\ &\leq \Delta_R \left(\int_{\xi \vee \zeta_1}^t e^{2\theta_1 \nu} d\nu \right) \left(\frac{2p-1}{2\theta_2} \right)^{\frac{1}{2}} \left(1 - e^{-2\theta_2(\zeta_2 - \zeta_1)} \right)^{\frac{1}{2}} \sup_{\nu \in [\zeta_1, \zeta_2]} \mathbb{E} [\|L_1(\nu)\|^{2p}]^{\frac{1}{2p}} \\ &\quad \times \left(\frac{1 - e^{-\theta_2 t}}{\theta_2} \sup_{\nu \in [0, t]} \|S_2(\nu)\|_{L_{2p}} + \left(\frac{2p-1}{2} \right)^{\frac{1}{2}} \left(\frac{1 - e^{-2\theta_2 t}}{\theta_2} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t]} \|L_2(\nu)\|_{L_{2p}} \right), \end{aligned}$$

for all $t \in [\xi, T]$, which then leads to

$$\begin{aligned} \mathbb{E} [|N_b(t)|^p]^{\frac{1}{p}} &\leq \frac{\Delta_R}{2} \left(\frac{2p-1}{2} \right)^{\frac{1}{2}} \frac{e^{2\theta_1 t} - e^{2\theta_1(\xi \vee \zeta_1)}}{\theta_1 \sqrt{\theta_2}} \left(1 - e^{-2\theta_2(\zeta_2 - \zeta_1)} \right)^{\frac{1}{2}} \sup_{\nu \in [\zeta_1, \zeta_2]} \mathbb{E} [\|L_1(\nu)\|^{2p}]^{\frac{1}{2p}} \\ &\quad \times \left(\frac{1 - e^{-\theta_2 t}}{\theta_2} \sup_{\nu \in [0, t]} \|S_2(\nu)\|_{L_{2p}} + \left(\frac{2p-1}{2} \right)^{\frac{1}{2}} \left(\frac{1 - e^{-2\theta_2 t}}{\theta_2} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t]} \|L_2(\nu)\|_{L_{2p}} \right), \end{aligned}$$

for all $t \in [\xi, T]$, thus proving (B.49). □

The final result considers a particular process resulting from the piecewise nature of the adaptation law.

Corollary B.5 Consider a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration \mathfrak{F}_t , and let $R \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_1 \times n_2} | \mathfrak{F}_t)$, $S \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_2} | \mathfrak{F}_t)$, $L_a, L_b \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_1 \times n_q} | \mathfrak{F}_t)$, and $L \in \mathcal{M}_2^{loc}(\mathbb{R}^{n_2 \times n_q} | \mathfrak{F}_t)$ satisfy

$$\mathbb{E} \left[\int_0^T \left(\|R(\nu)\|_F^{2p} + \|S(\nu)\|^2 + \|L(\nu)\|_F^{2p} + \|L_a(\nu)\|_F^{4p} + \|L_b(\nu)\|_F^{4p} \right) d\nu \right] < \infty, \quad \forall p \geq 1.$$

Additionally, for any $\theta_1, \theta_2 \in \mathbb{R}_{>0}$, $\mathbf{T}_s \in \mathbb{R}_{>0}$, and an \mathfrak{F}_t -adapted Brownian motion $Q_t \in \mathbb{R}^{n_q}$ define

$$N_{\{a,b\}}(t) = \int_{\mathbf{T}_s}^t e^{2(\theta_1 - \theta_2)\nu} \check{R}_{\{a,b\}}(\nu)^{\top} R(\nu) \check{R}(\nu) d\nu \in \mathbb{R}, \quad t \in [\mathbf{T}_s, T], \quad (\text{B.50})$$

where

$$\begin{aligned} \check{R}_a(\nu) &= L_a(\nu) \varsigma_1(\nu) \int_{(i^*(\nu)-1)\mathbf{T}_s}^{i^*(\nu)\mathbf{T}_s} e^{\theta_2 \beta} \varsigma_2(\beta) dQ_{\beta} \in \mathbb{R}^{n_1}, \quad \check{R}_b(\nu) = L_b(\nu) \int_{i^*(\nu)\mathbf{T}_s}^{\nu} e^{\theta_2 \beta} dQ_{\beta} \in \mathbb{R}^{n_1}, \\ \check{R}(\nu) &= \int_0^{\nu} e^{\theta_2 \beta} [S(\beta) d\beta + L(\beta) dQ_{\beta}] \in \mathbb{R}^{n_2}, \end{aligned}$$

and where $i^*(\nu) \doteq \max \{i \in \mathbb{N}_{\geq 1} : i\mathbf{T}_s \leq \nu\} = \lfloor \nu/\mathbf{T}_s \rfloor$. Furthermore, $\varsigma_{\{1,2\}} \in \mathcal{C}(\mathbb{R}_{\geq 0}, \mathbb{R})$ are any deterministic functions. Then, for all $(t, p) \in [\mathbf{T}_s, T] \times \mathbb{N}_{\geq 1}$:

$$\|N_a(t)\|_{L_p} \leq \frac{\Delta_R}{2} (p(4p-1))^{\frac{1}{2}} \frac{e^{2\theta_1 t} - e^{2\theta_1 \mathbf{T}_s}}{\theta_1 \sqrt{\theta_2}} (1 - e^{-2\theta_2 \mathbf{T}_s})^{\frac{1}{2}}$$

$$\begin{aligned} & \times \sup_{\nu \in [\mathbf{T}_s, t]} \left(\|L_a(\nu)\|_{L_{4p}} |\varsigma_1(\nu)| \sup_{\beta \in [(i^*(\nu)-1)\mathbf{T}_s, i^*(\nu)\mathbf{T}_s]} |\varsigma_2(\beta)| \right) \\ & \times \left(\frac{1 - e^{-\theta_2 t}}{\theta_2} \sup_{\nu \in [0, t]} \|S(\nu)\|_{L_{2p}} + \left(p \frac{2p-1}{2} \right)^{\frac{1}{2}} \left(\frac{1 - e^{-2\theta_2 t}}{\theta_2} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t]} \|L(\nu)\|_{L_{2p}} \right), \end{aligned} \quad (\text{B.51})$$

and

$$\begin{aligned} \|N_b(t)\|_{L_p} & \leq \frac{\Delta_R}{2} (p(4p-1))^{\frac{1}{2}} \frac{e^{2\theta_1 t} - e^{2\theta_1 \mathbf{T}_s}}{\theta_1 \sqrt{\theta_2}} (1 - e^{-2\theta_2 \mathbf{T}_s})^{\frac{1}{2}} \sup_{\nu \in [\mathbf{T}_s, t]} \|L_b(\nu)\|_{L_{4p}} \\ & \times \left(\frac{1 - e^{-\theta_2 t}}{\theta_2} \sup_{\nu \in [0, t]} \|S(\nu)\|_{L_{2p}} + \left(p \frac{2p-1}{2} \right)^{\frac{1}{2}} \left(\frac{1 - e^{-2\theta_2 t}}{\theta_2} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t]} \|L(\nu)\|_{L_{2p}} \right). \end{aligned} \quad (\text{B.52})$$

Proof. We provide a proof for the case $p \in \mathbb{N}_{\geq 2}$ only and begin by defining

$$\check{N}_{\{a,b\}}(t) = \int_{\mathbf{T}_s}^t e^{2(\theta_1 - \theta_2)\nu} \|\check{R}_{\{a,b\}}(\nu)\| \|\check{R}(\nu)\| d\nu \in \mathbb{R}, \quad t \in [\mathbf{T}_s, T]. \quad (\text{B.53})$$

Then,

$$\begin{aligned} |N_{\{a,b\}}(t)| & \leq \int_{\mathbf{T}_s}^t e^{2(\theta_1 - \theta_2)\nu} \|\check{R}_{\{a,b\}}(\nu)\| \|R(\nu)\|_F \|\check{R}(\nu)\| d\nu \\ & \leq \Delta_R \int_{\mathbf{T}_s}^t e^{2(\theta_1 - \theta_2)\nu} \|\check{R}_{\{a,b\}}(\nu)\| \|\check{R}(\nu)\| d\nu = \Delta_R \check{N}_{\{a,b\}}(t), \end{aligned}$$

and thus

$$\|N_{\{a,b\}}(t)\|_{L_p} \leq \Delta_R \|\check{N}_{\{a,b\}}(t)\|_{L_p}, \quad \forall (t, p) \in [\mathbf{T}_s, T] \times \mathbb{N}_{\geq 2}. \quad (\text{B.54})$$

Next, as in the proof of Lemma B.6, we have that

$$\mathbb{E} \left[\|\check{N}_{\{a,b\}}(t)\|^p \right]^{\frac{1}{p}} \leq \int_{\mathbf{T}_s}^t e^{2(\theta_1 - \theta_2)\nu} \mathbb{E} \left[\|\check{R}_{\{a,b\}}(\nu)\|^{2p} \right]^{\frac{1}{2p}} \mathbb{E} \left[\|\check{R}(\nu)\|^{2p} \right]^{\frac{1}{2p}} d\nu, \quad \forall t \in [\mathbf{T}_s, T]. \quad (\text{B.55})$$

Similarly, as in the proof of Lemma B.6, we further have that

$$\begin{aligned} \mathbb{E} \left[\|\check{R}(\nu)\|^{2p} \right]^{\frac{1}{2p}} & \leq e^{\theta_2 \nu} \left(\frac{1 - e^{-\theta_2 \nu}}{\theta_2} \sup_{\beta \in [0, \nu]} \|S(\beta)\|_{L_{2p}} \right. \\ & \quad \left. + \left(p \frac{2p-1}{2} \right)^{\frac{1}{2}} \left(\frac{1 - e^{-2\theta_2 \nu}}{\theta_2} \right)^{\frac{1}{2}} \sup_{\beta \in [0, \nu]} \|L(\beta)\|_{L_{2p}} \right), \quad \forall \nu \in [\mathbf{T}_s, t]. \end{aligned} \quad (\text{B.56})$$

Next, by the Cauchy-Schwarz inequality

$$\|\check{R}_a(\nu)\|_{L_{2p}} \leq \|L_a(\nu)\|_{L_{4p}} |\varsigma_1(\nu)| \left\| \int_{(i^*(\nu)-1)\mathbf{T}_s}^{i^*(\nu)\mathbf{T}_s} e^{\theta_2 \beta} \varsigma_2(\beta) dQ_\beta \right\|_{L_{4p}}, \quad \forall \nu \in [\mathbf{T}_s, t],$$

where we have used the fact that ς_1 is a deterministic function. Using Lemma B.2 to bound the Itô integral term, we get

$$\begin{aligned} & \|\check{R}_a(\nu)\|_{L_{2p}} \\ & \leq (p(4p-1))^{\frac{1}{2}} e^{\theta_2 i^*(\nu)\mathbf{T}_s} \left(\frac{1 - e^{-2\theta_2 \mathbf{T}_s}}{\theta_2} \right)^{\frac{1}{2}} \|L_a(\nu)\|_{L_{4p}} |\varsigma_1(\nu)| \sup_{\beta \in [(i^*(\nu)-1)\mathbf{T}_s, i^*(\nu)\mathbf{T}_s]} |\varsigma_2(\beta)| \\ & \leq (p(4p-1))^{\frac{1}{2}} e^{\theta_2 \nu} \left(\frac{1 - e^{-2\theta_2 \mathbf{T}_s}}{\theta_2} \right)^{\frac{1}{2}} \|L_a(\nu)\|_{L_{4p}} |\varsigma_1(\nu)| \sup_{\beta \in [(i^*(\nu)-1)\mathbf{T}_s, i^*(\nu)\mathbf{T}_s]} |\varsigma_2(\beta)|, \end{aligned} \quad (\text{B.57})$$

for all $(t, p) \in [\mathbf{T}_s, T] \times \mathbb{N}_{\geq 2}$, where we have used the fact that ς_2 is a deterministic function. Moreover, the second inequality follows from $i^*(\nu) \mathbf{T}_s \leq \nu, \forall \nu$.

Similarly, using the Cauchy-Schwarz inequality and Lemma B.2:

$$\begin{aligned} \|\check{R}_b(\nu)\|_{L_{2p}} &\leq (p(4p-1))^{\frac{1}{2}} e^{\theta_2 \nu} \left(\frac{1 - e^{2\theta_2(i^*(\nu)\mathbf{T}_s - \nu)}}{\theta_2} \right)^{\frac{1}{2}} \|L_b(\nu)\|_{L_{4p}} \\ &\leq (p(4p-1))^{\frac{1}{2}} e^{\theta_2 \nu} \left(\frac{1 - e^{-2\theta_2 \mathbf{T}_s}}{\theta_2} \right)^{\frac{1}{2}} \|L_b(\nu)\|_{L_{4p}}, \end{aligned} \quad (\text{B.58})$$

for all $(t, p) \in [\mathbf{T}_s, T] \times \mathbb{N}_{\geq 2}$, where, we obtain the second inequality by using the fact that $\nu - i^*(\nu) \mathbf{T}_s \leq \mathbf{T}_s, \forall \nu$. Substituting (B.56), and (B.57) into (B.55) leads to

$$\begin{aligned} \|\check{N}_a(t)\|_{L_p} &\leq \left(\int_{\mathbf{T}_s}^t e^{2\theta_1 \nu} d\nu \right) (p(4p-1))^{\frac{1}{2}} \left(\frac{1 - e^{-2\theta_2 \mathbf{T}_s}}{\theta_2} \right)^{\frac{1}{2}} \\ &\quad \times \sup_{\nu \in [\mathbf{T}_s, t]} \left(\|L_a(\nu)\|_{L_{4p}} |\varsigma_1(\nu)| \sup_{\beta \in [(i^*(\nu)-1)\mathbf{T}_s, i^*(\nu)\mathbf{T}_s]} |\varsigma_2(\beta)| \right) \\ &\quad \times \left(\frac{1 - e^{-\theta_2 t}}{\theta_2} \sup_{\nu \in [0, t]} \|S(\nu)\|_{L_{2p}} + \left(p \frac{2p-1}{2} \right)^{\frac{1}{2}} \left(\frac{1 - e^{-2\theta_2 t}}{\theta_2} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t]} \|L(\nu)\|_{L_{2p}} \right), \end{aligned}$$

for all $t \in [\mathbf{T}_s, T]$. Similarly, substituting (B.56), and (B.58) into (B.55) leads to

$$\begin{aligned} \|\check{N}_b(t)\|_{L_p} &\leq \left(\int_{\mathbf{T}_s}^t e^{2\theta_1 \nu} d\nu \right) (p(4p-1))^{\frac{1}{2}} \left(\frac{1 - e^{-2\theta_2 \mathbf{T}_s}}{\theta_2} \right)^{\frac{1}{2}} \sup_{\nu \in [\mathbf{T}_s, t]} \|L_b(\nu)\|_{L_{4p}} \\ &\quad \times \left(\frac{1 - e^{-\theta_2 t}}{\theta_2} \sup_{\nu \in [0, t]} \|S(\nu)\|_{L_{2p}} + \left(p \frac{2p-1}{2} \right)^{\frac{1}{2}} \left(\frac{1 - e^{-2\theta_2 t}}{\theta_2} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t]} \|L(\nu)\|_{L_{2p}} \right), \end{aligned}$$

for all $t \in [\mathbf{T}_s, T]$. Then, by solving the first integrals on the right hand side of the bounds above and substituting into (B.54) lead to (B.51) and (B.52), thus completing the proof. \square

C Reference Process

We provide the proof of Proposition 4.1 below.

Proof of Proposition 4.1. We consider the case $\mathbb{P}(Y_{N,0} = Y_0 \in U_N) = 1$ w.l.o.g., since the general case can be handled by following the approach in [106, Thm. 8.3.1].

We begin by observing that establishing the well-posedness of $Y_{N,t} = [(X_{N,t}^r)^\top \ (X_{N,t}^*)^\top]^\top$ is equivalent to establishing the individual well-posedness of the following systems:

$$dX_{N,t}^* = \bar{F}_{N,\mu}(t, X_{N,t}^*) dt + \bar{F}_{N,\sigma}(t, X_{N,t}^*) dW_t^*, \quad X_{N,0}^* = x_0^* \sim \xi_0^*, \quad (\text{C.1a})$$

$$dX_{N,t}^r = F_{N,\mu}(t, X_{N,t}^r, U_t^r) dt + F_{N,\sigma}(t, X_{N,t}^r) dW_t, \quad X_{N,0}^r = x_0 \sim \xi_0, \quad (\text{C.1b})$$

for $t \in [0, \tau_N \wedge T]$, where $\bar{F}_{N,\{\mu,\sigma\}}$ and $F_{N,\{\mu,\sigma\}}$ are defined analogously to $G_{N,\{\mu,\sigma\}}$ in (31).

The well-posedness of (C.1a) is straightforward to establish using Definition 1, Assumption 1 [48, Thm. 3.4], and due to $\bar{F}_{N,\{\mu,\sigma\}}(a) \equiv 0, \forall a \in \mathbb{R}^n$ with $\|a\| \geq 2N$.

Now, consider any $z \in \mathcal{M}_2([0, T], \mathbb{R}^n \mid \mathfrak{W}_{0,t})$, for any $t \in [0, T]$, and define

$$M(z(t)) = \int_0^t F_{N,\mu}(\nu, z(\nu), U_\nu^r) d\nu + \int_0^t F_{N,\sigma}(\nu, z(\nu)) dW_\nu, \quad t \in [0, T]. \quad (\text{C.2})$$

Let $f_N, \Lambda_{N,\mu}, p_N$, and $\Lambda_{N,\sigma}$ be the truncated versions of the functions f, Λ_μ, p , and Λ_σ , respectively, and where the truncation is defined as in (29). Then, we have that

$$M(z(t)) = \int_0^t (f_N(\nu, z(\nu)) + g(\nu)U_\nu^r + \Lambda_{N,\mu}(\nu, z(\nu))) d\nu + \int_0^t (p_N(\nu, z(\nu)) + \Lambda_{N,\sigma}(\nu, z(\nu))) dW_\nu,$$

for $t \in [0, T]$, where, from (27), we have that

$$U_\nu^r = \mathcal{F}_\omega (\Lambda_{N,\mu}^\parallel(\cdot, z))(\nu) + \mathcal{F}_{N,\omega} (p_N^\parallel(\cdot, z) + \Lambda_{N,\sigma}^\parallel(\cdot, z), W)(\nu).$$

Therefore, the previous expression can be expressed as

$$\begin{aligned} M(z(t)) &= \int_0^t (f_N(\nu, z(\nu)) + \Lambda_{N,\mu}(\nu, z(\nu))) d\nu + \int_0^t g(\nu) \mathcal{F}_\omega (\Lambda_{N,\mu}^\parallel(\cdot, z))(\nu) d\nu \\ &\quad + \int_0^t g(\nu) \mathcal{F}_{N,\omega} (p_N^\parallel(\cdot, z) + \Lambda_{N,\sigma}^\parallel(\cdot, z), W)(\nu) d\nu + \int_0^t (p_N(\nu, z(\nu)) + \Lambda_{N,\sigma}(\nu, z(\nu))) dW_\nu, \end{aligned} \quad (\text{C.3})$$

for $t \in [0, T]$. Using the definition of \mathcal{F}_ω in (15a), we have that

$$\int_0^t g(\nu) \mathcal{F}_\omega (\Lambda_{N,\mu}^\parallel(\cdot, z))(\nu) d\nu = -\omega \int_0^t \int_0^\nu g(\nu) e^{-\omega(\nu-\beta)} \Lambda_{N,\mu}^\parallel(\beta, z(\beta)) d\beta d\nu, \quad t \in [0, T].$$

Changing the order of integration in the double Lebesgue integral produces

$$\int_0^t g(\nu) \mathcal{F}_\omega (\Lambda_{N,\mu}^\parallel(\cdot, z))(\nu) d\nu = -\omega \int_0^t \left(\int_\nu^t g(\beta) e^{-\omega\beta} d\beta \right) e^{\omega\nu} \Lambda_{N,\mu}^\parallel(\nu, z(\nu)) d\nu, \quad t \in [0, T]. \quad (\text{C.4})$$

Next, using the definition of $\mathcal{F}_{N,\omega}$ in (27), we have that

$$\begin{aligned} &\int_0^t g(\nu) \mathcal{F}_{N,\omega} (p_N^\parallel(\cdot, z) + \Lambda_{N,\sigma}^\parallel(\cdot, z), W)(\nu) d\nu \\ &= -\omega \int_0^t \int_0^\nu g(\nu) e^{-\omega(\nu-\beta)} (p_N^\parallel(\beta, z(\beta)) + \Lambda_{N,\sigma}^\parallel(\beta, z(\beta))) dW_\beta d\nu, \quad t \in [0, T]. \end{aligned}$$

Applying Lemma B.1 to the above expression for

$$\begin{aligned} L(\nu) &= g(\nu) e^{-\omega\nu} \in \mathcal{C}([0, T]; \mathbb{R}^{n \times m}), \\ S(\beta) &= e^{\omega\beta} (p_N^\parallel(\beta, z(\beta)) + \Lambda_{N,\sigma}^\parallel(\beta, z(\beta))) \in \mathcal{M}([0, T]; \mathbb{R}^{m \times d} | \mathfrak{W}_{0,t}), Q_\beta = W_\beta, \end{aligned}$$

produces

$$\begin{aligned} &\int_0^t g(\nu) \mathcal{F}_{N,\omega} (p_N^\parallel(\cdot, z) + \Lambda_{N,\sigma}^\parallel(\cdot, z), W)(\nu) d\nu \\ &= -\omega \int_0^t \left(\int_\nu^t g(\beta) e^{-\omega\beta} d\beta \right) e^{\omega\nu} (p_N^\parallel(\nu, z(\nu)) + \Lambda_{N,\sigma}^\parallel(\nu, z(\nu))) dW_\nu, \quad t \in [0, T]. \end{aligned} \quad (\text{C.5})$$

Substituting (C.4) and (C.5) into (C.3) yields

$$\begin{aligned} M(z(t)) &= \int_0^t F_{N,\mu}(\nu, z(\nu), U_\nu^r) d\nu + \int_0^t F_{N,\sigma}(\nu, z(\nu)) dW_\nu \\ &= \int_0^t M_\mu(t, \nu, z(\nu)) d\nu + \int_0^t M_\sigma(t, \nu, z(\nu)) dW_\nu, \quad t \in [0, T], \end{aligned} \quad (\text{C.6})$$

where

$$M_\mu(t, \nu, z(\nu)) = f_N(\nu, z(\nu)) + \Lambda_{N,\mu}(\nu, z(\nu)) - \omega \left(\int_\nu^t g(\beta) e^{-\omega\beta} d\beta \right) e^{\omega\nu} \Lambda_{N,\mu}^\parallel(\nu, z(\nu)),$$

$$M_\sigma(t, \nu, z(\nu)) = p_N(\nu, z(\nu)) + \Lambda_{N,\sigma}(\nu, z(\nu)) - \omega \left(\int_\nu^t g(\beta) e^{-\omega\beta} d\beta \right) e^{\omega\nu} (p_N^\parallel(\nu, z(\nu)) + \Lambda_{N,\sigma}^\parallel(\nu, z(\nu))),$$

for $\nu \in [0, t]$, and $t \in [0, T]$. Using Assumption 1, we have that

$$\begin{aligned} &\left\| \left(\int_\nu^t g(\beta) e^{-\omega\beta} d\beta \right) e^{\omega\nu} \right\|_F \\ &\leq \int_\nu^t \|g(\beta)\|_F e^{-\omega\beta} d\beta e^{\omega\nu} \leq \Delta_g \frac{1 - e^{-\omega(t-\nu)}}{\omega} \leq \Delta_g \frac{1}{\omega}, \quad \forall \nu \in [0, t], t \in [0, T]. \end{aligned} \quad (\text{C.7})$$

Thus, we conclude that

$$\|M_\mu(t, \nu, z(\nu))\| \leq \|f_N(\nu, z(\nu))\| + \|\Lambda_{N,\mu}(\nu, z(\nu))\| + \Delta_g \|\Lambda_{N,\mu}^\parallel(\nu, z(\nu))\|, \quad (\text{C.8a})$$

$$\|M_\sigma(t, \nu, z(\nu))\|_F = \|p_N(\nu, z(\nu))\|_F + \|\Lambda_{N,\sigma}(\nu, z(\nu))\|_F + \Delta_g \|p_N^\parallel(\nu, z(\nu)) + \Lambda_{N,\sigma}^\parallel(\nu, z(\nu))\|_F, \quad (\text{C.8b})$$

for all $\nu \in [0, t]$, and $t \in [0, T]$.

Let us set $x(t) \equiv x_0$, and define the Picard iterates for (C.1b) as

$$x_k(t) = x_0 + \int_0^t F_{N,\mu}(\nu, x_{k-1}(\nu), U_\nu^r) d\nu + \int_0^t F_{N,\sigma}(\nu, x_{k-1}(\nu)) dW_\nu, \quad k \in \mathbb{N}, \quad t \in [0, T].$$

Then, by the definition of M in (C.2), we have that

$$x_k(t) = x_0 + M(x_{k-1}(t)) = x_0 + \int_0^t M_\mu(t, \nu, x_{k-1}(\nu)) d\nu + \int_0^t M_\sigma(t, \nu, x_{k-1}(\nu)) dW_\nu, \quad t \in [0, T].$$

Since the truncated functions f_N , $\Lambda_{N,\mu}$, p_N , and $\Lambda_{N,\sigma}$ agree with their non-truncated counterparts on $[0, \tau_N]$, we have that over the interval $[0, T] \supseteq [0, \tau_N]$, the Assumptions 1 and 4, the truncation definition in (31), along with the bounds in (C.8), imply linear growth of the integrands in the Picard iterates above. Therefore, as in the proof of [91, Thm. 2.3.1], we claim the existence of solutions to (C.1b) on $[0, T]$.

Similarly, using the assumptions of local Lipschitz continuity of f , and the global Lipschitz continuity of Λ_μ , p , and Λ_σ , we can use the same arguments as above to establish the uniqueness of solutions to (C.1b) on $[0, T]$ as in the proof of [91, Thm. 2.3.1].

Furthermore, as above, using the definition of the truncation in (31), we can use (C.8) and show linear growth bounds and Lipschitz continuity for $F_{N,\{\mu,\sigma\}}$, globally over \mathbb{R}^n and uniformly in $t \in \mathbb{R}_{\geq 0}$, thus implying the strong Markov property of the solutions by [91, Thm. 2.9.3].

Finally, since $G_{N,\{\mu,\sigma\}}(t, \cdot) = G_{\{\mu,\sigma\}}(t, \cdot)$ for all $t \in [0, \tau_N]$, we may invoke [48, Thm. 3.5], [60, Thm. 5.2.9] to conclude that $Y_{N,t}$ is a unique solution to (29) on $[0, \tau_N]$. \square

The next two results help us to derive the stochastic differential of $\mathcal{V}(Y_{N,t})$ in the proof of Lemma 4.1.

Proposition C.1 *Let $Y_{N,t}$ be the strong solution of (31), and let $\tau(t)$ be the stopping time defined in (34), Lemma 4.1. Moreover, define*

$$\mathcal{V}(Y_N) \doteq \|X_N^r - X_N^*\|^2, \quad \partial \mathcal{V}(Y_N) \doteq \partial_{X_N^r - X_N^*} \mathcal{V}(Y_N) = 2(X_N^r - X_N^*) \in \mathbb{R}^n.$$

Then,

$$\begin{aligned} & \int_0^{\tau(t)} e^{2\lambda\nu} \left(\nabla \mathcal{V}(Y_{N,\nu})^\top G_\mu(\nu, Y_{N,\nu}) + \frac{1}{2} \text{Tr} [H_\sigma(\nu, Y_{N,\nu}) \nabla^2 \mathcal{V}(Y_{N,\nu})] \right) d\nu \\ & \leq -2\lambda \int_0^{\tau(t)} e^{2\lambda\nu} \mathcal{V}(Y_{N,\nu}) d\nu + \int_0^{\tau(t)} e^{2\lambda\nu} \phi_U^r(\nu, Y_{N,\nu}) d\nu \\ & \quad + \int_0^{\tau(t)} e^{2\lambda\nu} \left(\phi_\mu^r(\nu, Y_{N,\nu}) + \phi_{\mu^\parallel}^r(\nu, Y_{N,\nu}) \right) d\nu, \quad (\text{C.9a}) \end{aligned}$$

$$\begin{aligned} & \int_0^{\tau(t)} e^{2\lambda\nu} \nabla \mathcal{V}(Y_{N,\nu})^\top G_\sigma(\nu, Y_{N,\nu}) d\widehat{W}_\nu \\ & = \int_0^{\tau(t)} e^{2\lambda\nu} \left(\phi_{\sigma^*}^r(\nu, Y_{N,\nu}) dW_\nu^* + [\phi_\sigma^r(\nu, Y_{N,\nu}) + \phi_{\sigma^\parallel}^r(\nu, Y_{N,\nu})] dW_\nu \right), \quad (\text{C.9b}) \end{aligned}$$

for all $t \in \mathbb{R}_{\geq 0}$, where $H_\sigma(\nu, Y_{N,\nu}) = G_\sigma(\nu, Y_{N,\nu}) G_\sigma(\nu, Y_{N,\nu})^\top$, $G_\mu(\nu, Y_{N,\nu})$ and $G_\sigma(\nu, Y_{N,\nu})$ are defined in (29). Additionally, we have defined

$$\begin{aligned} \phi_{\mu^\parallel}^r(\nu, Y_{N,\nu}) &= \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r), \\ \phi_\mu^r(\nu, Y_{N,\nu}) &= \partial \mathcal{V}(Y_N)^\top g(\nu)^\perp \Lambda_\mu^\perp(\nu, X_{N,\nu}^r) + \|F_\sigma(\nu, X_{N,\nu}^r)\|_F^2 + \|\bar{F}_\sigma(\nu, X_{N,\nu}^r)\|_F^2, \\ \phi_U^r(\nu, Y_{N,\nu}) &= \partial \mathcal{V}(Y_N)^\top g(\nu) U_\nu^r, \end{aligned}$$

and

$$\begin{aligned} \phi_{\sigma^\parallel}^r(\nu, Y_{N,\nu}) &= \partial \mathcal{V}(Y_N)^\top g(\nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r), \quad \phi_{\sigma^*}^r(\nu, Y_{N,\nu}) = -\partial \mathcal{V}(Y_N)^\top \bar{F}_\sigma(\nu, X_{N,\nu}^r), \\ \phi_\sigma^r(\nu, Y_{N,\nu}) &= \partial \mathcal{V}(Y_N)^\top g(\nu)^\perp F_\sigma^\perp(\nu, X_{N,\nu}^r). \end{aligned}$$

Proof. Using the definition of G_μ from (29) and the decomposition of F_μ in (3), we have that

$$\begin{aligned} \nabla \mathcal{V}(Y_{N,\nu})^\top G_\mu(\nu, Y_{N,\nu}) &= 2(X_{N,\nu}^r - X_{N,\nu}^*)^\top (\bar{F}_\mu(\nu, X_{N,\nu}^r) - \bar{F}_\mu(\nu, X_{N,\nu}^*)) \\ &\quad + 2(X_{N,\nu}^r - X_{N,\nu}^*)^\top (g(\nu)U_\nu^r + \Lambda_\mu(\nu, X_{N,\nu}^r)), \quad \nu \in [0, \tau(t)], \end{aligned}$$

which, due to Lemma 2.1 reduces to

$$\nabla \mathcal{V}(Y_{N,\nu})^\top G_\mu(\nu, Y_{N,\nu}) \leq -2\lambda \mathcal{V}(Y_{N,\nu}) + 2(X_{N,\nu}^r - X_{N,\nu}^*)^\top (g(\nu)U_\nu^r + \Lambda_\mu(\nu, X_{N,\nu}^r)), \quad (\text{C.10})$$

for all $\nu \in [0, \tau(t)]$. We develop the expression further by using (7) in Assumption 4 to write

$$\Lambda_\mu(\nu, X_{N,\nu}^r) = [g(\nu) \quad g(\nu)^\perp] \begin{bmatrix} \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) \\ \Lambda_\mu^\perp(\nu, X_{N,\nu}^r) \end{bmatrix} = g(\nu)\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) + g(\nu)^\perp \Lambda_\mu^\perp(\nu, X_{N,\nu}^r).$$

Substituting into (C.10) yields

$$\begin{aligned} \nabla \mathcal{V}(Y_{N,\nu})^\top G_\mu(\nu, Y_{N,\nu}) &\leq -2\lambda \mathcal{V}(Y_{N,\nu}) + 2(X_{N,\nu}^r - X_{N,\nu}^*)^\top g(\nu)^\perp \Lambda_\mu^\perp(\nu, X_{N,\nu}^r) \\ &\quad + 2(X_{N,\nu}^r - X_{N,\nu}^*)^\top g(\nu) (U_\nu^r + \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)), \quad \forall \nu \in [0, \tau(t)], \end{aligned}$$

which then leads to (C.9a) by further noting that

$$\begin{aligned} \frac{1}{2} \text{Tr} [H_\sigma(\nu, Y_{N,\nu}) \nabla^2 \mathcal{V}(Y_{N,\nu})] &= \|[F_\sigma(\nu, X_{N,\nu}^r) \quad -\bar{F}_\sigma(\nu, X_{N,\nu}^*)]\|_F^2 \\ &= \|F_\sigma(\nu, X_{N,\nu}^r)\|_F^2 + \|\bar{F}_\sigma(\nu, X_{N,\nu}^*)\|_F^2. \end{aligned}$$

Next, using the definition of G_σ from (29), we have that

$$\begin{aligned} &\int_0^{\tau(t)} e^{2\lambda\nu} \nabla \mathcal{V}(Y_{N,\nu})^\top G_\sigma(\nu, Y_{N,\nu}) d\widehat{W}_\nu \\ &= 2 \int_0^{\tau(t)} e^{2\lambda\nu} (X_{N,\nu}^r - X_{N,\nu}^*)^\top (F_\sigma(\nu, X_\nu^r) dW_\nu - \bar{F}_\sigma(\nu, X_\nu^*) dW_\nu^*) \\ &= -2 \int_0^{\tau(t)} e^{2\lambda\nu} (X_{N,\nu}^r - X_{N,\nu}^*)^\top \bar{F}_\sigma(\nu, X_\nu^*) dW_\nu^* \\ &\quad + 2 \int_0^{\tau(t)} e^{2\lambda\nu} (X_{N,\nu}^r - X_{N,\nu}^*)^\top (p(\nu, X_{N,\nu}^r) + \Lambda_\sigma(\nu, X_{N,\nu}^r)) dW_\nu, \quad t \in \mathbb{R}_{\geq 0}, \end{aligned}$$

where we have used the definition of F_σ in (1). Since (7) and (8) in Assumptions 4 and 5, respectively, along with Definition 3 imply that

$$\begin{aligned} p(\nu, X_{N,t}^r) + \Lambda_\sigma(\nu, X_{N,t}^r) &= g(\nu)^\perp p^\perp(\nu, X_{N,t}^r) + g(\nu)^\perp \Lambda_\sigma^\perp(\nu, X_{N,t}^r) + g(\nu)p^\parallel(\nu, X_{N,t}^r) + g(\nu)\Lambda_\sigma^\parallel(\nu, X_{N,t}^r) \\ &= g(\nu)^\perp F_\sigma^\perp(\nu, X_{N,t}^r) + g(\nu)F_\sigma^\parallel(\nu, X_{N,t}^r), \quad \forall \nu \in [0, \tau(t)], \end{aligned}$$

the previous integral equality can be re-written as

$$\begin{aligned} &\int_0^{\tau(t)} e^{2\lambda\nu} \nabla \mathcal{V}(Y_{N,\nu})^\top G_\sigma(\nu, Y_{N,\nu}) d\widehat{W}_\nu \\ &= -2 \int_0^{\tau(t)} e^{2\lambda\nu} (X_{N,\nu}^r - X_{N,\nu}^*)^\top \bar{F}_\sigma(\nu, X_{N,\nu}^*) dW_\nu^* + 2 \int_0^{\tau(t)} e^{2\lambda\nu} (X_{N,\nu}^r - X_{N,\nu}^*)^\top g(\nu)^\perp F_\sigma^\perp(\nu, X_{N,\nu}^r) dW_\nu \\ &\quad + 2 \int_0^{\tau(t)} e^{2\lambda\nu} (X_{N,\nu}^r - X_{N,\nu}^*)^\top g(\nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu, \quad t \in \mathbb{R}_{\geq 0}, \end{aligned}$$

thus establishing the expression in (C.9b). □

In the subsequent proposition, we derive the expression for how the reference feedback process U^r (27) enters the truncated joint process $Y_{N,t}$.

Proposition C.2 *Let $Y_{N,t}$ be the strong solution of (31), and let $\tau(t)$ be the stopping time defined in (34), Lemma 4.1. Then, for the term $\phi_{U^r}^r$ defined in the statement of Proposition C.1, we have that*

$$\begin{aligned} \int_0^{\tau(t)} e^{2\lambda\nu} \phi_{U^r}^r(\nu, Y_{N,\nu}) d\nu &= \int_0^{\tau(t)} \left(\hat{\mathcal{U}}_\mu^r(\tau(t), \nu, Y_N; \omega) d\nu + \hat{\mathcal{U}}_\sigma^r(\tau(t), \nu, Y_N; \omega) dW_\nu \right) \\ &\quad + \int_0^{\tau(t)} e^{2\lambda\nu} \left(\phi_{U_\mu}^r(\nu, Y_{N,\nu}; \omega) d\nu + \phi_{U_\sigma}^r(\nu, Y_{N,\nu}; \omega) dW_\nu \right), \quad t \in \mathbb{R}_{\geq 0}, \quad (\text{C.11}) \end{aligned}$$

where

$$\begin{aligned} \hat{\mathcal{U}}_\mu^r(\tau(t), \nu, Y_N; \omega) &= e^{-\omega\tau(t)} \frac{\omega}{2\lambda - \omega} \left(e^{\omega\tau(t)} \mathcal{P}^r(\tau(t), \nu) - e^{2\lambda\tau(t)} \partial \mathcal{V}(Y_{N,\tau(t)})^\top g(\tau(t)) \right) e^{\omega\nu} \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r), \\ \hat{\mathcal{U}}_\sigma^r(\tau(t), \nu, Y_N; \omega) &= e^{-\omega\tau(t)} \frac{\omega}{2\lambda - \omega} \left(e^{\omega\tau(t)} \mathcal{P}^r(\tau(t), \nu) - e^{2\lambda\tau(t)} \partial \mathcal{V}(Y_{N,\tau(t)})^\top g(\tau(t)) \right) e^{\omega\nu} F_\sigma^\parallel(\nu, X_{N,\nu}^r), \\ \phi_{U_\mu}^r(\nu, Y_{N,\nu}; \omega) &= \frac{\omega}{2\lambda - \omega} \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r), \\ \phi_{U_\sigma}^r(\nu, Y_{N,\nu}; \omega) &= \frac{\omega}{2\lambda - \omega} \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r), \end{aligned}$$

and

$$\mathcal{P}^r(\tau(t), \nu) = \int_\nu^{\tau(t)} e^{(2\lambda - \omega)\beta} d\beta \left[\partial \mathcal{V}(Y_{N,\beta})^\top g(\beta) \right].$$

Proof. Using the definition of $\phi_{U^r}^r$ in the statement of Proposition C.1, we have that

$$\begin{aligned} &\int_0^{\tau(t)} e^{2\lambda\nu} \phi_{U^r}^r(\nu, Y_{N,\nu}) d\nu \\ &= \int_0^{\tau(t)} e^{2\lambda\nu} \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) U_\nu^r d\nu \\ &= 2 \int_0^{\tau(t)} e^{2\lambda\nu} \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) \left(\mathcal{F}_\omega^\parallel(\Lambda_\mu^\parallel(\cdot, X^r))(\nu) + \mathcal{F}_{\mathcal{N},\omega}(p^\parallel(\cdot, X^r) + \Lambda_\sigma^\parallel(\cdot, X^r), W)(\nu) \right) d\nu, \end{aligned}$$

for all $t \geq 0$, where we have incorporated the definition of U^r from (27). Using the definitions of \mathcal{F}_ω and $\mathcal{F}_{\mathcal{N},\omega}(\cdot, W)$ in (15a) and (28), respectively, we get

$$\begin{aligned} &\int_0^{\tau(t)} e^{2\lambda\nu} \phi_{U^r}^r(\nu, Y_{N,\nu}) d\nu \\ &= \int_0^{\tau(t)} \int_0^\nu \left(-\omega e^{(2\lambda - \omega)\nu} \right) \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) \left(e^{\omega\beta} \Lambda_\mu^\parallel(\beta, X_{N,\beta}^r) d\beta \right) d\nu \\ &\quad + \int_0^{\tau(t)} \int_0^\nu \left(-\omega e^{(2\lambda - \omega)\nu} \right) \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) \left(e^{\omega\beta} F_\sigma^\parallel(\beta, X_{N,\beta}^r) dW_\beta \right) d\nu, \end{aligned}$$

for all $t \in \mathbb{R}_{\geq 0}$. Changing the order of integration in the first integral on the right hand side, and applying Lemma B.1 to the second integral, we get:

$$\begin{aligned} &\int_0^{\tau(t)} e^{2\lambda\nu} \phi_{U^r}^r(\nu, Y_{N,\nu}) d\nu \\ &= \int_0^{\tau(t)} \left(- \int_\nu^{\tau(t)} \omega e^{(2\lambda - \omega)\beta} \partial \mathcal{V}(Y_{N,\beta})^\top g(\beta) d\beta \right) e^{\omega\nu} \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu \end{aligned}$$

$$+ \int_0^{\tau(t)} \left(- \int_{\nu}^{\tau(t)} \omega e^{(2\lambda-\omega)\beta} \partial \mathcal{V}(Y_{N,\beta})^\top g(\beta) d\beta \right) e^{\omega\nu} F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu, \quad (\text{C.12})$$

for all $t \in \mathbb{R}_{\geq 0}$, where in the first integral, we switch between the variables β and ν after changing the order of integration. Now, the chain rule implies that

$$d_\beta \left[e^{(2\lambda-\omega)\beta} \partial \mathcal{V}(Y_{N,\beta})^\top g(\beta) \right] = \left(\frac{d}{d\beta} e^{(2\lambda-\omega)\beta} \right) \partial \mathcal{V}(Y_{N,\beta})^\top g(\beta) + e^{(2\lambda-\omega)\beta} d_\beta \left[\partial \mathcal{V}(Y_{N,\beta})^\top g(\beta) \right],$$

where $d_\beta[\cdot]$ in the last term is interpreted as the stochastic differential with respect to the variable β . Multiplying both sides by $-\frac{\omega}{2\lambda-\omega}$ yields

$$\begin{aligned} & -\frac{\omega}{2\lambda-\omega} d_\beta \left[e^{(2\lambda-\omega)\beta} \partial \mathcal{V}(Y_{N,\beta})^\top g(\beta) \right] \\ &= -\frac{\omega}{2\lambda-\omega} \left(\frac{d}{d\beta} e^{(2\lambda-\omega)\beta} \right) \partial \mathcal{V}(Y_{N,\beta})^\top g(\beta) - \frac{\omega}{2\lambda-\omega} e^{(2\lambda-\omega)\beta} d_\beta \left[\partial \mathcal{V}(Y_{N,\beta})^\top g(\beta) \right] \\ &= -\omega e^{(2\lambda-\omega)\beta} \partial \mathcal{V}(Y_{N,\beta})^\top g(\beta) - \frac{\omega}{2\lambda-\omega} e^{(2\lambda-\omega)\beta} d_\beta \left[\partial \mathcal{V}(Y_{N,\beta})^\top g(\beta) \right], \end{aligned}$$

and thus

$$\begin{aligned} -\omega e^{(2\lambda-\omega)\beta} \partial \mathcal{V}(Y_{N,\beta})^\top g(\beta) &= -\frac{\omega}{2\lambda-\omega} d_\beta \left[e^{(2\lambda-\omega)\beta} \partial \mathcal{V}(Y_{N,\beta})^\top g(\beta) \right] \\ &\quad + \frac{\omega}{2\lambda-\omega} e^{(2\lambda-\omega)\beta} d_\beta \left[\partial \mathcal{V}(Y_{N,\beta})^\top g(\beta) \right]. \end{aligned}$$

Integrating over the interval $[\nu, \tau(t)]$ leads to

$$\begin{aligned} & - \int_{\nu}^{\tau(t)} \omega e^{(2\lambda-\omega)\beta} \partial \mathcal{V}(Y_{N,\beta})^\top g(\beta) d\beta \\ &= -\frac{\omega}{2\lambda-\omega} e^{(2\lambda-\omega)\tau(t)} \partial \mathcal{V}(Y_{N,\tau(t)})^\top g(\tau(t)) + \frac{\omega}{2\lambda-\omega} e^{(2\lambda-\omega)\nu} \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) \\ &\quad + \frac{\omega}{2\lambda-\omega} \mathcal{P}^r(\tau(t), \nu), \end{aligned}$$

where

$$\mathcal{P}^r(\tau(t), \nu) = \int_{\nu}^{\tau(t)} e^{(2\lambda-\omega)\beta} d_\beta \left[\partial \mathcal{V}(Y_{N,\beta})^\top g(\beta) \right].$$

Substituting the expression above for the two identical inner integrals on the right hand side of (C.12) produces (C.11), thus completing the proof upon re-arranging terms. \square

The next result provides an alternative representation of \mathcal{P}^r that is amenable to the analysis of the reference process.

Proposition C.3 Recall the expression for $\mathcal{P}^r(\tau(t), \nu)$ in (41) in the statement of Lemma 4.1, which we restate below:

$$\mathcal{P}^r(\tau(t), \nu) = \int_{\nu}^{\tau(t)} e^{(2\lambda-\omega)\beta} d_\beta \left[\partial \mathcal{V}(Y_{N,\beta})^\top g(\beta) \right] \in \mathbb{R}^{1 \times m}, \quad 0 \leq \nu \leq \tau(t), \quad (\text{C.13})$$

where $\tau(t)$ is defined in (34), and $\partial \mathcal{V}(Y_{N,t}) \doteq \partial_{X_N^r - X_N^*} \mathcal{V}(Y_N) = 2(X_N^r - X_N^*) \in \mathbb{R}^n$. Then, $\mathcal{P}^r(\tau(t), \nu)$ admits the following representation:

$$\mathcal{P}^r(\tau(t), \nu) = \mathcal{P}_o^r(\tau(t), \nu) + \mathcal{P}_{ad}^r(\tau(t), \nu) \in \mathbb{R}^{1 \times m}, \quad 0 \leq \nu \leq \tau(t), \quad t \in \mathbb{R}_{\geq 0}, \quad (\text{C.14})$$

where

$$\mathcal{P}_o^r(\tau(t), \nu) = \int_{\nu}^{\tau(t)} e^{(2\lambda-\omega)\beta} (\mathcal{P}_\mu^r(\beta) d\beta + \mathcal{P}_\sigma^r(\beta) dW_\beta + \mathcal{P}_{\sigma_*}^r(\beta) dW_\beta^*)^\top \in \mathbb{R}^{1 \times m}, \quad (\text{C.15a})$$

$$\mathcal{P}_{ad}^r(\tau(t), \nu) = \int_{\nu}^{\tau(t)} e^{(2\lambda-\omega)\beta} \mathcal{P}_u^r(\beta)^\top d\beta \in \mathbb{R}^{1 \times m}, \quad (\text{C.15b})$$

and where $\mathcal{P}_\mu^r(\beta)$, $\mathcal{P}_\sigma^r(\beta) \in \mathbb{R}^m$ and $\mathcal{P}_\mu^r(\beta)$, $\mathcal{P}_\sigma^r(\beta) \in \mathbb{R}^{m \times d}$ are defined as

$$\begin{aligned}\mathcal{P}_\mu^r(\beta) &= \dot{g}(\beta)^\top \partial \mathcal{V}(Y_{N,\beta}) + 2g(\beta)^\top (\bar{F}_\mu(\beta, X_{N,\beta}^r) - \bar{F}_\mu(\beta, X_{N,\beta}^*) + \Lambda_\mu(\beta, X_{N,\beta}^r)), \\ \mathcal{P}_\mu^r(\beta) &= 2g(\beta)^\top g(\beta) U_\beta^r, \\ \mathcal{P}_\sigma^r(\beta) &= 2g(\beta)^\top F_\sigma(\beta, X_{N,\beta}^r), \quad \mathcal{P}_\sigma^r(\beta) = -2g(\beta)^\top \bar{F}_\sigma(\beta, X_{N,\beta}^*).\end{aligned}$$

Proof. We begin by writing $\partial \mathcal{V}(Y_{N,\beta})^\top g(\beta)$ as

$$\begin{aligned}\partial \mathcal{V}(Y_{N,\beta})^\top g(\beta) &= [\partial \mathcal{V}(Y_{N,\beta})^\top g_{\cdot,1}(\beta) \quad \cdots \quad \partial \mathcal{V}(Y_{N,\beta})^\top g_{\cdot,m}(\beta)] \\ &= [\sum_{i=1}^n \partial_i \mathcal{V}(Y_{N,\beta}) g_{i,1}(\beta) \quad \cdots \quad \sum_{i=1}^n \partial_i \mathcal{V}(Y_{N,\beta}) g_{i,m}(\beta)] \in \mathbb{R}^{1 \times m},\end{aligned}\tag{C.16}$$

where $g_{\cdot,j}(\beta) \in \mathbb{R}^n$ is the j -th column of $g(\beta)$, and

$$\partial_i \mathcal{V}(Y_{N,\beta}) \doteq \frac{\partial}{\partial (X_{N,\beta}^r - X_{N,\beta}^*)_i} \mathcal{V}(Y_{N,\beta}) = 2(X_{N,\beta}^r - X_{N,\beta}^*)_i \in \mathbb{R}, \quad i \in \{1, \dots, n\}.$$

Applying Itô's lemma to $\partial_i \mathcal{V}(Y_{N,\beta}) g_{i,j}(\beta) \in \mathbb{R}$, $(i, j) \in \{1, \dots, n\} \times \{1, \dots, m\}$, and using the truncated dynamics in (31) we get

$$\begin{aligned}d_\beta [\partial_i \mathcal{V}(Y_{N,\beta}) g_{i,j}(\beta)] &= \left(\partial_i \mathcal{V}(Y_{N,\beta}) \dot{g}_{i,j}(\beta) + \nabla \partial_i \mathcal{V}(Y_{N,\beta})^\top G_\mu(\beta, Y_{N,\beta}) g_{i,j}(\beta) \right) d\beta \\ &\quad + \nabla \partial_i \mathcal{V}(Y_{N,\beta})^\top G_\sigma(\beta, Y_{N,\beta}) g_{i,j}(\beta) d\widehat{W}_\beta,\end{aligned}\tag{C.17}$$

where we have used the straightforward fact that $\nabla^2 \partial_i \mathcal{V}(Y_{N,\beta}) = 0 \in \mathbb{S}^{2n}$. We have further replaced $G_{N,\mu}$ and $G_{N,\sigma}$ with G_μ and G_σ because from Proposition 4.1, $Y_{N,\beta}$ is also a strong solution of the joint process (29) for all $\beta \in [\nu, \tau(t)] \subseteq [0, \tau^*] \subseteq [0, \tau_N]$. Observe from (C.16) that

$$d_\beta [\partial \mathcal{V}(Y_{N,\beta})^\top g_{\cdot,j}(\beta)] = \sum_{i=1}^n d_\beta [\partial_i \mathcal{V}(Y_{N,\beta}) g_{i,j}(\beta)] \in \mathbb{R}, \quad j \in \{1, \dots, m\}.$$

It therefore follows from the expression in (C.17)

$$\begin{aligned}d_\beta [\partial \mathcal{V}(Y_{N,\beta})^\top g_{\cdot,j}(\beta)] &= \sum_{i=1}^n \left(\partial_i \mathcal{V}(Y_{N,\beta}) \dot{g}_{i,j}(\beta) + \nabla \partial_i \mathcal{V}(Y_{N,\beta})^\top G_\mu(\beta, Y_{N,\beta}) g_{i,j}(\beta) \right) d\beta \\ &\quad + \sum_{i=1}^n \nabla \partial_i \mathcal{V}(Y_{N,\beta})^\top G_\sigma(\beta, Y_{N,\beta}) g_{i,j}(\beta) d\widehat{W}_\beta,\end{aligned}$$

which can be further written as

$$\begin{aligned}d_\beta [\partial \mathcal{V}(Y_{N,\beta})^\top g_{\cdot,j}(\beta)] &= \left(\partial \mathcal{V}(Y_{N,\beta})^\top \dot{g}_{\cdot,j}(\beta) + g_{\cdot,j}(\beta)^\top \nabla \partial \mathcal{V}(Y_{N,\beta})^\top G_\mu(\beta, Y_{N,\beta}) \right) d\beta \\ &\quad + g_{\cdot,j}(\beta)^\top \nabla \partial \mathcal{V}(Y_{N,\beta})^\top G_\sigma(\beta, Y_{N,\beta}) d\widehat{W}_\beta \in \mathbb{R},\end{aligned}\tag{C.18}$$

for $j \in \{1, \dots, m\}$. Once again, (C.16) implies that

$$d_\beta [\partial \mathcal{V}(Y_{N,\beta})^\top g(\beta)] = [d_\beta [\partial \mathcal{V}(Y_{N,\beta})^\top g_{\cdot,1}(\beta)] \quad \cdots \quad d_\beta [\partial \mathcal{V}(Y_{N,\beta})^\top g_{\cdot,m}(\beta)]] \in \mathbb{R}^{1 \times m}.$$

Thus, we use the expression in (C.18) to write

$$\begin{aligned}d_\beta [\partial \mathcal{V}(Y_{N,\beta})^\top g(\beta)] &= \left(\partial \mathcal{V}(Y_{N,\beta})^\top \dot{g}(\beta) + (g(\beta)^\top \nabla \partial \mathcal{V}(Y_{N,\beta})^\top G_\mu(\beta, Y_{N,\beta}))^\top \right) d\beta \\ &\quad + (g(\beta)^\top \nabla \partial \mathcal{V}(Y_{N,\beta})^\top G_\sigma(\beta, Y_{N,\beta}) d\widehat{W}_\beta)^\top \in \mathbb{R}^{1 \times m}.\end{aligned}\tag{C.19}$$

Now, using the definition of the gradient of vector valued functions in Sec. 1.4, we observe that

$$\nabla \partial \mathcal{V}(Y_{N,\beta}) = 2 \left[\nabla (X_{N,\beta}^r - X_{N,\beta}^*)_1 \quad \cdots \quad \nabla (X_{N,\beta}^r - X_{N,\beta}^*)_n \right] = 2 \begin{bmatrix} \mathbb{I}_n \\ -\mathbb{I}_n \end{bmatrix} \in \mathbb{R}^{2n \times n}.$$

Thus, it follows from the definitions of the vector fields in (29) that

$$\begin{aligned}\nabla \partial \mathcal{V}(Y_{N,\beta})^\top G_\mu(\beta, Y_{N,\beta}) &= 2(F_\mu(\beta, X_{N,\beta}^r, U_\beta^r) - \bar{F}_\mu(\beta, X_{N,\beta}^*)) \\ &= 2\bar{F}_\mu(\beta, X_{N,\beta}^r) + 2\Lambda_\mu(\beta, X_{N,\beta}^r) - 2\bar{F}_\mu(\beta, X_{N,\beta}^*) + 2g(\beta)U_\beta^r \in \mathbb{R}^n, \\ \nabla \partial \mathcal{V}(Y_{N,\beta})^\top G_\sigma(\beta, Y_{N,\beta}) &= 2\begin{bmatrix} F_\sigma(\beta, X_{N,\beta}^r) & -\bar{F}_\sigma(\beta, X_{N,\beta}^*) \end{bmatrix} \in \mathbb{R}^{n \times 2d},\end{aligned}$$

where we have used the following decomposition from (3) in Definition 1:

$$F_\mu(\beta, X_{N,\beta}^r, U_\beta^r) = \bar{F}_\mu(\beta, X_{N,\beta}^r) + g(\beta)U_\beta^r + \Lambda_\mu(\beta, X_{N,\beta}^r).$$

Substituting the above identities into (C.19) produces

$$d_\beta \left[\partial \mathcal{V}(Y_{N,\beta})^\top g(\beta) \right] = [\mathcal{P}_\mu^r(\beta) + \mathcal{P}_\sigma^r(\beta)]^\top d\beta + [\mathcal{P}_\sigma^r(\beta)dW_\beta + \mathcal{P}_{\sigma_*}^r(\beta)dW_\beta^*]^\top \in \mathbb{R}^{1 \times m}.$$

Then, (C.14) is established by substituting the above into (C.13). □

The next result establishes the bounds for the pertinent entities in the last proposition.

Proposition C.4 *If the stopping time τ^* , defined in (34), Lemma 4.1, satisfies $\tau^* = t^*$, then the functions $\mathcal{P}_\mu^r(t) \in \mathbb{R}^m$ and $\mathcal{P}_\sigma^r(t), \mathcal{P}_{\sigma_*}^r(t) \in \mathbb{R}^{m \times d}$ defined in the statement of Proposition C.3 satisfy the following bounds for any $\mathbf{p} \in \mathbb{R}_{\geq 1}$:*

$$\sup_{\nu \in [0, t^*]} \|\mathcal{P}_\mu^r(\nu)\|_{L_{2p}}^{y_0} \leq 2\Delta_g \Delta_{\mathcal{P}_\mu}^r(\mathbf{p}) + \Delta_{\dot{g}} \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Y_{N,\nu})\|_{L_{2p}}^{y_0}, \quad (\text{C.20a})$$

$$\sum_{i \in \{\sigma, \sigma_*\}} \sup_{\nu \in [0, t^*]} \|\mathcal{P}_i^r(\nu)\|_{L_{2p}}^{y_0} \leq 2\Delta_g \Delta_{\mathcal{P}_\sigma}^r(\mathbf{p}), \quad (\text{C.20b})$$

where

$$\begin{aligned}\Delta_{\mathcal{P}_\mu}^r(\mathbf{p}) &= \sup_{\nu \in [0, t^*]} \left(\|\bar{F}_\mu(\nu, X_{N,\nu}^r) - \bar{F}_\mu(\nu, X_{N,\nu}^*)\|_{L_{2p}}^{y_0} + \|\Lambda_\mu(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} \right), \\ \Delta_{\mathcal{P}_\sigma}^r(\mathbf{p}) &= \sup_{\nu \in [0, t^*]} \left(\|\bar{F}_\sigma(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} + \|\Lambda_\sigma(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} \right) + \sup_{\nu \in [0, t^*]} \|\bar{F}_\sigma(\nu, X_{N,\nu}^*)\|_{L_{2p}}^{y_0}.\end{aligned}$$

Proof. We begin by using the definition of \mathcal{P}_μ^r in (C.15a) to obtain

$$\|\mathcal{P}_\mu^r(\nu)\| \leq \Delta_{\dot{g}} \|\partial \mathcal{V}(Y_{N,\nu})\| + 2\Delta_g (\|\bar{F}_\mu(\nu, X_{N,\nu}^r) - \bar{F}_\mu(\nu, X_{N,\nu}^*)\| + \|\Lambda_\mu(\nu, X_{N,\nu}^r)\|),$$

for all $\nu \in [0, t^*]$, where we have used the bounds on $g(t)$ and $\dot{g}(t)$ in Assumption 1. It then follows from the Minkowski's inequality that

$$\|\mathcal{P}_\mu^r(\nu)\|_{L_{2p}}^{y_0} \leq \Delta_{\dot{g}} \|\partial \mathcal{V}(Y_{N,\nu})\|_{L_{2p}}^{y_0} + 2\Delta_g \left(\|\bar{F}_\mu(\nu, X_{N,\nu}^r) - \bar{F}_\mu(\nu, X_{N,\nu}^*)\|_{L_{2p}}^{y_0} + \|\Lambda_\mu(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} \right),$$

for all $\nu \in [0, t^*]$, which in turn implies (C.20a).

Next, using the definitions of \mathcal{P}_σ^r and $\mathcal{P}_{\sigma_*}^r$, we obtain

$$\|\mathcal{P}_\sigma^r(\nu)\|_F \leq 2\Delta_g \|F_\sigma(\nu, X_{N,\nu}^r)\|_F, \quad \|\mathcal{P}_{\sigma_*}^r(\nu)\|_F \leq 2\Delta_g \|\bar{F}_\sigma(\nu, X_{N,\nu}^*)\|_F, \quad \forall \nu \in [0, t^*],$$

where we have used the bound on $g(t)$ in Assumption 1. It then follows from the decomposition $F_\sigma = \bar{F}_\sigma + \Lambda_\sigma$ in (3) that

$$\begin{aligned}\|\mathcal{P}_\sigma^r(\nu)\|_F &\leq 2\Delta_g (\|\bar{F}_\sigma(\nu, X_{N,\nu}^r)\|_F + \|\Lambda_\sigma(\nu, X_{N,\nu}^r)\|_F), \\ \|\mathcal{P}_{\sigma_*}^r(\nu)\|_F &\leq 2\Delta_g \|\bar{F}_\sigma(\nu, X_{N,\nu}^*)\|_F, \quad \forall \nu \in [0, t^*].\end{aligned}$$

From the Minkowski's inequality, we get

$$\begin{aligned}\|\mathcal{P}_\sigma^r(\nu)\|_{L_{2p}}^{y_0} &\leq 2\Delta_g (\|\bar{F}_\sigma(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} + \|\Lambda_\sigma(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0}), \\ \|\mathcal{P}_{\sigma_*}^r(\nu)\|_{L_{2p}}^{y_0} &\leq 2\Delta_g \|\bar{F}_\sigma(\nu, X_{N,\nu}^*)\|_{L_{2p}}^{y_0}, \quad \forall \nu \in [0, t^*],\end{aligned}$$

which then leads to (C.20b) in a straightforward manner. □

Next, we provide a result that is essential to the proof of the main results of the section.

Proposition C.5 Consider the following scalar processes:

$$\begin{aligned} N_1^r(t) &= \int_0^t e^{\omega\nu} M_\mu(t, \nu) \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu, & N_2^r(t) &= \int_0^t e^{\omega\nu} M_\mu(t, \nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu, \\ N_3^r(t) &= \int_0^t e^{\omega\nu} M_\sigma(t, \nu) \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu, & N_4^r(t) &= \int_0^t e^{\omega\nu} M_\sigma(t, \nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu, \end{aligned} \quad (\text{C.21})$$

where

$$M_\mu(t, \nu) = \int_\nu^t e^{(2\lambda-\omega)\beta} \mathcal{P}_\mu^r(\beta)^\top d\beta, \quad M_\sigma(t, \nu) = \int_\nu^t e^{(2\lambda-\omega)\beta} [\mathcal{P}_\sigma^r(\beta) dW_\beta + \mathcal{P}_{\sigma_*}^r(\beta) dW_\beta^*]^\top,$$

for $0 \leq \nu \leq t \leq T$, and where $\mathcal{P}_\mu^r(\beta) \in \mathbb{R}^m$ and $\mathcal{P}_\sigma^r(\beta), \mathcal{P}_{\sigma_*}^r(\beta) \in \mathbb{R}^{m \times d}$ are defined in the statement of Proposition C.3. If the stopping time τ^* , defined in (34), Lemma 4.1, satisfies $\tau^* = t^*$, then we have the following bounds:

$$\begin{aligned} \sum_{i=1}^4 |\mathbb{E}_{y_0} [N_i^r(t)]| &\leq \frac{e^{2\lambda t}}{\sqrt{\lambda\omega}} \Delta_{\mathcal{P}_1}^r(1, t^*) \sup_{\nu \in [0, t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)\|_{L_2}^{y_0} \\ &\quad + \left(\frac{e^{2\lambda t}}{\sqrt{\lambda\omega}} \Delta_{\mathcal{P}_2}^r(1, t^*) + \frac{e^{2\lambda t}}{\lambda} \Delta_{\mathcal{P}_3}^r(1, t^*) \right) \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu}^r)\|_{L_2}^{y_0}, \end{aligned} \quad (\text{C.22})$$

for all $t \in [0, t^*]$, and

$$\begin{aligned} \sum_{i=1}^4 \|N_i^r(t)\|_{L_p}^{y_0} &\leq \frac{e^{2\lambda t}}{\sqrt{\lambda\omega}} \Delta_{\mathcal{P}_1}^r(\mathbf{p}, t^*) \sup_{\nu \in [0, t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2\mathbf{p}}}^{y_0} \\ &\quad + \left(\frac{e^{2\lambda t}}{\sqrt{\lambda\omega}} \Delta_{\mathcal{P}_2}^r(\mathbf{p}, t^*) + \frac{e^{2\lambda t}}{\lambda} \Delta_{\mathcal{P}_3}^r(\mathbf{p}, t^*) \right) \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2\mathbf{p}}}^{y_0}, \end{aligned} \quad (\text{C.23})$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 2}$, where

$$\begin{aligned} \Delta_{\mathcal{P}_1}^r(\mathbf{p}, t^*) &= \frac{1}{2\sqrt{\lambda}} \sup_{\nu \in [0, t^*]} \|\mathcal{P}_\mu^r(\nu)\|_{L_{2\mathbf{p}}}^{y_0} + \frac{\mathbf{p}(\mathbf{p})}{2} \sup_{\nu \in [0, t^*]} \left(\|\mathcal{P}_\sigma^r(\nu)\|_{L_{2\mathbf{p}}}^{y_0} + \|\mathcal{P}_{\sigma_*}^r(\nu)\|_{L_{2\mathbf{p}}}^{y_0} \right), \\ \Delta_{\mathcal{P}_2}^r(\mathbf{p}, t^*) &= \mathbf{p}'(\mathbf{p}) \Delta_{\mathcal{P}_1}^r(\mathbf{p}, t^*), \quad \Delta_{\mathcal{P}_3}^r(\mathbf{p}, t^*) = \frac{\sqrt{m}}{2} \sup_{\nu \in [0, t^*]} \left(\|\mathcal{P}_\sigma^r(\nu)\|_{L_{2\mathbf{p}}}^{y_0} + \|\mathcal{P}_{\sigma_*}^r(\nu)\|_{L_{2\mathbf{p}}}^{y_0} \right), \end{aligned}$$

and where the constants $\mathbf{p}(\mathbf{p})$ and $\mathbf{p}'(\mathbf{p})$ are defined in (A.1a).

Proof. We begin with the process N_1^r and use the definition of M_μ to obtain

$$N_1^r(t) = \int_0^t e^{\omega\nu} \int_\nu^t e^{(2\lambda-\omega)\beta} [\mathcal{P}_\mu^r(\beta)]^\top d\beta \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu, \quad t \in [0, t^*].$$

Changing the order of integration and switching the variables β and ν leads to

$$N_1^r(t) = \int_0^t e^{(2\lambda-\omega)\nu} [\mathcal{P}_\mu^r(\nu)]^\top \left(\int_0^\nu e^{\omega\beta} \Lambda_\mu^\parallel(\beta, X_{N,\beta}^r) d\beta \right) d\nu, \quad t \in [0, t^*].$$

Hence,

$$|N_1^r(t)| \leq \int_0^t e^{(2\lambda-\omega)\nu} \|\mathcal{P}_\mu^r(\nu)\| \left(\int_0^\nu e^{\omega\beta} \|\Lambda_\mu^\parallel(\beta, X_{N,\beta}^r)\| d\beta \right) d\nu, \quad t \in [0, t^*].$$

Applying Proposition B.2 for $\{\theta_1, \theta_2, \xi, S_1, S_2\} = \{2\lambda, \omega, 0, \mathcal{P}_\mu^r, \Lambda_\mu^\parallel\}$ to the integral on the right hand side of the above inequality produces

$$\|N_1^r(t)\|_{L_p}^{y_0} \leq \frac{e^{2\lambda t}}{2\lambda\omega} \sup_{\nu \in [0, t]} \|\mathcal{P}_\mu^r(\nu)\|_{L_{2\mathbf{p}}}^{y_0} \sup_{\nu \in [0, t]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2\mathbf{p}}}^{y_0}$$

$$\leq \frac{e^{2\lambda t}}{2\lambda\omega} \sup_{\nu \in [0, t^*]} \|\mathcal{P}_\mu^r(\nu)\|_{L_{2p}}^{y_0} \sup_{\nu \in [0, t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0}, \quad \forall (t, p) \in [0, t^*] \times \mathbb{N}_{\geq 1}, \quad (\text{C.24})$$

where we have bounded $e^{2\lambda t} - 1$ by $e^{2\lambda t}$.

Next, from the definitions of $M_\mu(t, \nu)$ and N_2^r we obtain

$$\begin{aligned} N_2^r(t) &= \int_0^t e^{\omega\nu} M_\mu(t, \nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu \\ &= \int_0^t e^{\omega\nu} \int_\nu^t e^{(2\lambda-\omega)\beta} [\mathcal{P}_\mu^r(\beta)]^\top d\beta F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu \\ &= \int_0^t \int_\nu^t e^{(2\lambda-\omega)\beta} [\mathcal{P}_\mu^r(\beta)]^\top e^{\omega\nu} [F_\sigma^\parallel(\nu, X_{N,\nu}^r) \quad 0_{m,d}] d\beta d\widehat{W}_\nu, \end{aligned}$$

where recall from Definition 5 that $\widehat{W}_t = [(W_t)^\top \quad (W_t^*)^\top]^\top \in \mathbb{R}^{2d}$. Using the definition of \mathcal{P}_μ^r in Proposition C.3, and the regularity assumptions in Sec. 2.2, it is straightforward to show that $\mathcal{P}_\mu^r \in \mathcal{M}_2^{loc}(\mathbb{R}^m | \mathfrak{W}_t \times \mathfrak{W}_t^*)$ and $[F_\sigma^\parallel(\nu, X_{N,\nu}^r) \quad 0_{m,d}] \in \mathcal{M}_2^{loc}(\mathbb{R}^{m \times 2d} | \mathfrak{W}_t \times \mathfrak{W}_t^*)$. Therefore, we may apply Lemma B.1 to the right hand side of the previous expression and obtain

$$N_2^r(t) = \int_0^t e^{(2\lambda-\omega)\nu} [\mathcal{P}_\mu^r(\nu)]^\top \left(\int_0^\nu e^{\omega\beta} [F_\sigma^\parallel(\beta, X_{N,\beta}^r) \quad 0_{m,d}] d\widehat{W}_\beta \right) d\nu \in \mathbb{R},$$

and thus

$$|N_2^r(t)| \leq \int_0^t e^{(2\lambda-\omega)\nu} \|\mathcal{P}_\mu^r(\nu)\| \left\| \int_0^\nu e^{\omega\beta} [F_\sigma^\parallel(\beta, X_{N,\beta}^r) \quad 0_{m,d}] d\widehat{W}_\beta \right\| d\nu.$$

Since the process on the right hand side can be cast in the form of the process $N(t)$ in Lemma B.3 by setting

$$\begin{aligned} Q_t &= \widehat{W}_t \in \mathbb{R}^{2d} \quad (n_q = 2d), \quad \mathfrak{F}_t = \mathfrak{W}_t \times \mathfrak{W}_t^*, \quad \theta_1 = 2\lambda, \quad \theta_2 = \omega, \quad \xi = 0, \quad \zeta_1 = 0, \quad \zeta_2 = t, \\ S(t) &= \mathcal{P}_\mu^r(t) \in \mathcal{M}_2^{loc}(\mathbb{R}^m | \mathfrak{W}_t \times \mathfrak{W}_t^*) \quad (n_s = m), \\ L(t) &= [F_\sigma^\parallel(t, X_{N,t}^r) \quad 0_{m,d}] \in \mathcal{M}_2^{loc}(\mathbb{R}^{m \times 2d} | \mathfrak{W}_t \times \mathfrak{W}_t^*), \end{aligned}$$

we may apply Lemma B.3 to obtain the following bound:

$$\begin{aligned} \|N_2^r(t)\|_{L_p}^{y_0} &\leq \left(p \frac{2p-1}{2}\right)^{\frac{1}{2}} \frac{e^{2\lambda t}}{2\lambda\sqrt{\omega}} \sup_{\nu \in [0, t]} \|\mathcal{P}_\mu^r(\nu)\|_{L_{2p}}^{y_0} \sup_{\nu \in [0, t]} \|F_\sigma^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} \\ &\leq \left(p \frac{2p-1}{2}\right)^{\frac{1}{2}} \frac{e^{2\lambda t}}{2\lambda\sqrt{\omega}} \sup_{\nu \in [0, t^*]} \|\mathcal{P}_\mu^r(\nu)\|_{L_{2p}}^{y_0} \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0}, \end{aligned} \quad (\text{C.25})$$

for all $(t, p) \in [0, t^*] \times \mathbb{N}_{\geq 1}$.

Next, we consider the process $N_3^r(t)$ given by

$$\begin{aligned} N_3^r(t) &= \int_0^t e^{\omega\nu} M_\sigma(t, \nu) \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu \\ &= \int_0^t e^{\omega\nu} \left(\int_\nu^t e^{(2\lambda-\omega)\beta} [\mathcal{P}_\sigma^r(\beta) dW_\beta + \mathcal{P}_{\sigma_*}^r(\beta) dW_\beta^*]^\top \right) \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu \\ &= \int_0^t \int_\nu^t e^{\omega\nu} \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)^\top e^{(2\lambda-\omega)\beta} [\mathcal{P}_\sigma^r(\beta) \quad \mathcal{P}_{\sigma_*}^r(\beta)] d\widehat{W}_\beta d\nu. \end{aligned}$$

Using Lemma B.1, one sees that

$$\begin{aligned} N_3^r(t) &= \int_0^t \int_0^\nu e^{\omega\beta} \Lambda_\mu^\parallel(\beta, X_{N,\beta}^r)^\top e^{(2\lambda-\omega)\nu} [\mathcal{P}_\sigma^r(\nu) \quad \mathcal{P}_{\sigma_*}^r(\nu)] d\beta d\widehat{W}_\nu \\ &= \int_0^t e^{(2\lambda-\omega)\nu} \left(\int_0^\nu e^{\omega\beta} \Lambda_\mu^\parallel(\beta, X_{N,\beta}^r) d\beta \right)^\top [\mathcal{P}_\sigma^r(\nu) \quad \mathcal{P}_{\sigma_*}^r(\nu)] d\widehat{W}_\nu. \end{aligned}$$

Note that $N_3^r(t)$ can be cast in the form of the process $N(t)$ in Lemma B.4 by setting

$$\begin{aligned} Q_t &= \widehat{W}_t \in \mathbb{R}^{2d} (n_q = 2d), \quad \mathfrak{F}_t = \mathfrak{W}_t \times \mathfrak{W}_t^*, \quad \theta_1 = 2\lambda, \quad \theta_2 = \omega, \\ S(t) &= \Lambda_\mu^\parallel(t, X_{N,t}^r) \in \mathcal{M}_2^{loc}(\mathbb{R}^m | \mathfrak{W}_t \times \mathfrak{W}_t^*) (n_o = m), \\ L(t) &= [\mathcal{P}_\sigma^r(t) \quad \mathcal{P}_{\sigma_*}^r(t)] \in \mathcal{M}_2^{loc}(\mathbb{R}^{m \times 2d} | \mathfrak{W}_t \times \mathfrak{W}_t^*). \end{aligned}$$

Therefore, we use Lemma B.4 and obtain

$$\begin{aligned} \|N_3^r(t)\|_{L_p}^{y_0} &\leq (p(p-1))^{\frac{1}{2}} \frac{e^{2\lambda t}}{2\sqrt{2\lambda\omega}} \sup_{\nu \in [0,t]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2p}} \sup_{\nu \in [0,t]} \|\begin{bmatrix} \mathcal{P}_\sigma^r(\nu) & \mathcal{P}_{\sigma_*}^r(\nu) \end{bmatrix}\|_{L_{2p}} \\ &\leq (p(p-1))^{\frac{1}{2}} \frac{e^{2\lambda t}}{2\sqrt{2\lambda\omega}} \sup_{\nu \in [0,t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2p}} \sup_{\nu \in [0,t^*]} \|\begin{bmatrix} \mathcal{P}_\sigma^r(\nu) & \mathcal{P}_{\sigma_*}^r(\nu) \end{bmatrix}\|_{L_{2p}}, \end{aligned} \quad (C.26)$$

for all $(t, p) \in [0, t^*] \times \mathbb{N}_{\geq 2}$. Since

$$\|\begin{bmatrix} \mathcal{P}_\sigma^r(\nu) & \mathcal{P}_{\sigma_*}^r(\nu) \end{bmatrix}\|_F = \left(\|\mathcal{P}_\sigma^r(\nu)\|_F^2 + \|\mathcal{P}_{\sigma_*}^r(\nu)\|_F^2 \right)^{\frac{1}{2}} \leq \|\mathcal{P}_\sigma^r(\nu)\|_F + \|\mathcal{P}_{\sigma_*}^r(\nu)\|_F, \quad \forall \nu \in [0, t^*],$$

we may use the Minkowski's inequality to obtain

$$\|\begin{bmatrix} \mathcal{P}_\sigma^r(\nu) & \mathcal{P}_{\sigma_*}^r(\nu) \end{bmatrix}\|_{L_{2p}} \leq \|\mathcal{P}_\sigma^r(\nu)\|_{L_{2p}} + \|\mathcal{P}_{\sigma_*}^r(\nu)\|_{L_{2p}}, \quad \forall \nu \in [0, t^*]. \quad (C.27)$$

Hence, (C.26) can be further developed into

$$\|N_3^r(t)\|_{L_p}^{y_0} \leq (p(p-1))^{\frac{1}{2}} \frac{e^{2\lambda t}}{2\sqrt{2\lambda\omega}} \sup_{\nu \in [0,t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} \sup_{\nu \in [0,t^*]} \left(\|\mathcal{P}_\sigma^r(\nu)\|_{L_{2p}}^{y_0} + \|\mathcal{P}_{\sigma_*}^r(\nu)\|_{L_{2p}}^{y_0} \right), \quad (C.28)$$

for all $(t, p) \in [0, t^*] \times \mathbb{N}_{\geq 2}$. Additionally, using [91, Thm. 1.5.21], one obtains

$$\mathbb{E}_{y_0} [N_3^r(t)] = 0, \quad \forall t \in [0, t^*]. \quad (C.29)$$

Finally, consider the process $N_4^r(t)$ given as

$$\begin{aligned} N_4^r(t) &= \int_0^t e^{\omega\nu} M_\sigma(t, \nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu \\ &= \int_0^t e^{\omega\nu} \int_\nu^t e^{(2\lambda-\omega)\beta} [\mathcal{P}_\sigma^r(\beta) dW_\beta + \mathcal{P}_{\sigma_*}^r(\beta) dW_\beta^*]^\top F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu \\ &= \int_0^t e^{\omega\nu} \left(\int_\nu^t e^{(2\lambda-\omega)\beta} [\mathcal{P}_\sigma^r(\beta) \quad \mathcal{P}_{\sigma_*}^r(\beta)] d\widehat{W}_\beta \right)^\top [F_\sigma^\parallel(\nu, X_{N,\nu}^r) \quad 0_{m,d}] d\widehat{W}_\nu, \end{aligned}$$

which can be cast in the form of the process $N(t)$ in Lemma B.5 by choosing

$$\begin{aligned} Q_t &= \widehat{W}_t \in \mathbb{R}^{2d} (n_q = 2d), \quad \mathfrak{F}_t = \mathfrak{W}_t \times \mathfrak{W}_t^*, \quad \theta_1 = 2\lambda, \quad \theta_2 = \omega, \\ L_1(t) &= [\mathcal{P}_\sigma^r(t) \quad \mathcal{P}_{\sigma_*}^r(t)] \in \mathcal{M}_2^{loc}(\mathbb{R}^{m \times 2d} | \mathfrak{W}_t \times \mathfrak{W}_t^*) (n_l = m), \\ L_2(t) &= [F_\sigma^\parallel(t, X_{N,t}^r) \quad 0_{m,d}] \in \mathcal{M}_2^{loc}(\mathbb{R}^{m \times 2d} | \mathfrak{W}_t \times \mathfrak{W}_t^*). \end{aligned}$$

It then follows from Lemma B.5 that

$$\mathbb{E} [N_4^r(t)] \leq \sqrt{m} \frac{e^{2\lambda t}}{2\lambda} \sup_{\nu \in [0,t^*]} \mathbb{E} \left[\left\| \begin{bmatrix} \mathcal{P}_\sigma^r(\nu) & \mathcal{P}_{\sigma_*}^r(\nu) \end{bmatrix} F_\sigma^\parallel(\nu, X_{N,\nu}^r)^\top \right\|_F \right], \quad \forall t \in [0, t^*], \quad (C.30)$$

and

$$\begin{aligned} \|N_4^r(t)\|_{L_p} &\leq \frac{p\sqrt{(p-1)(2p-1)}}{4} \frac{e^{2\lambda t}}{\sqrt{\lambda\omega}} \sup_{\nu \in [0,t^*]} \|\begin{bmatrix} \mathcal{P}_\sigma^r(\nu) & \mathcal{P}_{\sigma_*}^r(\nu) \end{bmatrix}\|_{L_{2p}} \sup_{\nu \in [0,t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2p}} \\ &\quad + \sqrt{m} \frac{e^{2\lambda t}}{2\lambda} \sup_{\nu \in [0,t^*]} \left\| \begin{bmatrix} \mathcal{P}_\sigma^r(\nu) & \mathcal{P}_{\sigma_*}^r(\nu) \end{bmatrix} F_\sigma^\parallel(\nu, X_{N,\nu}^r)^\top \right\|_{L_p}, \end{aligned} \quad (C.31)$$

for all $(t, p) \in [0, t^*] \times \mathbb{N}_{\geq 2}$. Now, using the Cauchy-Schwarz inequality and (C.27), one obtains

$$\begin{aligned} \left\| [\mathcal{P}_\sigma^r(\nu) \quad \mathcal{P}_{\sigma_*}^r(\nu)] F_\sigma^\parallel(\nu, X_{N,\nu}^r)^\top \right\|_{L_q} &\leq \left\| [\mathcal{P}_\sigma^r(\nu) \quad \mathcal{P}_{\sigma_*}^r(\nu)] \right\|_{L_{2q}} \left\| F_\sigma^\parallel(\nu, X_{N,\nu}^r)^\top \right\|_{L_{2q}} \\ &\leq \left(\|\mathcal{P}_\sigma^r(\nu)\|_{L_{2q}}^{y_0} + \|\mathcal{P}_{\sigma_*}^r(\nu)\|_{L_{2q}}^{y_0} \right) \left\| F_\sigma^\parallel(\nu, X_{N,\nu}^r)^\top \right\|_{L_{2q}}^{y_0}, \end{aligned}$$

for all $(\nu, q) \in [0, t^*] \times \mathbb{N}_{\geq 1}$. Using this inequality for $q = 1$ and $q = p \in \mathbb{N}_{\geq 2}$, the bounds in (C.30) and (C.31) can be further developed into

$$\mathbb{E}_{y_0} [N_4^r(t)] \leq \sqrt{m} \frac{e^{2\lambda t}}{2\lambda} \sup_{\nu \in [0, t^*]} \left(\|\mathcal{P}_\sigma^r(\nu)\|_{L_2}^{y_0} + \|\mathcal{P}_{\sigma_*}^r(\nu)\|_{L_2}^{y_0} \right) \sup_{\nu \in [0, t^*]} \left\| F_\sigma^\parallel(\nu, X_{N,\nu}^r)^\top \right\|_{L_2}^{y_0}, \quad \forall t \in [0, t^*], \quad (\text{C.32})$$

and

$$\begin{aligned} \|N_4^r(t)\|_{L_p}^{y_0} &\leq \left(\frac{p\sqrt{(p-1)(2p-1)}}{4} \frac{e^{2\lambda t}}{\sqrt{\lambda\omega}} + \sqrt{m} \frac{e^{2\lambda t}}{2\lambda} \right) \\ &\quad \times \sup_{\nu \in [0, t^*]} \left(\|\mathcal{P}_\sigma^r(\nu)\|_{L_{2p}}^{y_0} + \|\mathcal{P}_{\sigma_*}^r(\nu)\|_{L_{2p}}^{y_0} \right) \sup_{\nu \in [0, t^*]} \left\| F_\sigma^\parallel(\nu, X_{N,\nu}^r)^\top \right\|_{L_{2p}}^{y_0}, \quad (\text{C.33}) \end{aligned}$$

for all $(t, p) \in [0, t^*] \times \mathbb{N}_{\geq 2}$. Then, the bound in (C.22) is established by adding together (C.24), (C.25), (C.29), and (C.32). Similarly, the bounds in (C.24), (C.25), (C.28), and (C.33), lead to (C.23). \square

Similar to the last proposition, the following result establishes the bound on the input to the reference system.

Proposition C.6 Consider the following scalar process:

$$N_{\mathcal{U}}^r(t) = \int_0^t e^{\omega\nu} M_{\mathcal{U}}(t, \nu) (\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu), \quad (\text{C.34})$$

where

$$M_{\mathcal{U}}(t, \nu) = \int_\nu^t e^{(2\lambda - \omega)\beta} \mathcal{P}_{\mathcal{U}}^r(\beta)^\top d\beta,$$

for $0 \leq \nu \leq t \leq T$, and where $\mathcal{P}_{\mathcal{U}}^r(\beta) \in \mathbb{R}^m$ is defined in the statement of Proposition C.3. If the stopping time τ^* , defined in (34), Lemma 4.1, satisfies $\tau^* = t^*$, then

$$\|N_{\mathcal{U}}^r(t)\|_{L_p}^{y_0} \leq \Delta_g^2 \frac{e^{2\lambda t}}{\lambda} \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^*]} \left\| \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) \right\|_{L_{2p}}^{y_0} + \left(p \frac{2p-1}{2} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t^*]} \left\| F_\sigma^\parallel(\nu, X_{N,\nu}^r) \right\|_{L_{2p}}^{y_0} \right)^2, \quad (\text{C.35})$$

for all $(t, p) \in [0, t^*] \times \mathbb{N}_{\geq 1}$.

Proof. We begin by decomposing the process $N_{\mathcal{U}}^r(t)$ as follows:

$$N_{\mathcal{U}}^r(t) = N_{\mathcal{U}_1}^r(t) + N_{\mathcal{U}_2}^r(t), \quad (\text{C.36})$$

where

$$N_{\mathcal{U}_1}^r(t) = \int_0^t e^{\omega\nu} M_{\mathcal{U}}(t, \nu) \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu, \quad N_{\mathcal{U}_2}^r(t) = \int_0^t e^{\omega\nu} M_{\mathcal{U}}(t, \nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu.$$

Now, consider the process $N_{\mathcal{U}_1}^r(t)$, which, using the definition of $M_{\mathcal{U}}$ can be written as

$$\begin{aligned} N_{\mathcal{U}_1}^r(t) &= \int_0^t e^{\omega\nu} M_{\mathcal{U}}(t, \nu) \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu \\ &= \int_0^t e^{\omega\nu} \int_\nu^t e^{(2\lambda - \omega)\beta} \mathcal{P}_{\mathcal{U}}^r(\beta)^\top d\beta \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu \\ &= \int_0^t \int_\nu^t e^{\omega\nu} e^{(2\lambda - \omega)\beta} \mathcal{P}_{\mathcal{U}}^r(\beta)^\top \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\beta d\nu. \end{aligned}$$

Since the integral above is a nested Lebesgue integral with t -continuous integrands, we may change the order of integration as follows:

$$\begin{aligned} N_{\mathcal{U}_1}^r(t) &= \int_0^t \int_0^\beta e^{\omega\nu} e^{(2\lambda-\omega)\beta} \mathcal{P}_{\mathcal{U}}^r(\beta)^\top \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu d\beta \\ &= \int_0^t e^{(2\lambda-\omega)\beta} \mathcal{P}_{\mathcal{U}}^r(\beta)^\top \left(\int_0^\beta e^{\omega\nu} \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu \right) d\beta \\ &= \int_0^t e^{(2\lambda-\omega)\nu} \mathcal{P}_{\mathcal{U}}^r(\nu)^\top \left(\int_0^\nu e^{\omega\beta} \Lambda_\mu^\parallel(\beta, X_{N,\beta}^r) d\beta \right) d\nu, \quad (\text{C.37}) \end{aligned}$$

where in the last integral we have switched between the variables β and ν .

Next, consider the process $N_{\mathcal{U}_2}^r(t)$, which, using the definition of $M_{\mathcal{U}}$ can be written as

$$\begin{aligned} N_{\mathcal{U}_2}^r(t) &= \int_0^t e^{\omega\nu} M_{\mathcal{U}}(t, \nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu \\ &= \int_0^t \int_\nu^t \left(e^{(2\lambda-\omega)\beta} \mathcal{P}_{\mathcal{U}}^r(\beta)^\top \right) \left(e^{\omega\nu} F_\sigma^\parallel(\nu, X_{N,\nu}^r) \right) d\beta dW_\nu \\ &= \int_0^t \int_\nu^t \left(e^{(2\lambda-\omega)\beta} \mathcal{P}_{\mathcal{U}}^r(\beta)^\top \right) \left(e^{\omega\nu} \begin{bmatrix} F_\sigma^\parallel(\nu, X_{N,\nu}^r) & 0_{m,d} \end{bmatrix} \right) d\beta d\widehat{W}_\nu, \end{aligned}$$

where once again recall from Definition 5 that $\widehat{W}_t = [(W_t)^\top \ (W_t^*)^\top]^\top \in \mathbb{R}^{2d}$. Using the definition of $P_{\mathcal{U}}^r$ in Proposition C.3, and the regularity assumptions in Sec. 2.2, it is straightforward to show that $P_{\mathcal{U}}^r \in \mathcal{M}_2^{loc}(\mathbb{R}^m | \mathfrak{W}_t \times \mathfrak{W}_t^*)$ and $[F_\sigma^\parallel(\nu, X_{N,\nu}^r) \ 0_{m,d}] \in \mathcal{M}_2^{loc}(\mathbb{R}^{m \times 2d} | \mathfrak{W}_t \times \mathfrak{W}_t^*)$. Therefore, we invoke Lemma B.1 to obtain

$$\begin{aligned} N_{\mathcal{U}_2}^r(t) &= \int_0^t e^{\omega\nu} M_{\mathcal{U}}(t, \nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu \\ &= \int_0^t \left(e^{(2\lambda-\omega)\nu} \mathcal{P}_{\mathcal{U}}^r(\nu)^\top \right) \int_0^\nu \left(e^{\omega\beta} \begin{bmatrix} F_\sigma^\parallel(\beta, X_{N,\beta}^r) & 0_{m,d} \end{bmatrix} \right) d\widehat{W}_\beta d\nu \\ &= \int_0^t e^{(2\lambda-\omega)\nu} \mathcal{P}_{\mathcal{U}}^r(\nu)^\top \left(\int_0^\nu e^{\omega\beta} F_\sigma^\parallel(\beta, X_{N,\beta}^r) dW_\beta \right) d\nu. \quad (\text{C.38}) \end{aligned}$$

Substituting (C.37) and (C.38) into (C.36) yields

$$N_{\mathcal{U}}^r(t) = \int_0^t e^{(2\lambda-\omega)\nu} \mathcal{P}_{\mathcal{U}}^r(\nu)^\top \left(\int_0^\nu e^{\omega\beta} [\Lambda_\mu^\parallel(\beta, X_{N,\beta}^r) d\beta + F_\sigma^\parallel(\beta, X_{N,\beta}^r) dW_\beta] \right) d\nu. \quad (\text{C.39})$$

Substituting the definition of $\mathcal{P}_{\mathcal{U}}^r$ from Proposition C.3 into the above:

$$\begin{aligned} \mathcal{P}_{\mathcal{U}}^r(t) &= 2g(t)^\top g(t) U_t^r \\ &\stackrel{(\star)}{=} 2g(t)^\top g(t) \left(\mathcal{F}_\omega(\Lambda_\mu^\parallel(\cdot, X_N^r))(t) + \mathcal{F}_{\mathcal{N},\omega}(F_\sigma^\parallel(\cdot, X_N^r), W)(t) \right) \\ &= 2g(t)^\top g(t) \left(-\omega \int_0^t e^{-\omega(t-\nu)} \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu - \omega \int_0^t e^{-\omega(t-\nu)} F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu \right) \\ &= -2\omega e^{-\omega t} g(t)^\top g(t) \left(\int_0^t e^{\omega\nu} [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \right), \end{aligned}$$

where (\star) follows from the definition of U_t^r in (27) for the truncated process (31). Substituting the above expression into (C.39) produces

$$\begin{aligned} N_{\mathcal{U}}^r(t) &= -2\omega \int_0^t e^{2(\lambda-\omega)\nu} \left(\int_0^\nu e^{\omega\beta} [\Lambda_\mu^\parallel(\beta, X_{N,\beta}^r) d\beta + F_\sigma^\parallel(\beta, X_{N,\beta}^r) dW_\beta] \right)^\top g(\nu)^\top g(\nu) \\ &\quad \times \left(\int_0^\nu e^{\omega\beta} [\Lambda_\mu^\parallel(\beta, X_{N,\beta}^r) d\beta + F_\sigma^\parallel(\beta, X_{N,\beta}^r) dW_\beta] \right) d\nu. \quad (\text{C.40}) \end{aligned}$$

Observe that $N_{\mathcal{U}}^r(t)$ can be expressed as the process $N(t)$ in the statement of Lemma B.6 by setting

$$\begin{aligned} Q_t &= \widehat{W}_t \in \mathbb{R}^{2d} \ (n_q = 2d), \quad \mathfrak{F}_t = \mathfrak{W}_t \times \mathfrak{W}_t^*, \quad \theta_1 = \lambda, \quad \theta_2 = \omega, \quad \xi = 0, \\ R(t) &= -2\omega g(t)^\top g(t) \in \mathcal{M}_2^{loc}(\mathbb{S}^m | \mathfrak{W}_t \times \mathfrak{W}_t^*), \\ S_1(t) &= S_2(t) = \Lambda_\mu^\parallel(t, X_{N,t}^r) \in \mathcal{M}_2^{loc}(\mathbb{R}^m | \mathfrak{W}_t \times \mathfrak{W}_t^*), \\ L_1(t) &= L_2(t) = [F_\sigma^\parallel(t, X_{N,t}^r) \quad 0_{m,d}] \in \mathcal{M}_2^{loc}(\mathbb{R}^{m \times 2d} | \mathfrak{W}_t \times \mathfrak{W}_t^*). \end{aligned}$$

Furthermore, as a consequence of Assumption 1, we may set Δ_R in the hypothesis of Lemma B.6 as $\Delta_R = 2\Delta_g^2\omega$. Hence, applying Lemma B.6 to the process $N_{\mathcal{U}}^r(t)$ in (C.40) leads to

$$\begin{aligned} \|N_{\mathcal{U}}^r(t)\|_{L_p}^{y_0} &\leq \Delta_g^2\omega \frac{e^{2\lambda t} - 1}{\lambda} \left(\frac{1 - e^{-\omega t}}{\omega} \sup_{\nu \in [0, t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} \right. \\ &\quad \left. + \left(\frac{2p-1}{2} \right)^{\frac{1}{2}} \left(\frac{1 - e^{-2\omega t}}{\omega} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} \right)^2, \end{aligned}$$

for all $(t, p) \in [0, t^*] \times \mathbb{N}_{\geq 1}$. Then, bounding $e^{2\lambda t} - 1$ by $e^{2\lambda t}$, and $1 - e^{-\omega t}$ and $1 - e^{-2\omega t}$ by 1 produces desired result in (C.35). \square

The next lemma establishes the bound on Ξ^r .

Lemma C.1 *If the stopping time τ^* , defined in (34), Lemma 4.1, satisfies $\tau^* = t^*$, then*

$$\mathbb{E}_{y_0} [e^{-2\lambda t} \Xi^r(t, Y_N)] \leq \frac{\Delta_{\Xi_1}^r(1, t^*)}{2\lambda}, \quad \forall t \in [0, t^*], \quad (\text{C.41a})$$

$$\|e^{-2\lambda t} \Xi^r(t, Y_N)\|_{L_p}^{y_0} \leq \frac{\Delta_{\Xi_1}^r(p, t^*)}{2\lambda} + \frac{\Delta_{\Xi_2}^r(p, t^*)}{2\sqrt{\lambda}}, \quad \forall (t, p) \in [0, t^*] \times \mathbb{N}_{\geq 2}, \quad (\text{C.41b})$$

where the term $\Xi^r(\tau(t), Y_N)$ is defined in (38), Lemma 4.1, and $\Delta_{\Xi_1}^r(p, t^*)$ and $\Delta_{\Xi_2}^r(p, t^*)$ are defined for $p \in \mathbb{N}_{\geq 1}$ as follows:

$$\begin{aligned} \Delta_{\Xi_1}^r(p, t^*) &= \sup_{\nu \in [0, t^*]} \left(\Delta_g^\perp \|\partial \mathcal{V}(Y_{N,\nu})\|_{L_{2p}}^{y_0} \|\Lambda_\mu^\perp(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} \right. \\ &\quad \left. + \mathbb{E}_{y_0} \left[F_\sigma(\nu, X_{N,\nu}^r)^{2p} \right]^{\frac{1}{p}} + \mathbb{E}_{y_0} \left[\bar{F}_\sigma(\nu, X_{N,\nu}^*)^{2p} \right]^{\frac{1}{p}} \right), \\ \Delta_{\Xi_2}^r(p, t^*) &= p(p) \sup_{\nu \in [0, t^*]} \left(\Delta_g^\perp \|F_\sigma^\perp(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} + \|\bar{F}_\sigma(\nu, X_{N,\nu}^*)\|_{L_{2p}}^{y_0} \right) \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Y_{N,\nu})\|_{L_{2p}}^{y_0}, \end{aligned}$$

where the constant $p(p)$ is defined in (A.1a).

Proof. We begin with the definition of Ξ^r in (38) to write

$$\begin{aligned} e^{-2\lambda t} \Xi^r(t, Y_N) &= e^{-2\lambda t} \int_0^t e^{2\lambda \nu} \phi_\mu^r(\nu, Y_{N,\nu}) d\nu \\ &\quad + e^{-2\lambda t} \int_0^t e^{2\lambda \nu} [\phi_\sigma^r(\nu, Y_{N,\nu}) - \phi_{\sigma^*}^r(\nu, Y_{N,\nu})] d\widehat{W}_\nu, \quad t \in [0, t^*], \quad (\text{C.42}) \end{aligned}$$

where $\widehat{W}_\nu \in \mathbb{R}^{2d}$ is defined in Definition 5. The regularity assumptions and the t -continuity of the strong solutions $X_{N,t}^r$ and $X_{N,t}^*$ imply that $[\phi_\sigma^r(\nu, Y_{N,\nu}) - \phi_{\sigma^*}^r(\nu, Y_{N,\nu})] \in \mathcal{M}_2^{loc}(\mathbb{R}^{1 \times 2d} | \mathfrak{W}_t \times \mathfrak{W}_t^*)$. Hence, taking the expectation of (C.42) and using the linearity of the expectation operator, along with [91, Thm. 1.5.21], leads to

$$\begin{aligned} \mathbb{E}_{y_0} [e^{-2\lambda t} \Xi^r(t, Y_N)] &= e^{-2\lambda t} \mathbb{E}_{y_0} \left[\int_0^t e^{2\lambda \nu} \phi_\mu^r(\nu, Y_{N,\nu}) d\nu \right] \\ &\leq e^{-2\lambda t} \mathbb{E}_{y_0} \left[\int_0^t e^{2\lambda \nu} |\phi_\mu^r(\nu, Y_{N,\nu})| d\nu \right], \quad t \in [0, t^*]. \quad (\text{C.43}) \end{aligned}$$

In the above, we have used the fact that since $\tau^* = t^*$, and t^* is a constant,

$$\mathbb{E} [e^{-2\lambda t}] = e^{-2\lambda t}, \forall t \in [0, t^*].$$

Moreover, (C.42) and the Minkowski's inequality imply that

$$\begin{aligned} \|e^{-2\lambda t} \Xi^r(t, Y_N)\|_{L_p}^{y_0} &\leq e^{-2\lambda t} \left\| \int_0^t e^{2\lambda \nu} |\phi_\mu^r(\nu, Y_{N,\nu})| d\nu \right\|_{L_p}^{y_0} \\ &\quad + e^{-2\lambda t} \left\| \int_0^t e^{2\lambda \nu} [\phi_\sigma^r(\nu, Y_{N,\nu}) \quad \phi_{\sigma_*}^r(\nu, Y_{N,\nu})] d\widehat{W}_\nu \right\|_{L_p}^{y_0}, \end{aligned} \quad (\text{C.44})$$

for all $(t, p) \in [0, t^*] \times \mathbb{N}_{\geq 2}$.

Now, the regularity assumptions and the t -continuity of the strong solutions $X_{N,t}^r$ and $X_{N,t}^*$ over $[0, t^*] \subseteq [0, \tau_N]$ imply that $\phi_\mu^r(\nu, Y_{N,\nu}) \in \mathcal{M}_2^{loc}(\mathbb{R}|\mathfrak{W}_t \times \mathfrak{W}_t^*)$. Thus, we invoke Proposition B.1 for $\{\theta, \xi, n_s, S\} = \{2\lambda, 0, 1, \phi_\mu^r(\cdot, Y_N)\}$ to obtain

$$\left\| \int_0^t e^{2\lambda \nu} |\phi_\mu^r(\nu, Y_{N,\nu})| d\nu \right\|_{L_p}^{y_0} \leq \frac{e^{2\lambda t}}{2\lambda} \sup_{\nu \in [0, t]} \|\phi_\mu^r(\nu, Y_{N,\nu})\|_{L_p}^{y_0} \leq \frac{e^{2\lambda t}}{2\lambda} \sup_{\nu \in [0, t^*]} \|\phi_\mu^r(\nu, Y_{N,\nu})\|_{L_p}^{y_0}, \quad (\text{C.45})$$

for all $(t, p) \in [0, t^*] \times \mathbb{N}_{\geq 1}$, where we have bounded $e^{2\lambda t} - 1$ by $e^{2\lambda t}$. Now, we obtain the following from the definition of ϕ_μ^r in (39):

$$\begin{aligned} |\phi_\mu^r(\nu, Y_{N,\nu})| &\leq \|\partial \mathcal{V}(Y_{N,\nu})\| \|g(\nu)^\perp\|_F \|\Lambda_\mu^\perp(\nu, X_{N,\nu}^r)\| + \|F_\sigma(\nu, X_{N,\nu}^r)\|_F^2 + \|\bar{F}_\sigma(\nu, X_{N,\nu}^*)\|_F^2 \\ &\leq \Delta_g^\perp \|\partial \mathcal{V}(Y_{N,\nu})\| \|\Lambda_\mu^\perp(\nu, X_{N,\nu}^r)\| + \|F_\sigma(\nu, X_{N,\nu}^r)\|_F^2 + \|\bar{F}_\sigma(\nu, X_{N,\nu}^*)\|_F^2, \end{aligned}$$

where we have used the bound on $g(t)^\perp$ from Assumption 1. Using the Minkowski's inequality followed by the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \|\phi_\mu^r(t, Y_{N,t})\|_{L_p}^{y_0} &\leq \Delta_g^\perp \|\partial \mathcal{V}(Y_{N,t})\|_{L_{2p}}^{y_0} \|\Lambda_\mu^\perp(\nu, X_{N,t}^r)\|_{L_{2p}}^{y_0} \\ &\quad + \mathbb{E}_{y_0} [F_\sigma(\nu, X_{N,t}^r)^{2p}]^{\frac{1}{p}} + \mathbb{E}_{y_0} [\bar{F}_\sigma(\nu, X_{N,t}^*)^{2p}]^{\frac{1}{p}}, \quad \forall (t, p) \in [0, t^*] \times \mathbb{N}_{\geq 1}. \end{aligned}$$

Substituting the above bound into (C.45) produces

$$\begin{aligned} \left\| \int_0^t e^{2\lambda \nu} |\phi_\mu^r(\nu, Y_{N,\nu})| d\nu \right\|_{L_p}^{y_0} &\leq \frac{e^{2\lambda t}}{2\lambda} \sup_{\nu \in [0, t^*]} \left(\Delta_g^\perp \|\partial \mathcal{V}(Y_{N,\nu})\|_{L_{2p}}^{y_0} \|\Lambda_\mu^\perp(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} \right. \\ &\quad \left. + \mathbb{E}_{y_0} [F_\sigma(\nu, X_{N,\nu}^r)^{2p}]^{\frac{1}{p}} + \mathbb{E}_{y_0} [\bar{F}_\sigma(\nu, X_{N,\nu}^*)^{2p}]^{\frac{1}{p}} \right), \end{aligned} \quad (\text{C.46})$$

for all $(t, p) \in [0, t^*] \times \mathbb{N}_{\geq 1}$. Next, using the definitions of $\phi_{\sigma_*}^r$ and ϕ_σ^r in (39), we obtain

$$\begin{aligned} \int_0^t e^{2\lambda \nu} [\phi_\sigma^r(\nu, Y_{N,\nu}) \quad \phi_{\sigma_*}^r(\nu, Y_{N,\nu})] d\widehat{W}_\nu \\ = \int_0^t e^{2\lambda \nu} [\partial \mathcal{V}(Y_{N,\nu})^\top g(\nu)^\perp F_\sigma^\perp(\nu, X_{N,\nu}^r) \quad -\partial \mathcal{V}(Y_{N,\nu})^\top \bar{F}_\sigma(\nu, X_{N,\nu}^*)] d\widehat{W}_\nu, \end{aligned}$$

where $\widehat{W}_\nu \in \mathbb{R}^{2d}$ is defined in Definition 5. Recall from the discussion after (C.42) that $[\phi_\sigma^r(\nu, Y_{N,\nu}) \quad \phi_{\sigma_*}^r(\nu, Y_{N,\nu})] \in \mathcal{M}_2^{loc}(\mathbb{R}^{1 \times 2d} | \mathfrak{W}_t \times \mathfrak{W}_t^*)$. Thus, we may use Lemma B.2 with $\{\theta, \xi, n_l, n_q, Q\} = \{2\lambda, 0, 1, 2d, \widehat{W}\}$, and

$$L(t) = [\partial \mathcal{V}(Y_{N,t})^\top g(t)^\perp F_\sigma^\perp(t, X_{N,t}^r) \quad -\partial \mathcal{V}(Y_{N,t})^\top \bar{F}_\sigma(\nu, X_{N,t}^*)],$$

to obtain

$$\left\| \int_0^t e^{2\lambda \nu} [\phi_\sigma^r(\nu, Y_{N,\nu}) \quad \phi_{\sigma_*}^r(\nu, Y_{N,\nu})] d\widehat{W}_\nu \right\|_{L_p}^{y_0}$$

$$\begin{aligned} &\leq \left(\mathbf{p} \frac{\mathbf{p}-1}{2} \right)^{\frac{1}{2}} \frac{e^{2\lambda t}}{2\sqrt{\lambda}} \sup_{\nu \in [0, t]} \left\| \begin{bmatrix} \partial \mathcal{V}(Y_{N, \nu})^\top g(t)^\perp F_\sigma^\perp(\nu, X_{N, \nu}^r) & -\partial \mathcal{V}(Y_{N, \nu})^\top \bar{F}_\sigma(\nu, X_{N, \nu}^*) \end{bmatrix} \right\|_{L_p}^{y_0} \\ &\leq \left(\mathbf{p} \frac{\mathbf{p}-1}{2} \right)^{\frac{1}{2}} \frac{e^{2\lambda t}}{2\sqrt{\lambda}} \sup_{\nu \in [0, t^*]} \left\| \begin{bmatrix} \partial \mathcal{V}(Y_{N, \nu})^\top g(t)^\perp F_\sigma^\perp(\nu, X_{N, \nu}^r) & -\partial \mathcal{V}(Y_{N, \nu})^\top \bar{F}_\sigma(\nu, X_{N, \nu}^*) \end{bmatrix} \right\|_{L_p}^{y_0}, \quad (\text{C.47}) \end{aligned}$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 2}$, where we have bounded $(e^{4\lambda t} - 1)^{\frac{1}{2}}$ by $e^{2\lambda t}$. Next, observe that

$$\begin{aligned} &\left\| \begin{bmatrix} \partial \mathcal{V}(Y_{N, t})^\top g(t)^\perp F_\sigma^\perp(t, X_{N, t}^r) & -\partial \mathcal{V}(Y_{N, t})^\top \bar{F}_\sigma(t, X_{N, t}^*) \end{bmatrix} \right\| \\ &= \left(\left\| \partial \mathcal{V}(Y_{N, t})^\top g(t)^\perp F_\sigma^\perp(t, X_{N, t}^r) \right\|^2 + \left\| \partial \mathcal{V}(Y_{N, t})^\top \bar{F}_\sigma(t, X_{N, t}^*) \right\|^2 \right)^{\frac{1}{2}} \\ &\leq \left(\left\| \partial \mathcal{V}(Y_{N, t})^\top g(t)^\perp F_\sigma^\perp(t, X_{N, t}^r) \right\| + \left\| \partial \mathcal{V}(Y_{N, t})^\top \bar{F}_\sigma(t, X_{N, t}^*) \right\| \right), \end{aligned}$$

and thus, an application of the Minkowski's inequality yields

$$\begin{aligned} &\left\| \begin{bmatrix} \partial \mathcal{V}(Y_{N, t})^\top g(t)^\perp F_\sigma^\perp(t, X_{N, t}^r) & -\partial \mathcal{V}(Y_{N, t})^\top \bar{F}_\sigma(t, X_{N, t}^*) \end{bmatrix} \right\|_{L_p}^{y_0} \\ &\leq \left(\left\| \partial \mathcal{V}(Y_{N, t})^\top g(t)^\perp F_\sigma^\perp(t, X_{N, t}^r) \right\|_{L_p}^{y_0} + \left\| \partial \mathcal{V}(Y_{N, t})^\top \bar{F}_\sigma(t, X_{N, t}^*) \right\|_{L_p}^{y_0} \right) \\ &\leq \left(\Delta_g^\perp \left\| \partial \mathcal{V}(Y_{N, t}) \right\|_{L_{2p}}^{y_0} \left\| F_\sigma^\perp(t, X_{N, t}^r) \right\|_{L_{2p}}^{y_0} + \left\| \partial \mathcal{V}(Y_{N, t}) \right\|_{L_{2p}}^{y_0} \left\| \bar{F}_\sigma(t, X_{N, t}^*) \right\|_{L_{2p}}^{y_0} \right), \end{aligned}$$

where we have used the bound on $g(t)^\perp$ from Assumption 1, and the Cauchy-Schwarz inequality to obtain the last bound. Substituting the above bound into (C.47) produces

$$\begin{aligned} &\left\| \int_0^t e^{2\lambda \nu} \begin{bmatrix} \phi_\sigma^r(\nu, Y_{N, \nu}) & \phi_{\sigma^*}^r(\nu, Y_{N, \nu}) \end{bmatrix} d\widehat{W}_\nu \right\|_{L_p}^{y_0} \\ &\leq \left(\mathbf{p} \frac{\mathbf{p}-1}{2} \right)^{\frac{1}{2}} \frac{e^{2\lambda t}}{2\sqrt{\lambda}} \sup_{\nu \in [0, t^*]} \left(\Delta_g^\perp \left\| \partial \mathcal{V}(Y_{N, \nu}) \right\|_{L_{2p}}^{y_0} \left\| F_\sigma^\perp(\nu, X_{N, \nu}^r) \right\|_{L_{2p}}^{y_0} \right. \\ &\quad \left. + \left\| \partial \mathcal{V}(Y_{N, \nu}) \right\|_{L_{2p}}^{y_0} \left\| \bar{F}_\sigma(\nu, X_{N, \nu}^*) \right\|_{L_{2p}}^{y_0} \right), \quad (\text{C.48}) \end{aligned}$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 2}$. Substituting (C.46) into (C.43) leads to (C.41a), and substituting (C.46) and (C.48) into (C.44) leads to (C.41b), thus completing the proof. \square

We next derive the bound on $\Xi_{\mathcal{U}}^r$ in the following lemma.

Lemma C.2 *If the stopping time τ^* , defined in (34), Lemma 4.1, satisfies $\tau^* = t^*$, then the term $\Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega)$ defined in (38), Lemma 4.1, satisfies the following bounds:*

$$\begin{aligned} &\mathbb{E}_{y_0} \left[e^{-(2\lambda + \omega)\tau(t)} \Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega) \right] \\ &\leq \frac{1}{|2\lambda - \omega|} (\Delta_{\mathcal{U}_1}^r(1, t^*) + \sqrt{\omega} \Delta_{\mathcal{U}_2}^r(1, t^*) + \omega \Delta_{\mathcal{U}_3}^r(1, t^*)), \quad \forall t \in [0, t^*], \quad (\text{C.49a}) \end{aligned}$$

$$\begin{aligned} &\left\| e^{-(2\lambda + \omega)\tau(t)} \Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega) \right\|_{L_p}^{y_0} \\ &\leq \frac{1}{|2\lambda - \omega|} (\Delta_{\mathcal{U}_1}^r(\mathbf{p}, t^*) + \sqrt{\omega} \Delta_{\mathcal{U}_2}^r(\mathbf{p}, t^*) + \omega \Delta_{\mathcal{U}_3}^r(\mathbf{p}, t^*)), \quad \forall (t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 2}, \quad (\text{C.49b}) \end{aligned}$$

where

$$\begin{aligned} \Delta_{\mathcal{U}_1}^r(\mathbf{p}, t^*) &= \sqrt{\lambda} \Delta_g \mathbf{p}(\mathbf{p}) \sup_{\nu \in [0, t^*]} \left\| \partial \mathcal{V}(Y_{N, \nu}) \right\|_{L_{2p}} \sup_{\nu \in [0, t^*]} \left\| F_\sigma^\parallel(\nu, X_{N, \nu}^r) \right\|_{L_{2p}} \\ &\quad + \left(\frac{\Delta_{\mathcal{P}_1}^r(\mathbf{p}, t^*)}{\sqrt{\lambda}} + 2\Delta_g \sup_{\nu \in [0, t^*]} \left\| \partial \mathcal{V}(Y_{N, \nu}) \right\|_{L_{2p}}^{y_0} + \frac{\Delta_g^2}{\lambda} \sup_{\nu \in [0, t^*]} \left\| \Lambda_\mu^\parallel(\nu, X_{N, \nu}^r) \right\|_{L_{2p}}^{y_0} \right) \\ &\quad \times \sup_{\nu \in [0, t^*]} \left\| \Lambda_\mu^\parallel(\nu, X_{N, \nu}^r) \right\|_{L_{2p}}^{y_0}, \end{aligned}$$

$$\Delta_{\mathcal{U}_2}^r(\mathbf{p}, t^*) = \left(\frac{\Delta_{\mathcal{P}_2}^r(\mathbf{p}, t^*)}{\sqrt{\lambda}} + \Delta_g \mathbf{p}'(\mathbf{p}) \left[\sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Y_{N, \nu})\|_{L_{2p}}^{y_0} + \frac{2\Delta_g}{\lambda} \sup_{\nu \in [0, t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N, \nu}^r)\|_{L_{2p}}^{y_0} \right] \right) \times \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N, \nu}^r)\|_{L_{2p}}^{y_0},$$

$$\Delta_{\mathcal{U}_3}^r(\mathbf{p}, t^*) = \left(\frac{\Delta_{\mathcal{P}_3}^r(\mathbf{p}, t^*)}{\lambda} + \frac{\Delta_g^2 \mathbf{p}'(\mathbf{p})^2}{\lambda} \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N, \nu}^r)\|_{L_{2p}}^{y_0} \right) \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N, \nu}^r)\|_{L_{2p}}^{y_0},$$

and where the constants $\Delta_{\mathcal{P}_1}^r(\mathbf{p}, t^*)$, $\Delta_{\mathcal{P}_2}^r(\mathbf{p}, t^*)$ and $\Delta_{\mathcal{P}_3}^r(\mathbf{p}, t^*)$ are defined in the statement of Proposition C.5. Furthermore, $\mathbf{p}(\mathbf{p})$ and $\mathbf{p}'(\mathbf{p})$ are defined in (A.1a).

Proof. We begin by recalling the definition of $\psi^r(\tau(t), \nu, Y_N)$ from (40)

$$\psi^r(\tau(t), \nu, Y_N) = \frac{\omega}{2\lambda - \omega} \left(e^{\omega(\tau(t)+\nu)} \mathcal{P}^r(\tau(t), \nu) - e^{(2\lambda\tau(t)+\omega\nu)} \partial \mathcal{V}(Y_{N, \tau(t)})^\top g(\tau(t)) \right) + \frac{2\lambda}{2\lambda - \omega} e^{(\omega\tau(t)+2\lambda\nu)} \partial \mathcal{V}(Y_{N, \nu})^\top g(\nu),$$

which, upon using the decomposition (C.14) in Proposition C.3, can be re-written as

$$\psi^r(\tau(t), \nu, Y_N) = \frac{\omega}{2\lambda - \omega} \left(e^{\omega(\tau(t)+\nu)} \mathcal{P}_o^r(\tau(t), \nu) - e^{(2\lambda\tau(t)+\omega\nu)} \partial \mathcal{V}(Y_{N, \tau(t)})^\top g(\tau(t)) \right) + \frac{\omega}{2\lambda - \omega} e^{\omega(\tau(t)+\nu)} \mathcal{P}_{ad}^r(\tau(t), \nu) + \frac{2\lambda}{2\lambda - \omega} e^{(\omega\tau(t)+2\lambda\nu)} \partial \mathcal{V}(Y_{N, \nu})^\top g(\nu),$$

We thus decompose $\psi^r(\tau(t), \nu, Y_N)$ in (40) as follows:

$$\psi^r(\tau(t), \nu, Y_N) = \sum_{i=1}^3 \psi_i^r(\tau(t), \nu, Y_N) + \psi_{ad}^r(\tau(t), \nu, Y_N) \in \mathbb{R}^{1 \times m}, \quad (\text{C.50})$$

where

$$\begin{aligned} \psi_1^r(\tau(t), \nu, Y_N) &= \frac{\omega}{2\lambda - \omega} e^{\omega(\tau(t)+\nu)} \mathcal{P}_o^r(\tau(t), \nu), \\ \psi_2^r(\tau(t), \nu, Y_N) &= -\frac{\omega}{2\lambda - \omega} e^{(2\lambda\tau(t)+\omega\nu)} \partial \mathcal{V}(Y_{N, \tau(t)})^\top g(\tau(t)), \\ \psi_3^r(\tau(t), \nu, Y_N) &= \frac{2\lambda}{2\lambda - \omega} e^{(\omega\tau(t)+2\lambda\nu)} \partial \mathcal{V}(Y_{N, \nu})^\top g(\nu), \end{aligned}$$

and

$$\psi_{ad}^r(\tau(t), \nu, Y_N) = \frac{\omega}{2\lambda - \omega} e^{\omega(\tau(t)+\nu)} \mathcal{P}_{ad}^r(\tau(t), \nu).$$

Next, using the definitions of $\mathcal{U}_\mu^r(\tau(t), \nu, Y_N; \omega)$ and $\mathcal{U}_\sigma^r(\tau(t), \nu, Y_N; \omega)$ in (39), the decomposition in (C.50) produces the following expression:

$$\begin{aligned} \mathcal{U}_\mu^r(\tau(t), \nu, Y_N; \omega) &= \left(\sum_{i=1}^3 \psi_i^r(\tau(t), \nu, Y_N) + \psi_{ad}^r(\tau(t), \nu, Y_N) \right) \Lambda_\mu^\parallel(\nu, X_{N, \nu}^r) \in \mathbb{R}, \\ \mathcal{U}_\sigma^r(\tau(t), \nu, Y_N; \omega) &= \left(\sum_{i=1}^3 \psi_i^r(\tau(t), \nu, Y_N) + \psi_{ad}^r(\tau(t), \nu, Y_N) \right) F_\sigma^\parallel(\nu, X_{N, \nu}^r) \in \mathbb{R}^{1 \times d}. \end{aligned}$$

Then, we may re-write $\Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega)$ in (38) as

$$\begin{aligned} \Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega) &= \int_0^t (\mathcal{U}_\mu^r(t, \nu, Y_N; \omega) d\nu + \mathcal{U}_\sigma^r(t, \nu, Y_N; \omega) dW_\nu) \\ &= \sum_{i=1}^3 \int_0^t \psi_i^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N, \nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N, \nu}^r) dW_\nu] \end{aligned}$$

$$+ \int_0^t \psi_{ad}^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu], \quad t \in [0, t^*],$$

where we have used $\tau(t) \in [0, t^*]$, $\forall t \in \mathbb{R}_{\geq 0}$, since $\tau^* = t^*$. Hence, exploiting the linearity of the expectation operator, along with Minkowski's inequality, we obtain

$$\begin{aligned} & \mathbb{E}_{y_0} \left[e^{-(2\lambda + \omega)\tau(t)} \Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega) \right] \\ &= e^{-(2\lambda + \omega)t} \sum_{i=1}^3 \mathbb{E}_{y_0} \left[\int_0^t \psi_i^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \right] \\ & \quad + e^{-(2\lambda + \omega)t} \mathbb{E}_{y_0} \left[\int_0^t \psi_{ad}^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \right], \quad \forall t \in [0, t^*], \quad (\text{C.51}) \end{aligned}$$

and

$$\begin{aligned} & \left\| e^{-(2\lambda + \omega)\tau(t)} \Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega) \right\|_{L_p}^{y_0} \\ & \leq e^{-(2\lambda + \omega)t} \sum_{i=1}^3 \left\| \int_0^t \psi_i^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \right\|_{L_p}^{y_0} \\ & \quad + e^{-(2\lambda + \omega)t} \left\| \int_0^t \psi_{ad}^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \right\|_{L_p}^{y_0}, \quad (\text{C.52}) \end{aligned}$$

for all $(t, p) \in [0, t^*] \times \mathbb{N}_{\geq 2}$. In the two expressions above, we have used the fact that since $\tau^* = t^*$, the definition of $\tau(t)$ in (34) leads to the conclusion that $\tau(t) = t$, for $t \in [0, t^*]$. Hence, the stopping time $\tau(t)$ is a deterministic function of t since t^* is a constant, and thus

$$\mathbb{E} \left[e^{-(2\lambda + \omega)\tau(t)} \right] = e^{-(2\lambda + \omega)\tau(t)} = e^{-(2\lambda + \omega)t}, \quad \forall t \in [0, t^*].$$

Next, using the definition of ψ_1^r in (C.50) produces

$$\begin{aligned} & \int_0^t \psi_1^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \\ &= \frac{\omega e^{\omega t}}{2\lambda - \omega} \int_0^t e^{\omega \nu} \mathcal{P}_\circ^r(t, \nu) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu], \quad t \in [0, t^*]. \quad (\text{C.53}) \end{aligned}$$

Using the definition of \mathcal{P}_\circ^r in (C.15a), one sees that

$$\int_0^t e^{\omega \nu} \mathcal{P}_\circ^r(t, \nu) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] = \sum_{i=1}^4 N_i^r(t), \quad \forall t \in [0, T],$$

where the processes $N_i^r(t)$, $i \in \{1, \dots, 4\}$, are defined in (C.21) in the statement of Proposition C.5 as follows:

$$\begin{aligned} N_1^r(t) &= \int_0^t e^{\omega \nu} M_\mu(t, \nu) \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu, & N_2^r(t) &= \int_0^t e^{\omega \nu} M_\mu(t, \nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu, \\ N_3^r(t) &= \int_0^t e^{\omega \nu} M_\sigma(t, \nu) \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu, & N_4^r(t) &= \int_0^t e^{\omega \nu} M_\sigma(t, \nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu, \end{aligned}$$

and

$$M_\mu(t, \nu) = \int_\nu^t e^{(2\lambda - \omega)\beta} \mathcal{P}_\mu^r(\beta)^\top d\beta, \quad M_\sigma(t, \nu) = \int_\nu^t e^{(2\lambda - \omega)\beta} [\mathcal{P}_\sigma^r(\beta) dW_\beta + \mathcal{P}_{\sigma_*}^r(\beta) dW_\beta^*]^\top.$$

Hence, we may write (C.53) as

$$\int_0^t \psi_1^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] = \frac{\omega e^{\omega t}}{2\lambda - \omega} \sum_{i=1}^4 N_i^r(t), \quad t \in [0, t^*].$$

It then follows from the linearity of the expectation operator and the Minkowski's inequality that

$$\mathbb{E}_{y_0} \left[\int_0^t \psi_1^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \right] \leq \frac{\omega e^{\omega t}}{|2\lambda - \omega|} \sum_{i=1}^4 |\mathbb{E}_{y_0} [N_i^r(t)]|, \forall t \in [0, t^*],$$

and

$$\begin{aligned} \left\| \int_0^t \psi_1^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \right\|_{L_p}^{y_0} \\ \leq \frac{\omega e^{\omega t}}{|2\lambda - \omega|} \sum_{i=1}^4 \|N_i^r(t)\|_{L_p}^{y_0}, \quad \forall (t, p) \in [0, t^*] \times \mathbb{N}_{\geq 2}. \end{aligned}$$

Hence, using Proposition C.5 leads to

$$\begin{aligned} \mathbb{E}_{y_0} \left[\int_0^t \psi_1^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \right] \\ \leq \frac{e^{(2\lambda + \omega)t}}{\sqrt{\lambda} |2\lambda - \omega|} \Delta_{\mathcal{P}_1}^r(1, t^*) \sup_{\nu \in [0, t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)\|_{L_2}^{y_0} \\ + \frac{e^{(2\lambda + \omega)t}}{|2\lambda - \omega|} \left(\frac{\sqrt{\omega}}{\sqrt{\lambda}} \Delta_{\mathcal{P}_2}^r(1, t^*) + \frac{\omega}{\lambda} \Delta_{\mathcal{P}_3}^r(1, t^*) \right) \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu}^r)\|_{L_2}^{y_0}, \quad (\text{C.54}) \end{aligned}$$

for all $t \in [0, t^*]$ and

$$\begin{aligned} \left\| \int_0^t \psi_1^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \right\|_{L_p}^{y_0} \\ \leq \frac{e^{(2\lambda + \omega)t}}{\sqrt{\lambda} |2\lambda - \omega|} \Delta_{\mathcal{P}_1}^r(p, t^*) \sup_{\nu \in [0, t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} \\ + \frac{e^{(2\lambda + \omega)t}}{|2\lambda - \omega|} \left(\frac{\sqrt{\omega}}{\sqrt{\lambda}} \Delta_{\mathcal{P}_2}^r(p, t^*) + \frac{\omega}{\lambda} \Delta_{\mathcal{P}_3}^r(p, t^*) \right) \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0}, \quad (\text{C.55}) \end{aligned}$$

for all $(t, p) \in [0, t^*] \times \mathbb{N}_{\geq 2}$.

Next, using the definition of ψ_2^r in (C.50) we see that

$$\begin{aligned} \int_0^t \psi_2^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \\ = -\frac{\omega e^{2\lambda t}}{2\lambda - \omega} \partial \mathcal{V}(Y_{N,t})^\top g(t) \int_0^t e^{\omega \nu} [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu], \quad t \in [0, t^*]. \end{aligned}$$

Hence

$$\begin{aligned} \left| \int_0^t \psi_2^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \right| \\ \leq \frac{\omega e^{2\lambda t}}{|2\lambda - \omega|} \Delta_g \|\partial \mathcal{V}(Y_{N,t})\| \left\| \int_0^t e^{\omega \nu} [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \right\|, \quad t \in [0, t^*], \end{aligned}$$

where we have used the bound on $g(t)$ from Assumption 1. It then follows from the Cauchy-Schwarz inequality and the Minkowski's inequality that

$$\begin{aligned} \left\| \int_0^t \psi_2^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \right\|_{L_p}^{y_0} \\ \leq \frac{\omega e^{2\lambda t}}{|2\lambda - \omega|} \Delta_g \|\partial \mathcal{V}(Y_{N,t})\|_{L_{2p}}^{y_0} \left\| \int_0^t e^{\omega \nu} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)\| d\nu \right\|_{L_{2p}}^{y_0} \end{aligned}$$

$$+ \frac{\omega e^{2\lambda t}}{|2\lambda - \omega|} \Delta_g \|\partial \mathcal{V}(Y_{N,t})\|_{L_{2p}}^{y_0} \left\| \int_0^t e^{\omega \nu} F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu \right\|_{L_{2p}}^{y_0}, \forall (t, p) \in [0, t^*] \times \mathbb{N}_{\geq 1}.$$

Since the regularity assumptions, along with the t -continuity of the strong solution $X_{N,t}^r$, implies that $\Lambda_\mu^\parallel(\cdot, X_N^r) \in \mathcal{M}_2^{loc}(\mathbb{R}^m | \mathfrak{Y}_t)$ and $F_\sigma^\parallel(\cdot, X_N^r) \in \mathcal{M}_2^{loc}(\mathbb{R}^{m \times d} | \mathfrak{Y}_t)$. Hence, we may invoke Proposition B.1 for $\{\theta, S(t), \xi\} = \{\omega, \Lambda_\mu^\parallel(\cdot, X_N^r)(n_s = m), 0\}$, and Lemma B.2 for $\{\theta, Q_t, L(t), \xi\} = \{\omega, W_t(n_q = d), F_\sigma^\parallel(\cdot, X_N^r)(n_l = m), 0\}$, and obtain

$$\begin{aligned} & \left\| \int_0^t \psi_2^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \right\|_{L_p}^{y_0} \\ & \leq \frac{e^{(2\lambda + \omega)t}}{|2\lambda - \omega|} \Delta_g \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Y_{N,\nu})\|_{L_{2p}}^{y_0} \sup_{\nu \in [0, t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} \\ & \quad + \frac{\sqrt{\omega} e^{(2\lambda + \omega)t}}{|2\lambda - \omega|} \Delta_g \left(p \frac{2p-1}{2} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Y_{N,\nu})\|_{L_{2p}}^{y_0} \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2p}}, \quad (\text{C.56}) \end{aligned}$$

for all $(t, p) \in [0, t^*] \times \mathbb{N}_{\geq 1}$, where we have bounded $e^{\omega t} - 1$ and $e^{2\omega t} - 1$ by $e^{\omega t}$ and $e^{2\omega t}$, respectively, and used the fact that $\|\partial \mathcal{V}(Y_{N,t})\|_{L_{2p}}^{y_0} \leq \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Y_{N,\nu})\|_{L_{2p}}^{y_0}$, $\forall t \in [0, t^*]$.

Next, using the definition of ψ_3^r in (C.50) we see that

$$\begin{aligned} & \int_0^t \psi_3^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \\ & = \frac{2\lambda e^{\omega t}}{2\lambda - \omega} \int_0^t e^{2\lambda \nu} \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu], \quad t \in [0, t^*]. \end{aligned}$$

It then follows from the linearity of the expectation operator and the Minkowski's inequality that

$$\begin{aligned} & \mathbb{E}_{y_0} \left[\int_0^t \psi_3^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \right] \\ & \leq \frac{2\lambda e^{\omega t}}{|2\lambda - \omega|} \mathbb{E}_{y_0} \left[\int_0^t e^{2\lambda \nu} \left| \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) \right| d\nu \right] \\ & \quad + \frac{2\lambda e^{\omega t}}{|2\lambda - \omega|} \left| \mathbb{E}_{y_0} \left[\int_0^t e^{2\lambda \nu} \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu \right] \right|, \quad \forall t \in [0, t^*], \quad (\text{C.57}) \end{aligned}$$

and

$$\begin{aligned} & \left\| \int_0^t \psi_3^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \right\|_{L_p}^{y_0} \\ & \leq \frac{2\lambda e^{\omega t}}{|2\lambda - \omega|} \left\| \int_0^t e^{2\lambda \nu} \left| \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) \right| d\nu \right\|_{L_p}^{y_0} \\ & \quad + \frac{2\lambda e^{\omega t}}{|2\lambda - \omega|} \left\| \int_0^t e^{2\lambda \nu} \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu \right\|_{L_p}^{y_0}, \quad \forall (t, p) \in [0, t^*] \times \mathbb{N}_{\geq 2}. \quad (\text{C.58}) \end{aligned}$$

Now, as before, Proposition B.1 for $\{\theta, S(t), \xi\} = \{2\lambda, \partial \mathcal{V}(Y_{N,t})^\top g(t) \Lambda_\mu^\parallel(t, X_{N,t}^r)(n_s = 1), 0\}$ yields

$$\begin{aligned} & \left\| \int_0^t e^{2\lambda \nu} \left| \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) \right| d\nu \right\|_{L_p}^{y_0} \\ & \leq \frac{e^{2\lambda t}}{2\lambda} \sup_{\nu \in [0, t^*]} \left\| \partial \mathcal{V}(Y_{N,\nu})^\top g(\nu) \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) \right\|_{L_p}, \quad \forall (t, p) \in [0, t^*] \times \mathbb{N}_{\geq 1}, \end{aligned}$$

where we have bounded $e^{2\lambda t} - 1$ by $e^{2\lambda t}$. It then follows from the Cauchy-Schwarz inequality and the bound on $g(t)$ from Assumption 1 that

$$\begin{aligned} & \left\| \int_0^t e^{2\lambda\nu} \left| \partial\mathcal{V}(Y_{N,\nu})^\top g(\nu) \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) \right| d\nu \right\|_{L_p}^{y_0} \\ & \leq \frac{e^{2\lambda t}}{2\lambda} \Delta_g \sup_{\nu \in [0, t^*]} \|\partial\mathcal{V}(Y_{N,\nu})\|_{L_{2p}} \sup_{\nu \in [0, t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2p}}, \quad \forall (t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 1}. \end{aligned} \quad (\text{C.59})$$

Next, we may write

$$\int_0^t e^{2\lambda\nu} \partial\mathcal{V}(Y_{N,\nu})^\top g(\nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu = \int_0^t e^{2\lambda\nu} \partial\mathcal{V}(Y_{N,\nu})^\top g(\nu) \begin{bmatrix} F_\sigma^\parallel(\nu, X_{N,\nu}^r) & 0_{m,d} \end{bmatrix} d\widehat{W}_\nu.$$

Furthermore, the regularity assumptions and the t -continuity of the strong solutions $X_{N,t}^r$ and $X_{N,t}^*$ imply that $(\partial\mathcal{V}(Y_{N,t})^\top g(t) \begin{bmatrix} F_\sigma^\parallel(t, X_{N,t}^r) & 0_{m,d} \end{bmatrix}) \in \mathcal{M}_2^{loc}(\mathbb{R}^{1 \times 2d} | \mathfrak{W}_t \times \mathfrak{W}_t^*)$. Hence, by [91, Thm. 1.5.21]

$$\mathbb{E}_{y_0} \left[\int_0^t e^{2\lambda\nu} \partial\mathcal{V}(Y_{N,\nu})^\top g(\nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu \right] = 0, \quad \forall t \in [0, t^*]. \quad (\text{C.60})$$

Moreover, Lemma B.2 for $\{\theta, Q_t, L(t), \xi\} = \{2\lambda, \widehat{W}_t(n_d = 2d), (\partial\mathcal{V}(Y_{N,t})^\top g(t) \begin{bmatrix} F_\sigma^\parallel(t, X_{N,t}^r) & 0_{m,d} \end{bmatrix}) (n_l = 1), 0\}$ produces

$$\begin{aligned} & \left\| \int_0^t e^{2\lambda\nu} \partial\mathcal{V}(Y_{N,\nu})^\top g(\nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu \right\|_{L_p}^{y_0} \\ & \leq \frac{e^{2\lambda t}}{2\sqrt{\lambda}} \left(\mathbf{p} \frac{\mathbf{p}-1}{2} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t^*]} \left\| \partial\mathcal{V}(Y_{N,\nu})^\top g(\nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r) \right\|_{L_p}, \quad \forall (t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 2}, \end{aligned}$$

where we have bounded $(e^{4\lambda t} - 1)^{\frac{1}{2}}$ by $e^{2\lambda t}$. It then follows from the Cauchy-Schwarz inequality and the bound on $g(t)$ from Assumption 1 that

$$\begin{aligned} & \left\| \int_0^t e^{2\lambda\nu} \partial\mathcal{V}(Y_{N,\nu})^\top g(\nu) F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu \right\|_{L_p}^{y_0} \\ & \leq \frac{e^{2\lambda t}}{2\sqrt{\lambda}} \Delta_g \left(\mathbf{p} \frac{\mathbf{p}-1}{2} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t^*]} \|\partial\mathcal{V}(Y_{N,\nu})\|_{L_{2p}} \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2p}}, \quad \forall (t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 2}. \end{aligned} \quad (\text{C.61})$$

Substituting (C.59) and (C.60) into (C.57) yields

$$\begin{aligned} & \mathbb{E}_{y_0} \left[\int_0^t \psi_3^r(t, \nu, Y_N) \left[\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu \right] \right] \\ & \leq \frac{e^{(2\lambda + \omega)t}}{|2\lambda - \omega|} \Delta_g \sup_{\nu \in [0, t^*]} \|\partial\mathcal{V}(Y_{N,\nu})\|_{L_2}^{y_0} \sup_{\nu \in [0, t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)\|_{L_2}^{y_0} \quad \forall t \in [0, t^*]. \end{aligned} \quad (\text{C.62})$$

Similarly, substituting (C.59) and (C.61) into (C.58) yields

$$\begin{aligned} & \left\| \int_0^t \psi_3^r(t, \nu, Y_N) \left[\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu \right] \right\|_{L_p}^{y_0} \\ & \leq \frac{e^{(2\lambda + \omega)t}}{|2\lambda - \omega|} \Delta_g \sup_{\nu \in [0, t^*]} \|\partial\mathcal{V}(Y_{N,\nu})\|_{L_{2p}}^{y_0} \sup_{\nu \in [0, t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} \\ & \quad + \frac{\sqrt{\lambda} e^{(2\lambda + \omega)t}}{|2\lambda - \omega|} \Delta_g \left(\mathbf{p} \frac{\mathbf{p}-1}{2} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t^*]} \|\partial\mathcal{V}(Y_{N,\nu})\|_{L_{2p}}^{y_0} \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0}, \end{aligned} \quad (\text{C.63})$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 2}$.

Finally, using the definition of ψ_{ad}^r in (C.50) we see that

$$\int_0^t \psi_{ad}^r(t, \nu, Y_N) \left[\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu \right]$$

$$= \frac{\omega e^{\omega t}}{2\lambda - \omega} \int_0^t e^{\omega \nu} \mathcal{P}_{ad}^r(t, \nu) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu], \quad t \in [0, t^*].$$

Using the definition of \mathcal{P}_{ad}^r in (C.15b), Proposition C.3, then leads to

$$\begin{aligned} & \int_0^t \psi_{ad}^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \\ &= \frac{\omega e^{\omega t}}{2\lambda - \omega} \int_0^t e^{\omega \nu} \left(\int_\nu^t e^{(2\lambda - \omega)\beta} \mathcal{P}_{\mathcal{U}}^r(\beta)^\top d\beta \right) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \\ &= \frac{\omega e^{\omega t}}{2\lambda - \omega} N_{\mathcal{U}}^r(t), \quad t \in [0, t^*], \end{aligned}$$

where the process $N_{\mathcal{U}}^r$ is defined in (C.34), Proposition C.6. Thus, we invoke Proposition C.6 and obtain

$$\begin{aligned} & \left\| \int_0^t \psi_{ad}^r(t, \nu, Y_N) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r) d\nu + F_\sigma^\parallel(\nu, X_{N,\nu}^r) dW_\nu] \right\|_{L_p}^{y_0} \\ & \leq \frac{e^{(2\lambda + \omega)t}}{2\lambda |2\lambda - \omega|} \Delta_g^2 \left(\sup_{\nu \in [0, t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} + \sqrt{\omega} \left(p \frac{2p-1}{2} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} \right)^2, \quad (\text{C.64}) \end{aligned}$$

for all $(t, p) \in [0, t^*] \times \mathbb{N}_{\geq 1}$.

Substituting (C.54), (C.56), (C.62), and (C.64) into (C.51), followed by grouping terms that are $\propto \{\omega, \sqrt{\omega}, 1\} / |2\lambda - \omega|$ leads to (C.49a). Similarly, substituting (C.55), (C.56), (C.63), and (C.64) into (C.52) leads to (C.49b), thus concluding the proof. \square

Next, we derive the expressions for the terms that constitute the bound in Proposition C.5.

Proposition C.7 Consider the stopping times τ^* and t^* , as defined in (34), Lemma 4.1, and assume that $\tau^* = t^*$. If Assumptions 1 - 5 hold, then with p^* defined in Assumption 3, and

$$\tilde{\mathcal{E}}_{r,y_0}(2p) \doteq \sup_{\nu \in [0, t^*]} \|X_{N,\nu}^r - X_{N,\nu}^*\|_{L_{2p}}^{y_0}, \quad \mathcal{E}_{r,y_0}(2p) \doteq \sup_{\nu \in [0, t^*]} \|X_{N,\nu}^r\|_{L_{2p}}^{y_0}, \quad p \in \{1, \dots, p^*\}, \quad (\text{C.65})$$

the following bounds hold for all $p \in \{1, \dots, p^*\}$:

$$\Delta_{\mathcal{P}_1}^r(p, t^*) \leq \hat{\Delta}_1^r(p) + \hat{\Delta}_2^r(p) \sqrt{\mathcal{E}_{r,y_0}(2p)} + \hat{\Delta}_3^r \mathcal{E}_{r,y_0}(2p) + \hat{\Delta}_4^r \tilde{\mathcal{E}}_{r,y_0}(2p), \quad (\text{C.66a})$$

$$\Delta_{\mathcal{P}_3}^r(p, t^*) \leq \sqrt{m} \Delta_g (2\Delta_p + \Delta_\sigma) + \sqrt{m} \Delta_g \Delta_\sigma \sqrt{\mathcal{E}_{r,y_0}(2p)}, \quad (\text{C.66b})$$

where $\Delta_{\mathcal{P}_1}^r(p, t^*)$ and $\Delta_{\mathcal{P}_3}^r(p, t^*)$ are defined in the statement of Proposition C.5. Furthermore, the constants $p(p)$ and $\mathbb{1}_{Lip}(f)$ are defined in (A.1), and the constants $\hat{\Delta}_1^r(p)$, $\hat{\Delta}_2^r(p)$, $\hat{\Delta}_3^r$, and $\hat{\Delta}_4^r$ are defined in (A.2).

Proof. We begin with the term $\Delta_{\mathcal{P}_\mu}^r$ defined in Proposition C.4 and the definition $\bar{F}_\mu = f$ to get

$$\Delta_{\mathcal{P}_\mu}^r(p) = \sup_{\nu \in [0, t^*]} \left(\|f(\nu, X_{N,\nu}^r) - f(\nu, X_{N,\nu}^*)\|_{L_{2p}}^{y_0} + \|\Lambda_\mu(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} \right) \quad (\text{C.67a})$$

$$\leq \sup_{\nu \in [0, t^*]} \left(\|f(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} + \|f(\nu, X_{N,\nu}^*)\|_{L_{2p}}^{y_0} + \|\Lambda_\mu(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} \right), \quad (\text{C.67b})$$

where the second inequality is due to the Minkowski inequality. Using Assumptions 1, 3, and 4, the bound in (C.67b) can be developed into

$$\Delta_{\mathcal{P}_\mu}^r(p) \leq \Delta_f (2 + \Delta_*) + \Delta_\mu + (\Delta_f + \Delta_\mu) \mathcal{E}_{r,y_0}(2p), \quad (\text{C.68})$$

where we have also used the subadditivity of the square root function. If additionally Assumption 8 holds, then we obtain a tighter bound by developing (C.67a) to get

$$\Delta_{\mathcal{P}_\mu}^r(p) \leq \Delta_\mu + L_f \tilde{\mathcal{E}}_{r,y_0}(2p) + \Delta_\mu \mathcal{E}_{r,y_0}(2p). \quad (\text{C.69})$$

We may combine the bounds in (C.68) and (C.69) by writing

$$\begin{aligned} \Delta_{\mathcal{P}_\mu}^r(\mathbf{p}) &\leq \Delta_f (2 + \Delta_\star) (1 - \mathbb{1}_{\text{Lip}}(f)) + \Delta_\mu \\ &\quad + (\Delta_f (1 - \mathbb{1}_{\text{Lip}}(f)) + \Delta_\mu) \mathcal{E}_{r,y_0}(2\mathbf{p}) + L_f \mathbb{1}_{\text{Lip}}(f) \tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p}), \end{aligned} \quad (\text{C.70})$$

where $\mathbb{1}_{\text{Lip}}(f)$ is defined in (A.1b).

Next, we have the following from Proposition C.4:

$$\Delta_{\mathcal{P}_\sigma}^r(\mathbf{p}) = \sup_{\nu \in [0, t^\star]} \left(\|p(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} + \|\Lambda_\sigma(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} \right) + \sup_{\nu \in [0, t^\star]} \|p(\nu, X_{N,\nu}^\star)\|_{L_{2p}}^{y_0},$$

where we have used the definition $\bar{F}_\sigma = p$. It then follows from Assumptions 1, 3, and 4 that

$$\Delta_{\mathcal{P}_\sigma}^r(\mathbf{p}) \leq 2\Delta_p + \Delta_\sigma + \Delta_\sigma \sqrt{\mathcal{E}_{r,y_0}(2\mathbf{p})}, \quad (\text{C.71})$$

where we have also used the following due to Cauchy-Schwarz inequality:

$$\left\| \|X_{N,\nu}^r\|_{L_{2p}}^{\frac{1}{2}} \right\|_{L_{2p}}^{y_0} = \left(\|X_{N,\nu}^r\|_{L_p}^{y_0} \right)^{\frac{1}{2}} \leq \left(\|X_{N,\nu}^r\|_{L_{2p}}^{y_0} \right)^{\frac{1}{2}},$$

and thus

$$\sup_{\nu \in [0, t^\star]} \left\| \|X_{N,\nu}^r\|_{L_{2p}}^{\frac{1}{2}} \right\|_{L_{2p}}^{y_0} \leq \sup_{\nu \in [0, t^\star]} \left(\|X_{N,\nu}^r\|_{L_{2p}}^{y_0} \right)^{\frac{1}{2}} = \left(\sup_{\nu \in [0, t^\star]} \|X_{N,\nu}^r\|_{L_{2p}}^{y_0} \right)^{\frac{1}{2}} = \sqrt{\mathcal{E}_{r,y_0}(2\mathbf{p})}. \quad (\text{C.72})$$

Using (C.70) and (C.71), we can now bound the terms in (C.20) as follows:

$$\begin{aligned} \sup_{\nu \in [0, t^\star]} \|\mathcal{P}_\mu^r(\nu)\|_{L_{2p}}^{y_0} &\leq 2\Delta_g [\Delta_f (2 + \Delta_\star) (1 - \mathbb{1}_{\text{Lip}}(f)) + \Delta_\mu] \\ &\quad + 2\Delta_g (\Delta_f (1 - \mathbb{1}_{\text{Lip}}(f)) + \Delta_\mu) \mathcal{E}_{r,y_0}(2\mathbf{p}) + 2(\Delta_g L_f \mathbb{1}_{\text{Lip}}(f) + \Delta_g) \tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p}), \end{aligned} \quad (\text{C.73a})$$

$$\sum_{i \in \{\sigma, \sigma_\star\}} \sup_{\nu \in [0, t^\star]} \|\mathcal{P}_i^r(\nu)\|_{L_{2p}}^{y_0} \leq 2\Delta_g (2\Delta_p + \Delta_\sigma) + 2\Delta_g \Delta_\sigma \sqrt{\mathcal{E}_{r,y_0}(2\mathbf{p})}, \quad (\text{C.73b})$$

where we have used the definition $\partial \mathcal{V}(Y_N) = 2(X_N^r - X_N^\star)$ from the statement of Lemma 4.1. Now, recall the following definitions of $\Delta_{\mathcal{P}_1}^r(\mathbf{p}, t^\star)$ and $\Delta_{\mathcal{P}_3}^r(\mathbf{p}, t^\star)$ from the statement of Proposition C.5:

$$\begin{aligned} \Delta_{\mathcal{P}_1}^r(\mathbf{p}, t^\star) &= \frac{1}{2\sqrt{\lambda}} \sup_{\nu \in [0, t^\star]} \|\mathcal{P}_\mu^r(\nu)\|_{L_{2p}}^{y_0} + \frac{\mathbf{p}(\mathbf{p})}{2} \sup_{\nu \in [0, t^\star]} \left(\|\mathcal{P}_\sigma^r(\nu)\|_{L_{2p}}^{y_0} + \|\mathcal{P}_{\sigma_\star}^r(\nu)\|_{L_{2p}}^{y_0} \right), \\ \Delta_{\mathcal{P}_3}^r(\mathbf{p}, t^\star) &= \frac{\sqrt{m}}{2} \sup_{\nu \in [0, t^\star]} \left(\|\mathcal{P}_\sigma^r(\nu)\|_{L_{2p}}^{y_0} + \|\mathcal{P}_{\sigma_\star}^r(\nu)\|_{L_{2p}}^{y_0} \right). \end{aligned}$$

Substituting (C.73) into the above then leads to (C.66). □

The final result of the section derives the bounds due to Lemmas C.1 and C.2.

Lemma C.3 Consider the stopping times τ^\star and t^\star , as defined in (34), Lemma 4.1, and assume that $\tau^\star = t^\star$. If Assumptions 1 - 5 hold, then with \mathbf{p}^\star defined in Assumption 3, and

$$\tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p}) \doteq \sup_{\nu \in [0, t^\star]} \|X_{N,\nu}^r - X_{N,\nu}^\star\|_{L_{2p}}^{y_0}, \quad \mathcal{E}_{r,y_0}(2\mathbf{p}) \doteq \sup_{\nu \in [0, t^\star]} \|X_{N,\nu}^r\|_{L_{2p}}^{y_0}, \quad \mathbf{p} \in \{1, \dots, \mathbf{p}^\star\}, \quad (\text{C.74})$$

the following bound holds for all $(t, \mathbf{p}) \in \mathbb{R}_{\geq 0} \times \{2, \dots, \mathbf{p}^\star\}$:

$$\begin{aligned} &\left\| e^{-2\lambda\tau(t)} \Xi^r(\tau(t), Y_N) \right\|_{L_p}^{y_0} + \left\| e^{-(2\lambda + \omega)\tau(t)} \Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega) \right\|_{L_p}^{y_0} \\ &\leq \Delta_{\circ}^r(\mathbf{p}, \omega) + \tilde{\Delta}_{\circ}^r(\mathbf{p}, \omega) (\mathcal{E}_{r,y_0}(2\mathbf{p}))^{\frac{1}{2}} + \tilde{\Delta}_{\circ}^r(\mathbf{p}, \omega) \mathcal{E}_{r,y_0}(2\mathbf{p}) + \Delta_{\circ}^r(\mathbf{p}, \omega) \tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p}) \\ &\quad + \left(\tilde{\Delta}_{\otimes}^r(\mathbf{p}, \omega) \mathcal{E}_{r,y_0}(2\mathbf{p}) + \Delta_{\otimes}^r(\mathbf{p}, \omega) \tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p}) \right) (\mathcal{E}_{r,y_0}(2\mathbf{p}))^{\frac{1}{2}} \end{aligned}$$

$$+ \left(\check{\Delta}_{\oplus}^r(\mathbf{p}, \omega) \mathcal{E}_{r,y_0}(2\mathbf{p}) + \Delta_{\oplus}^r(\mathbf{p}, \omega) \tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p}) \right) (\mathcal{E}_{r,y_0}(2\mathbf{p})), \quad (\text{C.75})$$

where

$$\begin{aligned} \Delta_{\circ}^r(\mathbf{p}, \omega) &= \frac{\Delta_{\circ_1}^r}{2\lambda} + \frac{1}{|2\lambda - \omega|} (\Delta_{\circ_2}^r(\mathbf{p}) + \sqrt{\omega} \Delta_{\circ_3}^r(\mathbf{p}) + \omega \Delta_{\circ_4}^r(\mathbf{p})), \\ \check{\Delta}_{\ominus}^r(\mathbf{p}, \omega) &= \frac{\Delta_{\ominus_1}^r}{2\lambda} + \frac{1}{|2\lambda - \omega|} (\Delta_{\ominus_2}^r(\mathbf{p}) + \sqrt{\omega} \Delta_{\ominus_3}^r(\mathbf{p}) + \omega \Delta_{\ominus_4}^r(\mathbf{p})), \\ \Delta_{\odot}^r(\mathbf{p}, \omega) &= \frac{\Delta_{\odot_2}^r}{2\lambda} + \frac{\Delta_{\odot_3}^r(\mathbf{p})}{2\sqrt{\lambda}} + \frac{1}{|2\lambda - \omega|} (\Delta_{\odot_5}^r(\mathbf{p}) + \sqrt{\omega} \Delta_{\odot_7}^r(\mathbf{p})), \\ \check{\Delta}_{\odot}^r(\mathbf{p}, \omega) &= \frac{\Delta_{\odot_1}^r}{2\lambda} + \frac{1}{|2\lambda - \omega|} (\Delta_{\odot_4}^r(\mathbf{p}) + \sqrt{\omega} \Delta_{\odot_6}^r(\mathbf{p}) + \omega \Delta_{\odot_8}^r(\mathbf{p})), \\ \Delta_{\otimes}^r(\mathbf{p}, \omega) &= \frac{1}{|2\lambda - \omega|} (\Delta_{\otimes_3}^r(\mathbf{p}) + \sqrt{\omega} \Delta_{\otimes_5}^r(\mathbf{p})), \quad \check{\Delta}_{\otimes}^r(\mathbf{p}, \omega) = \frac{\Delta_{\otimes_1}^r(\mathbf{p})}{2\sqrt{\lambda}} + \frac{1}{|2\lambda - \omega|} (\Delta_{\otimes_2}^r(\mathbf{p}) + \sqrt{\omega} \Delta_{\otimes_4}^r(\mathbf{p})), \\ \Delta_{\oplus}^r(\mathbf{p}, \omega) &= \frac{\Delta_{\oplus_1}^r}{2\lambda} + \frac{\Delta_{\oplus_3}^r(\mathbf{p})}{|2\lambda - \omega|}, \quad \check{\Delta}_{\oplus}^r(\mathbf{p}, \omega) = \frac{\Delta_{\oplus_2}^r(\mathbf{p})}{|2\lambda - \omega|}, \end{aligned}$$

and the constants $\Delta_i^r(\mathbf{p})$, $i \in \{\circ_1, \dots, \circ_4, \ominus_1, \dots, \ominus_4, \odot_1, \dots, \odot_8, \otimes_1, \dots, \otimes_5, \oplus_1, \oplus_2, \oplus_3\}$, are defined in (A.3) - (A.7) in Section A.1. Additionally, the term

$$\mathbb{E}_{y_0} \left[e^{-2\lambda\tau(t)} \Xi^r(\tau(t), Y_N) \right] + \mathbb{E}_{y_0} \left[e^{-(2\lambda+\omega)\tau(t)} \Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega) \right],$$

can be bounded for all $t \in \mathbb{R}_{\geq 0}$ by the right hand side of (C.75) evaluated at $\mathbf{p} = 1$.

Proof. We begin by combining the results from Lemmas C.1 and C.2 to obtain the following bounds that hold for all $t \in [0, t^*]$:

$$\begin{aligned} \mathbb{E}_{y_0} \left[e^{-2\lambda t} \Xi^r(t, Y_N) \right] + \mathbb{E}_{y_0} \left[e^{-(2\lambda+\omega)\tau(t)} \Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega) \right] \\ \leq \frac{\Delta_{\Xi_1}^r(1, t^*)}{2\lambda} + \frac{1}{|2\lambda - \omega|} (\Delta_{\mathcal{U}_1}^r(1, t^*) + \sqrt{\omega} \Delta_{\mathcal{U}_2}^r(1, t^*) + \omega \Delta_{\mathcal{U}_3}^r(1, t^*)), \quad (\text{C.76a}) \end{aligned}$$

$$\begin{aligned} \|e^{-2\lambda t} \Xi^r(t, Y_N)\|_{L_p}^{y_0} + \|e^{-(2\lambda+\omega)\tau(t)} \Xi_{\mathcal{U}}^r(\tau(t), Y_N; \omega)\|_{L_p}^{y_0} \\ \leq \frac{\Delta_{\Xi_1}^r(\mathbf{p})}{2\lambda} + \frac{\Delta_{\Xi_2}^r(\mathbf{p})}{2\sqrt{\lambda}} + \frac{1}{|2\lambda - \omega|} (\Delta_{\mathcal{U}_1}^r(\mathbf{p}, t^*) + \sqrt{\omega} \Delta_{\mathcal{U}_2}^r(\mathbf{p}, t^*) + \omega \Delta_{\mathcal{U}_3}^r(\mathbf{p}, t^*)), \quad \forall \mathbf{p} \in \mathbb{N}_{\geq 2}. \quad (\text{C.76b}) \end{aligned}$$

Next, we use the definitions $\bar{F}_{\sigma}(\nu, \cdot) = p(\nu, \cdot)$, $\partial \mathcal{V}(Y_N) = 2(X_N^r - X_N^*)$, along with Assumptions 1, 4, and 5 to arrive at the following bounds:

$$\sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Y_{N,\nu})\|_{L_{2p}}^{y_0} \leq 2\tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p}), \quad \sup_{\nu \in [0, t^*]} \|\Lambda_{\mu}^{\perp}(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} \leq \Delta_{\mu}^{\perp} + \Delta_{\mu}^{\perp} \mathcal{E}_{r,y_0}(2\mathbf{p}), \quad (\text{C.77a})$$

$$\sup_{\nu \in [0, t^*]} \|F_{\sigma}(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} \leq \Delta_p + \Delta_{\sigma} + \Delta_{\sigma} \sqrt{\mathcal{E}_{r,y_0}(2\mathbf{p})}, \quad (\text{C.77b})$$

$$\sup_{\nu \in [0, t^*]} \|F_{\sigma}^{\perp}(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} \leq \Delta_p^{\perp} + \Delta_{\sigma}^{\perp} + \Delta_{\sigma}^{\perp} \sqrt{\mathcal{E}_{r,y_0}(2\mathbf{p})}, \quad \sup_{\nu \in [0, t^*]} \|\bar{F}_{\sigma}(\nu, X_{N,\nu}^*)\|_{L_{2p}}^{y_0} \leq \Delta_p, \quad (\text{C.77c})$$

where we have further used the subadditivity of the square root function. We have also used the manipulations in (C.72).

Now, recall the definitions of $\Delta_{\Xi_1}^r(\mathbf{p})$ and $\Delta_{\Xi_2}^r(\mathbf{p})$ in Lemma C.1 for $\mathbf{p} \in \mathbb{N}_{\geq 1}$:

$$\begin{aligned} \Delta_{\Xi_1}^r(\mathbf{p}, t^*) &= \sup_{\nu \in [0, t^*]} \left(\Delta_g^{\perp} \|\partial \mathcal{V}(Y_{N,\nu})\|_{L_{2p}}^{y_0} \|\Lambda_{\mu}^{\perp}(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} \right. \\ &\quad \left. + \mathbb{E}_{y_0} \left[F_{\sigma}(\nu, X_{N,\nu}^r)^{2p} \right]^{\frac{1}{p}} + \mathbb{E}_{y_0} \left[\bar{F}_{\sigma}(\nu, X_{N,\nu}^*)^{2p} \right]^{\frac{1}{p}} \right), \\ \Delta_{\Xi_2}^r(\mathbf{p}, t^*) &= \mathbf{p}(\mathbf{p}) \sup_{\nu \in [0, t^*]} \left(\Delta_g^{\perp} \|F_{\sigma}^{\perp}(\nu, X_{N,\nu}^r)\|_{L_{2p}}^{y_0} + \|\bar{F}_{\sigma}(\nu, X_{N,\nu}^*)\|_{L_{2p}}^{y_0} \right) \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Y_{N,\nu})\|_{L_{2p}}^{y_0}. \end{aligned}$$

We can bound these terms by using (C.77) as follows:

$$\begin{aligned}\Delta_{\Xi_1}^r(\mathbf{p}, t^*) &\leq \Delta_p^2 + 2\Delta_g^\perp \Delta_\mu^\perp \tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p}) (1 + \mathcal{E}_{r,y_0}(2\mathbf{p})) + \left(\Delta_p + \Delta_\sigma + \Delta_\sigma \sqrt{\mathcal{E}_{r,y_0}(2\mathbf{p})} \right)^2, \\ \Delta_{\Xi_2}^r(\mathbf{p}, t^*) &\leq 2\mathbf{p}(\mathbf{p}) \left(\Delta_g^\perp (\Delta_p^\perp + \Delta_\sigma^\perp) + \Delta_p + \Delta_g^\perp \Delta_\sigma^\perp \sqrt{\mathcal{E}_{r,y_0}(2\mathbf{p})} \right) \tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p}),\end{aligned}$$

for all $\mathbf{p} \in \{1, \mathbf{p}^*\}$. We can therefore write the above as

$$\Delta_{\Xi_1}^r(\mathbf{p}, t^*) \leq \Delta_{\odot_1}^r + \Delta_{\odot_1}^r (\mathcal{E}_{r,y_0}(2\mathbf{p}))^{\frac{1}{2}} + \Delta_{\odot_1}^r \mathcal{E}_{r,y_0}(2\mathbf{p}) + \Delta_{\odot_2}^r \tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p}) + \Delta_{\odot_1}^r \mathcal{E}_{r,y_0}(2\mathbf{p}) \tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p}), \quad (\text{C.79a})$$

$$\Delta_{\Xi_2}^r(\mathbf{p}, t^*) \leq \Delta_{\odot_3}^r(\mathbf{p}) \tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p}) + \Delta_{\odot_1}^r(\mathbf{p}) (\mathcal{E}_{r,y_0}(2\mathbf{p}))^{\frac{1}{2}} \tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p}), \quad (\text{C.79b})$$

for all $\mathbf{p} \in \{1, \mathbf{p}^*\}$.

Next, recall the terms $\Delta_{\mathcal{U}_1}^r$, $\Delta_{\mathcal{U}_2}^r$, and $\Delta_{\mathcal{U}_3}^r$ in Lemma C.2:

$$\begin{aligned}\Delta_{\mathcal{U}_1}^r(\mathbf{p}, t^*) &= \sqrt{\lambda} \Delta_g \mathbf{p}(\mathbf{p}) \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Y_{N,\nu})\|_{L_{2\mathbf{p}}} \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2\mathbf{p}}} \\ &\quad + \left(\frac{\Delta_{\mathcal{P}_1}^r(\mathbf{p}, t^*)}{\sqrt{\lambda}} + 2\Delta_g \left[\sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Y_{N,\nu})\|_{L_{2\mathbf{p}}}^{y_0} + \frac{\Delta_g}{2\lambda} \sup_{\nu \in [0, t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2\mathbf{p}}}^{y_0} \right] \right) \\ &\quad \times \sup_{\nu \in [0, t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2\mathbf{p}}}^{y_0}, \\ \Delta_{\mathcal{U}_2}^r(\mathbf{p}, t^*) &= \mathbf{p}'(\mathbf{p}) \left(\frac{\Delta_{\mathcal{P}_1}^r(\mathbf{p}, t^*)}{\sqrt{\lambda}} + \Delta_g \left[\sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Y_{N,\nu})\|_{L_{2\mathbf{p}}}^{y_0} + \frac{2\Delta_g}{\lambda} \sup_{\nu \in [0, t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2\mathbf{p}}}^{y_0} \right] \right) \\ &\quad \times \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2\mathbf{p}}}^{y_0}, \\ \Delta_{\mathcal{U}_3}^r(\mathbf{p}, t^*) &= \left(\frac{\Delta_{\mathcal{P}_3}^r(\mathbf{p}, t^*)}{\lambda} + \frac{\Delta_g^2 \mathbf{p}'(\mathbf{p})^2}{\lambda} \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2\mathbf{p}}}^{y_0} \right) \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2\mathbf{p}}}^{y_0},\end{aligned}$$

where we have further used the definition $\Delta_{\mathcal{P}_2}^r(\mathbf{p}, t^*) = \mathbf{p}'(\mathbf{p}) \Delta_{\mathcal{P}_1}^r(\mathbf{p}, t^*)$ from Proposition C.5. In order to bound these terms, in addition to (C.66) in Proposition C.7, and the bounds in (C.77) above, we also appeal to the following due to Assumptions 1, 4, and 5:

$$\begin{aligned}\sup_{\nu \in [0, t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2\mathbf{p}}}^{y_0} &\leq \Delta_\mu^\parallel + \Delta_\mu^\parallel \mathcal{E}_{r,y_0}(2\mathbf{p}), \\ \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu}^r)\|_{L_{2\mathbf{p}}}^{y_0} &\leq \Delta_p^\parallel + \Delta_\sigma^\parallel + \Delta_\sigma^\parallel \sqrt{\mathcal{E}_{r,y_0}(2\mathbf{p})}.\end{aligned} \quad (\text{C.80})$$

Using the above then leads to

$$\begin{aligned}\Delta_{\mathcal{U}_1}^r(\mathbf{p}, t^*) &\leq \Delta_{\odot_2}^r(\mathbf{p}) + \Delta_{\odot_2}^r(\mathbf{p}) (\mathcal{E}_{r,y_0}(2\mathbf{p}))^{\frac{1}{2}} + \Delta_{\odot_4}^r(\mathbf{p}) \mathcal{E}_{r,y_0}(2\mathbf{p}) + \Delta_{\odot_5}^r(\mathbf{p}) \tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p}) \\ &\quad + \left(\Delta_{\odot_2}^r(\mathbf{p}) \mathcal{E}_{r,y_0}(2\mathbf{p}) + \Delta_{\odot_3}^r(\mathbf{p}) \tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p}) \right) (\mathcal{E}_{r,y_0}(2\mathbf{p}))^{\frac{1}{2}} \\ &\quad + \left(\Delta_{\odot_2}^r \mathcal{E}_{r,y_0}(2\mathbf{p}) + \Delta_{\odot_3}^r \tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p}) \right) \mathcal{E}_{r,y_0}(2\mathbf{p}), \quad (\text{C.81a})\end{aligned}$$

$$\begin{aligned}\Delta_{\mathcal{U}_2}^r(\mathbf{p}, t^*) &\leq \Delta_{\odot_3}^r(\mathbf{p}) + \Delta_{\odot_3}^r(\mathbf{p}) (\mathcal{E}_{r,y_0}(2\mathbf{p}))^{\frac{1}{2}} + \Delta_{\odot_6}^r(\mathbf{p}) \mathcal{E}_{r,y_0}(2\mathbf{p}) + \Delta_{\odot_7}^r(\mathbf{p}) \tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p}) \\ &\quad + \left(\Delta_{\odot_4}^r(\mathbf{p}) \mathcal{E}_{r,y_0}(2\mathbf{p}) + \Delta_{\odot_5}^r(\mathbf{p}) \tilde{\mathcal{E}}_{r,y_0}(2\mathbf{p}) \right) (\mathcal{E}_{r,y_0}(2\mathbf{p}))^{\frac{1}{2}}, \quad (\text{C.81b})\end{aligned}$$

$$\Delta_{\mathcal{U}_3}^r(\mathbf{p}, t^*) \leq \Delta_{\odot_4}^r(\mathbf{p}) + \Delta_{\odot_4}^r(\mathbf{p}) (\mathcal{E}_{r,y_0}(2\mathbf{p}))^{\frac{1}{2}} + \Delta_{\odot_8}^r(\mathbf{p}) \mathcal{E}_{r,y_0}(2\mathbf{p}). \quad (\text{C.81c})$$

The proof is then concluded by substituting (C.79) - (C.81) into (C.76).

□

D True (Uncertain) Process

We first begin by deriving a couple of expressions regarding the \mathcal{L}_1 -DRAC feedback operator $\mathcal{F}_{\mathcal{L}_1}$ and the reference feedback operator \mathcal{F}_r .

Proposition D.1 *Consider the \mathcal{L}_1 -DRAC feedback operator $\mathcal{F}_{\mathcal{L}_1}$ and the reference feedback operator \mathcal{F}_r defined in (13) and (27), respectively. Assume that the closed-loop system (90) admits $z \in \mathcal{M}_2(\mathbb{R}^n \mid \mathfrak{W}_t)$ as a unique strong solution, for all $t \in [0, \tau]$, where $\tau \in [0, T]$ ⁸ is a stopping time. Then,*

$$(\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r)(z)(t) = \omega e^{-\omega t} \tilde{\mathcal{F}}(z, t), \quad t \in [0, \tau], \quad (\text{D.1})$$

where $\tilde{\mathcal{F}}(z, t)$ is defined as

$$\tilde{\mathcal{F}}(z, t) = \int_0^t e^{\omega \nu} \left(\Sigma_\nu^\parallel(z) - \hat{\Lambda}^\parallel(\nu) d\nu \right), \quad (\text{D.2})$$

for $t \in [0, \tau]$, $z_0 = z(0)$, and

$$\begin{aligned} \hat{\Lambda}^\parallel(t) &= 0_m \mathbb{1}_{[0, T_s) \cap [0, \tau]}(t) \\ &\quad - \sum_{i \in \mathbb{N}_{\geq 1}} \left(\lambda_s (1 - e^{\lambda_s T_s})^{-1} \Theta_{ad}(iT_s) \int_{(i-1)T_s}^{iT_s} e^{-\lambda_s(iT_s - \nu)} \Sigma_\nu(z) d\nu \right) \mathbb{1}_{[iT_s, (i+1)T_s) \cap [0, \tau]}(t). \end{aligned} \quad (\text{D.3})$$

In the expressions above, we have (formally) defined

$$\Sigma_t^{\{\cdot, \parallel\}}(z) \doteq \Lambda_\mu^{\{\cdot, \parallel\}}(t, z(t)) dt + F_\sigma^{\{\cdot, \parallel\}}(t, z(t)) dW_t, \quad t \in [0, \tau]. \quad (\text{D.4})$$

Proof. Since z is assumed to be the unique strong solution of (90), we have that

$$\begin{aligned} dz(t) &= F_\mu(t, z(t), \mathcal{F}_{\mathcal{L}_1}(z)(t)) dt + F_\sigma(t, z(t)) dW_t \\ &= (f(t, z(t)) + g(t) \mathcal{F}_{\mathcal{L}_1}(z)(t) + \Lambda_\mu(t, z(t))) dt + F_\sigma(t, z(t)) dW_t, \quad z(0) = z_0 \sim \xi_0 \text{ (}\mathbb{P}\text{-a.s.)}, \end{aligned} \quad (\text{D.5})$$

where we have set $U_{\mathcal{L}_1, t} = \mathcal{F}_{\mathcal{L}_1}(z)(t)$. Next, using the definition of $\mathcal{F}_{\mathcal{L}_1}$ from (14) we set

$$\mathcal{F}_{\mathcal{L}_1}(z) = \mathcal{F}_\omega(\hat{\Lambda}^\parallel), \quad \hat{\Lambda}^\parallel = \mathcal{F}_{T_s}^\parallel(\hat{\Lambda}), \quad \hat{\Lambda} = \mathcal{F}_{T_s}(\hat{z}, z), \quad \hat{z} = \mathcal{F}_{\lambda_s}(z). \quad (\text{D.6})$$

Hence, using the definition of \mathcal{F}_r from (27), we obtain

$$\begin{aligned} ((\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r)(z))(t) &= -\omega \int_0^t e^{-\omega(t-\nu)} \hat{\Lambda}^\parallel(\nu) d\nu - \mathcal{F}_\omega(\Lambda_\mu^\parallel(\cdot, z))(t) - \mathcal{F}_{N, \omega}(F_\sigma^\parallel(\cdot, z), W)(t) \\ &= \omega e^{-\omega t} \int_0^t e^{\omega \nu} \left(\Sigma_\nu^\parallel(z) - \hat{\Lambda}^\parallel(\nu) d\nu \right), \end{aligned}$$

for $t \in [0, \tau]$, where we have used the definition of $\Sigma_t^\parallel(z)$ in (D.4). Then, for any stopping time $\hat{\tau} \in [0, \tau]$ one obtains the following:

$$\begin{aligned} ((\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r)(z))(t) &= \omega e^{-\omega t} \left(\int_0^{\hat{\tau}} e^{\omega \nu} \left(\Sigma_\nu^\parallel(z) - \hat{\Lambda}^\parallel(\nu) d\nu \right) + \int_{\hat{\tau}}^t e^{\omega \nu} \left(\Sigma_\nu^\parallel(z) - \hat{\Lambda}^\parallel(\nu) d\nu \right) \right) \\ &= \omega e^{-\omega t} \left(\hat{\mathcal{F}}(\hat{\tau}) + \int_{\hat{\tau}}^t e^{\omega \nu} \left(\Sigma_\nu^\parallel(z) - \hat{\Lambda}^\parallel(\nu) d\nu \right) \right), \end{aligned}$$

for $t \in [\hat{\tau}, \tau]$, where we have defined

$$\hat{\mathcal{F}}(t) \doteq \int_0^t e^{\omega \nu} \left(\Sigma_\nu^\parallel(z) - \hat{\Lambda}^\parallel(\nu) d\nu \right).$$

Since the temporal instances iT_s , $i \in \mathbb{N}$, are constant, and hence stopping times, we can write

$$((\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r)(z))(t) = \begin{cases} \omega e^{-\omega t} \hat{\mathcal{F}}(t) = \omega e^{-\omega t} \int_0^t e^{\omega \nu} \left(\Sigma_\nu^\parallel - \hat{\Lambda}^\parallel(\nu) d\nu \right), & t \in [0, T_s) \cap [0, \tau] \\ \omega e^{-\omega t} \left(\hat{\mathcal{F}}(iT_s) + \int_{iT_s}^t e^{\omega \nu} \left(\Sigma_\nu^\parallel(z) - \hat{\Lambda}^\parallel(\nu) d\nu \right) \right), & t \in [iT_s, (i+1)T_s) \cap [0, \tau] \end{cases},$$

⁸For example, given any stopping time $\tau' \in \mathbb{R}_{\geq 0}$, we can choose $\tau \doteq \tau' \wedge T$.

for $i \in \mathbb{N}_{\geq 1}$, thus proving (D.2) since $\hat{\mathcal{F}}(t) = \tilde{\mathcal{F}}(z, t)$, over each interval $[iT_s, (i+1)T_s) \cap [0, \tau]$, for $i \in \mathbb{N}_{\geq 1}$.

Next, we derive the expression for the adaptive estimate $\hat{\Lambda}^\parallel(t)$, which is defined in (D.6) as

$$\hat{\Lambda}^\parallel = \mathcal{F}_{T_s}^\parallel(\hat{\Lambda}), \quad \hat{\Lambda} = \mathcal{F}_{T_s}(\hat{z}, z).$$

It then follows from the definitions of the adaptation law operators \mathcal{F}_{T_s} and $\mathcal{F}_{T_s}^\parallel$ in (15b) that

$$\hat{\Lambda}^\parallel(t) = \sum_{i \in \mathbb{N}_{\geq 0}} \Theta_{ad}(iT_s) \hat{\Lambda}(t) \mathbb{1}_{[iT_s, (i+1)T_s) \cap [0, \tau]}(t), \quad (\text{D.7a})$$

$$\hat{\Lambda}(t) = 0_n \mathbb{1}_{[0, T_s)}(t) + \lambda_s (1 - e^{\lambda_s T_s})^{-1} \sum_{i \in \mathbb{N}_{\geq 1}} \tilde{z}(iT_s) \mathbb{1}_{[iT_s, (i+1)T_s) \cap [0, \tau]}(t), \quad (\text{D.7b})$$

where $\tilde{z} \doteq \hat{z} - z$. Moreover, $\hat{z} = \mathcal{F}_{\lambda_s}(z)$ is the prediction process defined via the operator \mathcal{F}_{λ_s} in (15c) as follows:

$$\hat{z}(t) = z_0 + \int_0^t \left(-\lambda_s \mathbb{I}_n \tilde{z}(\nu) + f(\nu, z(\nu)) + g(\nu) \mathcal{F}_{\mathcal{L}_1}(z)(\nu) + \hat{\Lambda}(\nu) \right) d\nu, \quad t \in [0, \tau].$$

We may write the integral equation above in its differential form as follows:

$$d\hat{z}(t) = \left(-\lambda_s \mathbb{I}_n \tilde{z}(t) + f(t, z(t)) + g(t) \mathcal{F}_{\mathcal{L}_1}(z)(t) + \hat{\Lambda}(t) \right) dt, \quad \hat{z}(0) = z_0, \quad t \in [0, \tau].$$

Subtracting (D.5) from the above equation then leads to differential equation for the error process $\tilde{z} \doteq \hat{z} - z$ as follows:

$$\begin{aligned} d\tilde{z}(t) &= \left(-\lambda_s \mathbb{I}_n \tilde{z}(t) + \hat{\Lambda}(t) - \Lambda_\mu(t, z(t)) \right) dt - F_\sigma(t, z(t)) dW_t \\ &= \left(-\lambda_s \mathbb{I}_n \tilde{z}(t) + \hat{\Lambda}(t) \right) dt - \Sigma_t(z), \quad \tilde{z}(0) = 0_n, \quad t \in [0, \tau], \end{aligned} \quad (\text{D.8})$$

where Σ_t is given in (D.4). Therefore, we can write the integral form over the stopping times iT_s , $i \in \mathbb{N}$, as follows

$$\begin{aligned} \tilde{z}(t) &= \int_0^t \left(-\lambda_s \mathbb{I}_n \tilde{z}(\nu) + \hat{\Lambda}(\nu) \right) d\nu - \int_0^t \Sigma_\nu(z), \quad t \in [0, T_s) \cap [0, \tau], \\ \tilde{z}(t) &= \tilde{z}(iT_s) + \int_{iT_s}^t \left(-\lambda_s \mathbb{I}_n \tilde{z}(\nu) + \hat{\Lambda}(\nu) \right) d\nu - \int_{iT_s}^t \Sigma_\nu(z), \quad t \in [iT_s, (i+1)T_s) \cap [0, \tau], \end{aligned}$$

for $i \in \mathbb{N}_{\geq 1}$. Substituting the adaptive estimate $\hat{\Lambda}(t)$ from (D.7) into the above equations, we obtain

$$\begin{aligned} \tilde{z}(t) &= - \int_0^t \lambda_s \mathbb{I}_n \tilde{z}(\nu) d\nu - \int_0^t \Sigma_\nu(z), \quad t \in [0, T_s) \cap [0, \tau], \\ \tilde{z}(t) &= \tilde{z}(iT_s) + \int_{iT_s}^t \left(-\lambda_s \mathbb{I}_n \tilde{z}(\nu) + \lambda_s (1 - e^{\lambda_s T_s})^{-1} \tilde{z}(iT_s) \right) d\nu - \int_{iT_s}^t \Sigma_\nu(z), \quad t \in [iT_s, (i+1)T_s) \cap [0, \tau], \end{aligned}$$

for $i \in \mathbb{N}_{\geq 1}$. Since the expressions above represent linear integral equations in the variable \tilde{z} , passing through their linear differential forms, we obtain their solutions as follows (see e.g. [107, Sec. 5.4.2]):

$$\begin{aligned} \tilde{z}(t) &= - \int_0^t e^{-\lambda_s(t-\nu)} \Sigma_\nu(z), \quad t \in ([0, T_s) \cap [0, \tau]), \\ \tilde{z}(t) &= e^{-\lambda_s(t-iT_s)} \tilde{z}(iT_s) + \lambda_s (1 - e^{\lambda_s T_s})^{-1} \tilde{z}(iT_s) \int_{iT_s}^t e^{-\lambda_s(t-\nu)} d\nu \\ &\quad - \int_{iT_s}^t e^{-\lambda_s(t-\nu)} \Sigma_\nu(z), \quad t \in ([iT_s, (i+1)T_s) \cap [0, \tau]). \end{aligned}$$

Hence, at the temporal instances $t = iT_s$, $i \in \mathbb{N}_{\geq 1}$, the error process \tilde{z} is given by

$$\tilde{z}(iT_s) = - \int_{(i-1)T_s}^{iT_s} e^{-\lambda_s(iT_s-\nu)} \Sigma_\nu(z), \quad i \in \mathbb{N}_{\geq 1}.$$

Substituting the above into (D.7b) then leads to

$$\begin{aligned} \hat{\Lambda}(t) &= 0_n, \quad t \in [0, T_s) \cap [0, \tau], \\ \hat{\Lambda}(t) &= -\lambda_s (1 - e^{\lambda_s T_s})^{-1} \int_{(i-1)T_s}^{iT_s} e^{-\lambda_s(iT_s-\nu)} \Sigma_\nu(\nu), \quad t \in [iT_s, (i+1)T_s) \cap [0, \tau], \quad i \in \mathbb{N}_{\geq 1}. \end{aligned}$$

Thus, we obtain the expression for the matched adaptive estimate in (D.3) by substituting the above into (D.7a). \square

We will need the following trivial result.

Proposition D.2 *We have the following identities:*

$$\begin{aligned} \Theta_{ad}(i\mathbf{T}_s) \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s \nu} \Sigma_\nu(z) &= \Theta_{ad}(i\mathbf{T}_s) \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s \nu} \widetilde{\Sigma}_\nu(i\mathbf{T}_s, z) \\ &\quad + \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s \nu} [\Lambda_\mu^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) d\nu + F_\sigma^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) dW_\nu], \end{aligned} \quad (\text{D.9})$$

and

$$\int_{i\mathbf{T}_s}^t e^{\omega \nu} \Sigma_\nu^\parallel(z) = \int_{i\mathbf{T}_s}^t e^{\omega \nu} \widetilde{\Sigma}_\nu^\parallel(i\mathbf{T}_s, z) + \int_{i\mathbf{T}_s}^t e^{\omega \nu} [\Lambda_\mu^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) d\nu + F_\sigma^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) dW_\nu], \quad (\text{D.10})$$

for all $t \in [i\mathbf{T}_s, \tau]$, $i \in \mathbb{N}_{\geq 1}$, where $\Sigma_t^{\{\cdot, \parallel\}}(z)$ is defined in (D.4), and

$$\widetilde{\Sigma}_t^{\{\cdot, \parallel\}}(i\mathbf{T}_s, z) = \widetilde{\Lambda}_\mu^{\{\cdot, \parallel\}}(t, i\mathbf{T}_s, z) dt + \widetilde{F}_\sigma^{\{\cdot, \parallel\}}(t, i\mathbf{T}_s, z) dW_t, \quad (\text{D.11a})$$

$$\widetilde{\Lambda}_\mu^{\{\cdot, \parallel\}}(t, i\mathbf{T}_s, z) = \Lambda_\mu^{\{\cdot, \parallel, \perp\}}(t, z(t)) - \Lambda_\mu^{\{\cdot, \parallel, \perp\}}(i\mathbf{T}_s, z(i\mathbf{T}_s)), \quad (\text{D.11b})$$

$$\widetilde{F}_\sigma^{\{\cdot, \parallel\}}(t, i\mathbf{T}_s, z) = F_\sigma^{\{\cdot, \parallel, \perp\}}(t, z(t)) - F_\sigma^{\{\cdot, \parallel, \perp\}}(i\mathbf{T}_s, z(i\mathbf{T}_s)). \quad (\text{D.11c})$$

Proof. We begin by adding and subtracting the terms $\Lambda_\mu(i\mathbf{T}_s, z(i\mathbf{T}_s))$ and $F_\sigma(i\mathbf{T}_s, z(i\mathbf{T}_s))$ to obtain the following expression:

$$\begin{aligned} \Theta_{ad}(i\mathbf{T}_s) \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s \nu} \Sigma_\nu(z) &= \Theta_{ad}(i\mathbf{T}_s) \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s \nu} \widetilde{\Sigma}_\nu(i\mathbf{T}_s, z) \\ &\quad + \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s \nu} [\Lambda_\mu^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) d\nu + F_\sigma^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) dW_\nu], \end{aligned} \quad (\text{D.12})$$

where we have used the definition of Θ_{ad} from (15b) to observe that

$$\begin{aligned} \Theta_{ad}(i\mathbf{T}_s) \Lambda_\mu(i\mathbf{T}_s, z(i\mathbf{T}_s)) &= [\mathbb{I}_m \quad 0_{m, n-m}] \bar{g}(i\mathbf{T}_s)^{-1} \Lambda_\mu(i\mathbf{T}_s, z(i\mathbf{T}_s)) \\ &= [\mathbb{I}_m \quad 0_{m, n-m}] [g(i\mathbf{T}_s) \quad g(i\mathbf{T}_s)^\perp]^{-1} \Lambda_\mu(i\mathbf{T}_s, z(i\mathbf{T}_s)), \end{aligned}$$

which further leads to

$$\Theta_{ad}(i\mathbf{T}_s) \Lambda_\mu(i\mathbf{T}_s, z(i\mathbf{T}_s)) = [\mathbb{I}_m \quad 0_{m, n-m}] \begin{bmatrix} \Lambda_\mu^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) \\ \Lambda_\mu^\perp(i\mathbf{T}_s, z(i\mathbf{T}_s)) \end{bmatrix} = \Lambda_\mu^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)).$$

An identical line of reasoning leads to $\Theta_{ad}(i\mathbf{T}_s) F_\sigma(i\mathbf{T}_s, z(i\mathbf{T}_s)) = F_\sigma^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s))$. The identity in (D.9) is thus established.

The expression in (D.10) follows similarly by adding and subtracting $\Lambda_\mu^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s))$ and $F_\sigma^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s))$ to and from the integrand of $\int_{i\mathbf{T}_s}^t e^{\omega \nu} \Sigma_\nu^\parallel(z)$. \square

The following result is a further development of Proposition D.1.

Proposition D.3 *Assume that the closed-loop system (90) admits $z \in \mathcal{M}_2(\mathbb{R}^n \mid \mathfrak{W}_t)$ as a unique strong solution, for all $t \in [0, \tau]$, where $\tau \in [0, T]$ is a stopping time. Then, the process $\widetilde{\mathcal{F}}(z, t)$ defined in the statement of Proposition D.1 can be expressed as follows:*

$$\begin{aligned} \widetilde{\mathcal{F}}(z, t) &= \mathbb{1}_{<\mathbf{T}_s}(t) \int_0^t e^{\omega \nu} \Sigma_\nu^\parallel(z) + \mathbb{1}_{\geq \mathbf{T}_s}(t) \int_0^{\mathbf{T}_s} e^{\omega \nu} \Lambda_\mu^\parallel(\nu, z(\nu)) d\nu \\ &\quad + \mathbb{1}_{\geq \mathbf{T}_s}(t) \int_0^t e^{\omega \nu} \left[\left(\widetilde{\mathcal{H}}_\mu^\parallel(\nu, z) + e^{(\lambda_s - \omega)\nu} \widetilde{\mathcal{H}}_\mu(\nu, z) \right) d\nu + \left(\widetilde{\mathcal{H}}_\sigma^\parallel(\nu, z) + e^{(\lambda_s - \omega)\nu} \widetilde{\mathcal{H}}_\sigma(\nu, z) \right) dW_\nu \right] \\ &\quad + \mathbb{1}_{\geq \mathbf{T}_s}(t) \int_0^t e^{\omega \nu} \left(\widetilde{\mathcal{G}}_{\sigma_1}(\nu, z) + e^{(\lambda_s - \omega)\nu} \widetilde{\mathcal{G}}_{\sigma_2}(\nu, z) \right) dW_\nu \end{aligned}$$

$$+ \mathbb{1}_{\geq \mathbf{T}_s}(t) \int_0^t e^{\omega \nu} [\mathcal{G}_\mu(\nu, z) d\nu + (\mathcal{G}_{\sigma_1}(\nu, z) + \mathcal{G}_{\sigma_2}(\nu, z)) dW_\nu], \quad (\text{D.13})$$

for $t \in [0, \tau]$, where $\Sigma_\nu^\parallel(z)$ is defined in (D.4), and for $\nu \in [0, t]$, we have

$$\begin{aligned} \tilde{\mathcal{H}}_\mu^\parallel(\nu, z) &= \mathbb{1}_{< \mathbf{T}_s}(\nu) 0_m + \sum_{i=1}^{i^*(t)-1} \mathbb{1}_{[i\mathbf{T}_s, (i+1)\mathbf{T}_s)}(\nu) \tilde{\Lambda}_\mu^\parallel(\nu, i\mathbf{T}_s, z) + \mathbb{1}_{\geq i^*(t)\mathbf{T}_s}(\nu) \tilde{\Lambda}_\mu^\parallel(\nu, i^*(t)\mathbf{T}_s, z), \\ \tilde{\mathcal{H}}_\mu(\nu, z) &= - \sum_{i=1}^{i^*(t)-1} \mathbb{1}_{[(i-1)\mathbf{T}_s, i\mathbf{T}_s)}(\nu) \tilde{\gamma}_i(\mathbf{T}_s) \Theta_{ad}(i\mathbf{T}_s) \tilde{\Lambda}_\mu(\nu, i\mathbf{T}_s, z) \\ &\quad - \mathbb{1}_{[(i^*(t)-1)\mathbf{T}_s, i^*(t)\mathbf{T}_s)}(\nu) \tilde{\gamma}^*(t, \mathbf{T}_s) \Theta_{ad}(i^*(t)\mathbf{T}_s) \tilde{\Lambda}_\mu(\nu, i^*(t)\mathbf{T}_s, z) + \mathbb{1}_{\geq i^*(t)\mathbf{T}_s}(\nu) 0_m, \\ \tilde{\mathcal{H}}_\sigma^\parallel(\nu, z) &= \mathbb{1}_{< \mathbf{T}_s}(\nu) 0_m + \sum_{i=1}^{i^*(t)-1} \mathbb{1}_{[i\mathbf{T}_s, (i+1)\mathbf{T}_s)}(\nu) \tilde{F}_\sigma^\parallel(\nu, i\mathbf{T}_s, z) + \mathbb{1}_{\geq i^*(t)\mathbf{T}_s}(\nu) \tilde{F}_\sigma^\parallel(\nu, i^*(t)\mathbf{T}_s, z), \\ \tilde{\mathcal{H}}_\sigma(\nu, z) &= - \sum_{i=1}^{i^*(t)-1} \mathbb{1}_{[(i-1)\mathbf{T}_s, i\mathbf{T}_s)}(\nu) \tilde{\gamma}_i(\mathbf{T}_s) \Theta_{ad}(i\mathbf{T}_s) \tilde{F}_\sigma(\nu, i\mathbf{T}_s, z) \\ &\quad - \mathbb{1}_{[(i^*(t)-1)\mathbf{T}_s, i^*(t)\mathbf{T}_s)}(\nu) \tilde{\gamma}^*(t, \mathbf{T}_s) \Theta_{ad}(i^*(t)\mathbf{T}_s) \tilde{F}_\sigma(\nu, i^*(t)\mathbf{T}_s, z) + \mathbb{1}_{\geq i^*(t)\mathbf{T}_s}(\nu) 0_m. \end{aligned}$$

Furthermore, we have defined the following for $\nu \in [0, t]$:

$$\begin{aligned} \mathcal{G}_\mu(\nu, z) &= \mathbb{1}_{< \mathbf{T}_s}(\nu) 0_m + \sum_{i=1}^{i^*(t)-1} \mathbb{1}_{[i\mathbf{T}_s, (i+1)\mathbf{T}_s)}(\nu) (1 - e^{-\lambda_s \mathbf{T}_s}) \Lambda_\mu^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) \\ &\quad + \mathbb{1}_{\geq i^*(t)\mathbf{T}_s}(\nu) (1 - e^{-\lambda_s \mathbf{T}_s}) \Lambda_\mu^\parallel(i^*(t)\mathbf{T}_s, z(i^*(t)\mathbf{T}_s)), \\ \tilde{\mathcal{G}}_{\sigma_1}(\nu, z) &= \mathbb{1}_{< \mathbf{T}_s}(\nu) \tilde{F}_\sigma^\parallel(\nu, 0, z) + \mathbb{1}_{\geq \mathbf{T}_s}(\nu) 0_{m,d}, \\ \tilde{\mathcal{G}}_{\sigma_2}(\nu, z) &= - \sum_{i=1}^{i^*(t)-1} \mathbb{1}_{[(i-1)\mathbf{T}_s, i\mathbf{T}_s)}(\nu) \tilde{F}_\sigma^\parallel(i\mathbf{T}_s, (i-1)\mathbf{T}_s, z) \tilde{\gamma}_i(\mathbf{T}_s) + \mathbb{1}_{\geq (i^*(t)-1)\mathbf{T}_s}(\nu) 0_{m,d}, \\ \mathcal{G}_{\sigma_1}(\nu, z) &= \sum_{i=1}^{i^*(t)-1} \mathbb{1}_{[(i-1)\mathbf{T}_s, i\mathbf{T}_s)}(\nu) F_\sigma^\parallel((i-1)\mathbf{T}_s, z((i-1)\mathbf{T}_s)) (1 - \tilde{\gamma}_i(\mathbf{T}_s) e^{(\lambda_s - \omega)\nu}) \\ &\quad + \mathbb{1}_{[(i^*(t)-1)\mathbf{T}_s, i^*(t)\mathbf{T}_s)}(\nu) F_\sigma^\parallel(i^*(t)\mathbf{T}_s, z(i^*(t)\mathbf{T}_s)) + \mathbb{1}_{\geq i^*(t)\mathbf{T}_s}(\nu) 0_{m,d}, \\ \mathcal{G}_{\sigma_2}(\nu, z) &= \mathbb{1}_{< (i^*(t)-1)\mathbf{T}_s}(\nu) 0_{m,d} \\ &\quad - F_\sigma^\parallel(i^*(t)\mathbf{T}_s, z(i^*(t)\mathbf{T}_s)) \left(\mathbb{1}_{[(i^*(t)-1)\mathbf{T}_s, i^*(t)\mathbf{T}_s)}(\nu) \tilde{\gamma}^*(t, \mathbf{T}_s) e^{(\lambda_s - \omega)\nu} - \mathbb{1}_{\geq i^*(t)\mathbf{T}_s}(\nu) \right). \end{aligned}$$

The expressions above are defined using the entities in (D.11) and the following:

$$\begin{aligned} i^*(t) &\doteq \max \{i \in \mathbb{N}_{\geq 1} : i\mathbf{T}_s \leq t\} = \left\lfloor \frac{t}{\mathbf{T}_s} \right\rfloor, \\ \tilde{\gamma}_i(\mathbf{T}_s) &\doteq e^{(\omega - \lambda_s)i\mathbf{T}_s} \frac{\lambda_s}{\omega} \frac{e^{\omega \mathbf{T}_s} - 1}{e^{\lambda_s \mathbf{T}_s} - 1} = \int_{i\mathbf{T}_s}^{(i+1)\mathbf{T}_s} e^{\omega \nu} d\nu \left(\int_{i\mathbf{T}_s}^{(i+1)\mathbf{T}_s} e^{\lambda_s \nu} d\nu \right)^{-1}, \\ \tilde{\gamma}^*(t, \mathbf{T}_s) &\doteq e^{(\omega - \lambda_s)i^*(t)\mathbf{T}_s} \frac{\lambda_s}{\omega} \frac{e^{\omega(t - i^*(t)\mathbf{T}_s)} - 1}{e^{\lambda_s \mathbf{T}_s} - 1} = \int_{i^*(t)\mathbf{T}_s}^t e^{\omega \nu} d\nu \left(\int_{i^*(t)\mathbf{T}_s}^{(i^*(t)+1)\mathbf{T}_s} e^{\lambda_s \nu} d\nu \right)^{-1}. \end{aligned}$$

Proof. We have from (D.3) that $\hat{\Lambda}^\parallel(\nu) = 0_m$, for $\nu \in [0, \mathbf{T}_s)$, and hence

$$\begin{aligned} \tilde{\mathcal{F}}(z, t) &= \int_0^t e^{\omega \nu} (\Sigma_\nu^\parallel(z) - \hat{\Lambda}^\parallel(\nu) d\nu) \\ &= \int_0^t e^{\omega \nu} \Sigma_\nu^\parallel(z) = \int_0^t e^{\omega \nu} \Lambda_\mu^\parallel(\nu, z(\nu)) d\nu + \int_0^t e^{\omega \nu} F_\sigma^\parallel(\nu, z(\nu)) dW_\nu, \quad \forall t \in [0, \mathbf{T}_s) \cap [0, \tau]. \quad (\text{D.14}) \end{aligned}$$

Similarly, over the interval $[iT_s, (i+1)T_s]$, $i \in \mathbb{N}_{\geq 1}$, the adaptive estimate is given by (D.3) as

$$\hat{\Lambda}^\parallel(\nu) = -\frac{\lambda_s}{1 - e^{\lambda_s T_s}} \Theta_{ad}(iT_s) \int_{(i-1)T_s}^{iT_s} e^{-\lambda_s(iT_s - \beta)} \Sigma_\beta(z),$$

which implies that

$$\int_{iT_s}^t e^{\omega\nu} \hat{\Lambda}^\parallel(\nu) d\nu = -\frac{\lambda_s}{1 - e^{\lambda_s T_s}} \Theta_{ad}(iT_s) \left(\int_{iT_s}^t e^{\omega\nu} d\nu \right) \int_{(i-1)T_s}^{iT_s} e^{-\lambda_s(iT_s - \nu)} \Sigma_\nu(z),$$

for all $(t, i) \in ([iT_s, (i+1)T_s] \cap [0, \tau]) \times \mathbb{N}_{\geq 1}$, where note that we have replaced β with ν as the integration variable in the last term. Since

$$\begin{aligned} -\frac{\lambda_s}{1 - e^{\lambda_s T_s}} \int_{(i-1)T_s}^{iT_s} e^{-\lambda_s(iT_s - \nu)} \Sigma_\nu(z) d\nu &= -\frac{\lambda_s e^{-\lambda_s iT_s}}{1 - e^{\lambda_s T_s}} \int_{(i-1)T_s}^{iT_s} e^{\lambda_s \nu} \Sigma_\nu(z) d\nu \\ &= \left(\int_{iT_s}^{(i+1)T_s} e^{\lambda_s \nu} d\nu \right)^{-1} \int_{(i-1)T_s}^{iT_s} e^{\lambda_s \nu} \Sigma_\nu(z) d\nu, \end{aligned}$$

the expression above the last can be written as

$$\int_{iT_s}^t e^{\omega\nu} \hat{\Lambda}^\parallel(\nu) d\nu = \frac{\int_{iT_s}^t e^{\omega\nu} d\nu}{\int_{(i+1)T_s}^{iT_s} e^{\lambda_s \nu} d\nu} \Theta_{ad}(iT_s) \int_{(i-1)T_s}^{iT_s} e^{\lambda_s \nu} \Sigma_\nu(z) d\nu, \quad \forall (t, i) \in ([iT_s, (i+1)T_s] \cap [0, \tau]) \times \mathbb{N}_{\geq 1}.$$

It then follows from (D.9), Proposition D.2, that

$$\begin{aligned} \int_{iT_s}^t e^{\omega\nu} \hat{\Lambda}^\parallel(\nu) d\nu &= \frac{\int_{iT_s}^t e^{\omega\nu} d\nu}{\int_{(i+1)T_s}^{iT_s} e^{\lambda_s \nu} d\nu} \Theta_{ad}(iT_s) \int_{(i-1)T_s}^{iT_s} e^{\lambda_s \nu} \tilde{\Sigma}_\nu(iT_s, z) d\nu \\ &\quad + \frac{\int_{iT_s}^t e^{\omega\nu} d\nu}{\int_{(i+1)T_s}^{iT_s} e^{\lambda_s \nu} d\nu} \int_{(i-1)T_s}^{iT_s} e^{\lambda_s \nu} \Lambda_\mu^\parallel(iT_s, z(iT_s)) d\nu \\ &\quad + \frac{\int_{iT_s}^t e^{\omega\nu} d\nu}{\int_{(i+1)T_s}^{iT_s} e^{\lambda_s \nu} d\nu} \int_{(i-1)T_s}^{iT_s} e^{\lambda_s \nu} F_\sigma^\parallel(iT_s, z(iT_s)) dW_\nu, \end{aligned}$$

for all $(t, i) \in ([iT_s, (i+1)T_s] \cap [0, \tau]) \times \mathbb{N}_{\geq 1}$. Additionally, we have from (D.10), Proposition D.2, that

$$\int_{iT_s}^t e^{\omega\nu} \Sigma_\nu^\parallel(z) d\nu = \int_{iT_s}^t e^{\omega\nu} \tilde{\Sigma}_\nu^\parallel(iT_s, z) d\nu + \int_{iT_s}^t e^{\omega\nu} [\Lambda_\mu^\parallel(iT_s, z(iT_s)) d\nu + F_\sigma^\parallel(iT_s, z(iT_s)) dW_\nu],$$

for all $(t, i) \in ([iT_s, (i+1)T_s] \cap [0, \tau]) \times \mathbb{N}_{\geq 1}$. Hence, we use the two expressions above and conclude that

$$\begin{aligned} \tilde{\mathcal{F}}(z, t) &= \int_{iT_s}^t e^{\omega\nu} \tilde{\Sigma}_\nu^\parallel(iT_s, z) d\nu - \frac{\int_{iT_s}^t e^{\omega\nu} d\nu}{\int_{(i+1)T_s}^{iT_s} e^{\lambda_s \nu} d\nu} \Theta_{ad}(iT_s) \int_{(i-1)T_s}^{iT_s} e^{\lambda_s \nu} \tilde{\Sigma}_\nu(iT_s, z) d\nu \\ &\quad + \left(\int_{iT_s}^t e^{\omega\nu} d\nu - \frac{\int_{iT_s}^t e^{\omega\nu} d\nu}{\int_{(i+1)T_s}^{iT_s} e^{\lambda_s \nu} d\nu} \int_{(i-1)T_s}^{iT_s} e^{\lambda_s \nu} d\nu \right) \Lambda_\mu^\parallel(iT_s, z(iT_s)) \\ &\quad + F_\sigma^\parallel(iT_s, z(iT_s)) \left(\int_{iT_s}^t e^{\omega\nu} dW_\nu - \frac{\int_{iT_s}^t e^{\omega\nu} d\nu}{\int_{(i+1)T_s}^{iT_s} e^{\lambda_s \nu} d\nu} \int_{(i-1)T_s}^{iT_s} e^{\lambda_s \nu} dW_\nu \right), \quad (\text{D.15}) \end{aligned}$$

for all $(t, i) \in ([iT_s, (i+1)T_s] \cap [0, \tau]) \times \mathbb{N}_{\geq 1}$.

Using (D.14) and (D.15), we can write $\tilde{\mathcal{F}}(z, t)$ as

$$\tilde{\mathcal{F}}(z, t) = \mathbb{1}_{<T_s}(t) \int_0^t e^{\omega\nu} \Sigma_\nu^\parallel(z) d\nu + \mathbb{1}_{\geq T_s}(t) \int_0^{T_s} e^{\omega\nu} \Lambda_\mu^\parallel(\nu, z(\nu)) d\nu$$

$$+ \mathbb{1}_{\geq \mathbf{T}_s}(t) \left(\tilde{\mathcal{F}}_o(z, t) + \hat{\mathcal{F}}_\mu(z, t) + \hat{\mathcal{F}}_\sigma(z, t) \right), \quad t \in [0, \tau], \quad (\text{D.16})$$

where we have defined the following for $t \in [\mathbf{T}_s, \infty) \cap [0, \tau]$:

$$\begin{aligned} \tilde{\mathcal{F}}_o(z, t) &= \sum_{i=1}^{i^*(t)-1} \left(\int_{i\mathbf{T}_s}^{(i+1)\mathbf{T}_s} e^{\omega\nu} \tilde{\Sigma}_\nu^\parallel(i\mathbf{T}_s, z) - \tilde{\gamma}_i(\mathbf{T}_s) \Theta_{ad}(i\mathbf{T}_s) \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s \nu} \tilde{\Sigma}_\nu(i\mathbf{T}_s, z) \right) \\ &\quad + \int_{i^*(t)\mathbf{T}_s}^t e^{\omega\nu} \tilde{\Sigma}_\nu^\parallel(i^*(t)\mathbf{T}_s, z) - \tilde{\gamma}^*(t, \mathbf{T}_s) \Theta_{ad}(i^*(t)\mathbf{T}_s) \int_{(i^*(t)-1)\mathbf{T}_s}^{i^*(t)\mathbf{T}_s} e^{\lambda_s \nu} \tilde{\Sigma}_\nu(i^*(t)\mathbf{T}_s, z), \\ \hat{\mathcal{F}}_\mu(z, t) &= \sum_{i=1}^{i^*(t)-1} \left(\int_{i\mathbf{T}_s}^{(i+1)\mathbf{T}_s} e^{\omega\nu} d\nu - \tilde{\gamma}_i(\mathbf{T}_s) \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s \nu} d\nu \right) \Lambda_\mu^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) \\ &\quad + \left(\int_{i^*(t)\mathbf{T}_s}^t e^{\omega\nu} d\nu - \tilde{\gamma}^*(t, \mathbf{T}_s) \int_{(i^*(t)-1)\mathbf{T}_s}^{i^*(t)\mathbf{T}_s} e^{\lambda_s \nu} d\nu \right) \Lambda_\mu^\parallel(i^*(t)\mathbf{T}_s, z(i^*(t)\mathbf{T}_s)), \\ \hat{\mathcal{F}}_\sigma(z, t) &= \int_0^{\mathbf{T}_s} e^{\omega\nu} F_\sigma^\parallel(\nu, z(\nu)) dW_\nu + \sum_{i=1}^{i^*(t)-1} F_\sigma^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) \left(\int_{i\mathbf{T}_s}^{(i+1)\mathbf{T}_s} e^{\omega\nu} dW_\nu - \tilde{\gamma}_i(\mathbf{T}_s) \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s \nu} dW_\nu \right) \\ &\quad + F_\sigma^\parallel(i^*(t)\mathbf{T}_s, z(i^*(t)\mathbf{T}_s)) \left(\int_{i^*(t)\mathbf{T}_s}^t e^{\omega\nu} dW_\nu - \tilde{\gamma}^*(t, \mathbf{T}_s) \int_{(i^*(t)-1)\mathbf{T}_s}^{i^*(t)\mathbf{T}_s} e^{\lambda_s \nu} dW_\nu \right). \end{aligned}$$

Now, we re-write $\tilde{\mathcal{F}}_o$ as

$$\begin{aligned} \tilde{\mathcal{F}}_o(z, t) &= \int_{\mathbf{T}_s}^{i^*(t)\mathbf{T}_s} e^{\omega\nu} \left(\sum_{i=1}^{i^*(t)-1} \mathbb{1}_{[i\mathbf{T}_s, (i+1)\mathbf{T}_s)}(\nu) \tilde{\Sigma}_\nu^\parallel(i\mathbf{T}_s, z) \right) + \int_{i^*(t)\mathbf{T}_s}^t e^{\omega\nu} \tilde{\Sigma}_\nu^\parallel(i^*(t)\mathbf{T}_s, z) \\ &\quad - \int_0^{(i^*(t)-1)\mathbf{T}_s} e^{\lambda_s \nu} \left(\sum_{i=1}^{i^*(t)-1} \mathbb{1}_{[(i-1)\mathbf{T}_s, i\mathbf{T}_s)}(\nu) \tilde{\gamma}_i(\mathbf{T}_s) \Theta_{ad}(i\mathbf{T}_s) \tilde{\Sigma}_\nu(i\mathbf{T}_s, z) \right) \\ &\quad - \int_{(i^*(t)-1)\mathbf{T}_s}^{i^*(t)\mathbf{T}_s} e^{\lambda_s \nu} \tilde{\gamma}^*(t, \mathbf{T}_s) \Theta_{ad}(i^*(t)\mathbf{T}_s) \tilde{\Sigma}_\nu(i^*(t)\mathbf{T}_s, z), \end{aligned}$$

for $t \in [\mathbf{T}_s, \infty) \cap [0, \tau]$. Thus, using the definition of $\tilde{\Sigma}_t^{\{\cdot, \parallel\}}$ ($i\mathbf{T}_s, z$) in (D.11), we can further express $\tilde{\mathcal{F}}_o$ as

$$\tilde{\mathcal{F}}_o(z, t) = \int_0^t e^{\omega\nu} \left[\left(\tilde{\mathcal{H}}_\mu^\parallel(\nu, z) + e^{(\lambda_s - \omega)\nu} \tilde{\mathcal{H}}_\mu(\nu, z) \right) d\nu + \left(\tilde{\mathcal{H}}_\sigma^\parallel(\nu, z) + e^{(\lambda_s - \omega)\nu} \tilde{\mathcal{H}}_\sigma(\nu, z) \right) dW_\nu \right], \quad (\text{D.17})$$

for $t \in [\mathbf{T}_s, \infty) \cap [0, \tau]$.

We next manipulate the terms $\hat{\mathcal{F}}_\mu$ and $\hat{\mathcal{F}}_\sigma$, and start by observing that

$$\begin{aligned} \int_{i\mathbf{T}_s}^{(i+1)\mathbf{T}_s} e^{\omega\nu} d\nu - \tilde{\gamma}_i(\mathbf{T}_s) \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s \nu} d\nu &= \left(\int_{i\mathbf{T}_s}^{(i+1)\mathbf{T}_s} e^{\omega\nu} d\nu \right) \left(1 - \left(\int_{i\mathbf{T}_s}^{(i+1)\mathbf{T}_s} e^{\lambda_s \nu} d\nu \right)^{-1} \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s \nu} d\nu \right) \\ &= \left(\int_{i\mathbf{T}_s}^{(i+1)\mathbf{T}_s} e^{\omega\nu} d\nu \right) (1 - e^{-\lambda_s \mathbf{T}_s}), \end{aligned}$$

and, similarly

$$\int_{i^*(t)\mathbf{T}_s}^t e^{\omega\nu} d\nu - \tilde{\gamma}^*(t, \mathbf{T}_s) \int_{(i^*(t)-1)\mathbf{T}_s}^{i^*(t)\mathbf{T}_s} e^{\lambda_s \nu} d\nu = \left(\int_{i^*(t)\mathbf{T}_s}^t e^{\omega\nu} d\nu \right) (1 - e^{-\lambda_s \mathbf{T}_s}).$$

Hence, $\hat{\mathcal{F}}_\mu$, defined in (D.16), can be written as

$$\begin{aligned}\hat{\mathcal{F}}_\mu(z, t) = & \sum_{i=1}^{i^*(t)-1} \left(\int_{i\mathbf{T}_s}^{(i+1)\mathbf{T}_s} e^{\omega\nu} d\nu \right) (1 - e^{-\lambda_s \mathbf{T}_s}) \Lambda_\mu^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) \\ & + \left(\int_{i^*(t)\mathbf{T}_s}^t e^{\omega\nu} d\nu \right) (1 - e^{-\lambda_s \mathbf{T}_s}) \Lambda_\mu^\parallel(i^*(t)\mathbf{T}_s, z(i^*(t)\mathbf{T}_s)),\end{aligned}$$

for $t \in [\mathbf{T}_s, \infty) \cap [0, \tau]$, which we can further write as

$$\hat{\mathcal{F}}_\mu(z, t) = \int_0^t e^{\omega\nu} \mathcal{G}_\mu(\nu, z) d\nu, \quad t \in [\mathbf{T}_s, \infty) \cap [0, \tau]. \quad (\text{D.18})$$

Next, by adding and subtracting $F_\sigma^\parallel(0, z(0))$, we obtain

$$\begin{aligned}\int_0^{\mathbf{T}_s} e^{\omega\nu} F_\sigma^\parallel(\nu, z(\nu)) dW_\nu &= \int_0^{\mathbf{T}_s} e^{\omega\nu} (F_\sigma^\parallel(\nu, z(\nu)) - F_\sigma^\parallel(0, z(0))) dW_\nu + F_\sigma^\parallel(0, z(0)) \int_0^{\mathbf{T}_s} e^{\omega\nu} dW_\nu \\ &= \int_0^{\mathbf{T}_s} e^{\omega\nu} \tilde{F}_\sigma^\parallel(\nu, 0, z) dW_\nu + F_\sigma^\parallel(0, z(0)) \int_0^{\mathbf{T}_s} e^{\omega\nu} dW_\nu.\end{aligned} \quad (\text{D.19})$$

Furthermore, we obtain the following by adding and subtracting $F_\sigma^\parallel((i-1)\mathbf{T}_s, z((i-1)\mathbf{T}_s))$:

$$\begin{aligned}F_\sigma^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) &= (F_\sigma^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) - F_\sigma^\parallel((i-1)\mathbf{T}_s, z((i-1)\mathbf{T}_s))) + F_\sigma^\parallel((i-1)\mathbf{T}_s, z((i-1)\mathbf{T}_s)) \\ &= \tilde{F}_\sigma^\parallel(i\mathbf{T}_s, (i-1)\mathbf{T}_s, z) + F_\sigma^\parallel((i-1)\mathbf{T}_s, z((i-1)\mathbf{T}_s)).\end{aligned}$$

Using this expression, we may write

$$\begin{aligned}F_\sigma^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) \tilde{\gamma}_i(\mathbf{T}_s) \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s \nu} dW_\nu &= \tilde{F}_\sigma^\parallel(i\mathbf{T}_s, (i-1)\mathbf{T}_s, z) \tilde{\gamma}_i(\mathbf{T}_s) \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s \nu} dW_\nu \\ &\quad + F_\sigma^\parallel((i-1)\mathbf{T}_s, z((i-1)\mathbf{T}_s)) \tilde{\gamma}_i(\mathbf{T}_s) \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s \nu} dW_\nu,\end{aligned}$$

which allows us to further write

$$\begin{aligned}& \sum_{i=1}^{i^*(t)-1} F_\sigma^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) \left(\int_{i\mathbf{T}_s}^{(i+1)\mathbf{T}_s} e^{\omega\nu} dW_\nu - \tilde{\gamma}_i(\mathbf{T}_s) \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s \nu} dW_\nu \right) \\ &= \sum_{i=1}^{i^*(t)-1} \left(F_\sigma^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) \int_{i\mathbf{T}_s}^{(i+1)\mathbf{T}_s} e^{\omega\nu} dW_\nu - F_\sigma^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) \tilde{\gamma}_i(\mathbf{T}_s) \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s \nu} dW_\nu \right) \\ &= \sum_{i=1}^{i^*(t)-1} \left(F_\sigma^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) \int_{i\mathbf{T}_s}^{(i+1)\mathbf{T}_s} e^{\omega\nu} dW_\nu - F_\sigma^\parallel((i-1)\mathbf{T}_s, z((i-1)\mathbf{T}_s)) \tilde{\gamma}_i(\mathbf{T}_s) \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s \nu} dW_\nu \right) \\ &\quad - \sum_{i=1}^{i^*(t)-1} \tilde{F}_\sigma^\parallel(i\mathbf{T}_s, (i-1)\mathbf{T}_s, z) \tilde{\gamma}_i(\mathbf{T}_s) \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s \nu} dW_\nu. \quad (\text{D.20})\end{aligned}$$

Then, adding (D.19) and (D.20), we obtain

$$\begin{aligned}& \int_0^{\mathbf{T}_s} e^{\omega\nu} F_\sigma^\parallel(\nu, z(\nu)) dW_\nu + \sum_{i=1}^{i^*(t)-1} F_\sigma^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) \left(\int_{i\mathbf{T}_s}^{(i+1)\mathbf{T}_s} e^{\omega\nu} dW_\nu - \tilde{\gamma}_i(\mathbf{T}_s) \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s \nu} dW_\nu \right) \\ &= \int_0^{\mathbf{T}_s} e^{\omega\nu} \tilde{F}_\sigma^\parallel(\nu, 0, z) dW_\nu - \sum_{i=1}^{i^*(t)-1} \tilde{F}_\sigma^\parallel(i\mathbf{T}_s, (i-1)\mathbf{T}_s, z) \tilde{\gamma}_i(\mathbf{T}_s) \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s \nu} dW_\nu \\ &\quad + \sum_{i=1}^{i^*(t)-1} F_\sigma^\parallel((i-1)\mathbf{T}_s, z((i-1)\mathbf{T}_s)) \left(\int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\omega\nu} [1 - \tilde{\gamma}_i(\mathbf{T}_s) e^{(\lambda_s - \omega)\nu}] dW_\nu \right)\end{aligned}$$

$$+ F_\sigma^\parallel((i^*(t) - 1)\mathbf{T}_s, z((i^*(t) - 1)\mathbf{T}_s)) \int_{(i^*(t)-1)\mathbf{T}_s}^{i^*(t)\mathbf{T}_s} e^{\omega\nu} dW_\nu, \quad (\text{D.21})$$

where we have used the manipulation

$$\begin{aligned} F_\sigma^\parallel(0, z(0)) \int_0^{\mathbf{T}_s} e^{\omega\nu} dW_\nu + \sum_{i=1}^{i^*(t)-1} F_\sigma^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) \int_{i\mathbf{T}_s}^{(i+1)\mathbf{T}_s} e^{\omega\nu} dW_\nu \\ = \sum_{i=1}^{i^*(t)-1} F_\sigma^\parallel((i-1)\mathbf{T}_s, z((i-1)\mathbf{T}_s)) \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\omega\nu} dW_\nu \\ + F_\sigma^\parallel((i^*(t) - 1)\mathbf{T}_s, z((i^*(t) - 1)\mathbf{T}_s)) \int_{(i^*(t)-1)\mathbf{T}_s}^{i^*(t)\mathbf{T}_s} e^{\omega\nu} dW_\nu, \end{aligned}$$

which allows us to write

$$\begin{aligned} F_\sigma^\parallel(0, z(0)) \int_0^{\mathbf{T}_s} e^{\omega\nu} dW_\nu + \sum_{i=1}^{i^*(t)-1} F_\sigma^\parallel(i\mathbf{T}_s, z(i\mathbf{T}_s)) \int_{i\mathbf{T}_s}^{(i+1)\mathbf{T}_s} e^{\omega\nu} dW_\nu \\ - \sum_{i=1}^{i^*(t)-1} F_\sigma^\parallel((i-1)\mathbf{T}_s, z((i-1)\mathbf{T}_s)) \tilde{\gamma}_i(\mathbf{T}_s) \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s\nu} dW_\nu \\ = \sum_{i=1}^{i^*(t)-1} F_\sigma^\parallel((i-1)\mathbf{T}_s, z((i-1)\mathbf{T}_s)) \left(\int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\omega\nu} [1 - \tilde{\gamma}_i(\mathbf{T}_s) e^{(\lambda_s - \omega)\nu}] dW_\nu \right) \\ + F_\sigma^\parallel((i^*(t) - 1)\mathbf{T}_s, z((i^*(t) - 1)\mathbf{T}_s)) \int_{(i^*(t)-1)\mathbf{T}_s}^{i^*(t)\mathbf{T}_s} e^{\omega\nu} dW_\nu. \end{aligned}$$

Substituting (D.21) for the first two terms in the definition of $\hat{\mathcal{F}}_\sigma$ in (D.16), we obtain

$$\begin{aligned} \hat{\mathcal{F}}_\sigma(z, t) = \int_0^{\mathbf{T}_s} e^{\omega\nu} \tilde{F}_\sigma^\parallel(\nu, 0, z) dW_\nu - \sum_{i=1}^{i^*(t)-1} \tilde{F}_\sigma^\parallel(i\mathbf{T}_s, (i-1)\mathbf{T}_s, z) \tilde{\gamma}_i(\mathbf{T}_s) \int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\lambda_s\nu} dW_\nu \\ + \sum_{i=1}^{i^*(t)-1} F_\sigma^\parallel((i-1)\mathbf{T}_s, z((i-1)\mathbf{T}_s)) \left(\int_{(i-1)\mathbf{T}_s}^{i\mathbf{T}_s} e^{\omega\nu} [1 - \tilde{\gamma}_i(\mathbf{T}_s) e^{(\lambda_s - \omega)\nu}] dW_\nu \right) \\ + F_\sigma^\parallel((i^*(t) - 1)\mathbf{T}_s, z((i^*(t) - 1)\mathbf{T}_s)) \int_{(i^*(t)-1)\mathbf{T}_s}^{i^*(t)\mathbf{T}_s} e^{\omega\nu} dW_\nu \\ + F_\sigma^\parallel(i^*(t)\mathbf{T}_s, z(i^*(t)\mathbf{T}_s)) \left(\int_{i^*(t)\mathbf{T}_s}^t e^{\omega\nu} dW_\nu - \tilde{\gamma}^*(t, \mathbf{T}_s) \int_{(i^*(t)-1)\mathbf{T}_s}^{i^*(t)\mathbf{T}_s} e^{\lambda_s\nu} dW_\nu \right), \end{aligned}$$

for $t \in [\mathbf{T}_s, \infty) \cap [0, \tau]$. We can therefore re-write the previous expression as

$$\begin{aligned} \hat{\mathcal{F}}_\sigma(z, t) = \int_0^t e^{\omega\nu} \left(\tilde{\mathcal{G}}_{\sigma_1}(\nu, z) + e^{(\lambda_s - \omega)\nu} \tilde{\mathcal{G}}_{\sigma_2}(\nu, z) \right) dW_\nu \\ + \int_0^t e^{\omega\nu} (\mathcal{G}_{\sigma_1}(\nu, z) + \mathcal{G}_{\sigma_2}(\nu, z)) dW_\nu, \quad t \in [\mathbf{T}_s, \infty) \cap [0, \tau]. \quad (\text{D.22}) \end{aligned}$$

Adding the derived expressions for $\tilde{\mathcal{F}}_\sigma$, $\hat{\mathcal{F}}_\mu$, and $\hat{\mathcal{F}}_\sigma$ in (D.17), (D.18) and (D.22), respectively, leads to

$$\begin{aligned} \tilde{\mathcal{F}}_\sigma(z, t) + \hat{\mathcal{F}}_\mu(z, t) + \hat{\mathcal{F}}_\sigma(z, t) \\ = \int_0^t e^{\omega\nu} \left[\left(\tilde{\mathcal{H}}_\mu^\parallel(\nu, z) + e^{(\lambda_s - \omega)\nu} \tilde{\mathcal{H}}_\mu(\nu, z) \right) d\nu + \left(\tilde{\mathcal{H}}_\sigma^\parallel(\nu, z) + e^{(\lambda_s - \omega)\nu} \tilde{\mathcal{H}}_\sigma(\nu, z) \right) dW_\nu \right] \\ + \int_0^t e^{\omega\nu} \left(\tilde{\mathcal{G}}_{\sigma_1}(\nu, z) + e^{(\lambda_s - \omega)\nu} \tilde{\mathcal{G}}_{\sigma_2}(\nu, z) \right) dW_\nu \end{aligned}$$

$$+ \int_0^t e^{\omega \nu} [\mathcal{G}_\mu(\nu, z) d\nu + (\mathcal{G}_{\sigma_1}(\nu, z) + \mathcal{G}_{\sigma_2}(\nu, z)) dW_\nu], \quad t \in [\mathbf{T}_s, \infty) \cap [0, \tau].$$

Substituting the above into (D.16) gives us the desired result in (D.13). \square

Next, we provide the proof of Proposition 4.3.

Proof of Proposition 4.3. It follows from Definitions 8 and 9 that we need to show the well posedness of

$$\begin{aligned} dZ_{N,t} &= dX_{N,t} - dX_{N,t}^r \\ &= [F_{N,\mu}(t, X_{N,t}, U_{\mathcal{L}_1,t}) - F_{N,\mu}(t, X_{N,t}^r, U_t^r)] dt + [F_{N,\sigma}(t, X_{N,t}) - F_{N,\sigma}(t, X_{N,t}^r)] dW_t \\ &= [F_{N,\mu}(t, X_{N,t}, U_{\mathcal{L}_1,t}) - F_{N,\mu}(t, X_{N,t}^r, U_t^r)] dt + \dot{F}_{N,\sigma}(t, Z_{N,t}) dW_t, \end{aligned} \quad (\text{D.23})$$

for $t \in [0, T]$, where $Z_{N,0} = 0$, $U_{\mathcal{L}_1} = \mathcal{F}_{\mathcal{L}_1}(X_N)$, $U^r = \mathcal{F}_r(X_N^r)$, the functions $F_{N,\{\mu,\sigma\}}$ are defined analogously to $J_{N,\{\mu,\sigma\}}$ in (99), and we have set

$$\dot{F}_{N,\sigma}(t, Z_{N,t}) \doteq F_{N,\sigma}(t, X_{N,t}) - F_{N,\sigma}(t, X_{N,t}^r).$$

Let us denote by f_N , $\Lambda_{N,\mu}$, p_N , and $\Lambda_{N,\sigma}$ the truncated versions of the functions f , Λ_μ , p , and Λ_σ , respectively, and where the truncation is defined as in (99). Then, by adding and subtracting $g(t)\mathcal{F}_r(X_N)(t)$, along with Definition 1, we get

$$\begin{aligned} F_{N,\mu}(t, X_{N,t}, U_{\mathcal{L}_1,t}) - F_{N,\mu}(t, X_{N,t}^r, U_t^r) &= F_{N,\mu}(t, X_{N,t}, \mathcal{F}_{\mathcal{L}_1}(X_N)(t)) - F_{N,\mu}(t, X_{N,t}^r, \mathcal{F}_r(X_N^r)(t)) \\ &= \dot{F}_{N,\mu}(t, Z_{N,t}) + g(t)[(\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r)(X_{N,t})], \end{aligned}$$

where we have defined

$$\dot{F}_{N,\mu}(t, Z_{N,t}) \doteq F_{N,\mu}(t, X_{N,t}, \mathcal{F}_r(X_N)(t)) - F_{N,\mu}(t, X_{N,t}^r, \mathcal{F}_r(X_N^r)(t)).$$

Substituting the above into (D.23)

$$dZ_{N,t} = \dot{F}_{N,\mu}(t, Z_{N,t}) dt + \dot{F}_{N,\sigma}(t, Z_{N,t}) dW_t + g(t)[(\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r)(X_{N,t})] dt, \quad Z_{N,0} = 0,$$

for $t \in [0, T]$. Now, consider any $z \in \mathcal{M}_2([0, T], \mathbb{R}^n \mid \mathfrak{W}_{0,t})$, for any $t \in [0, T]$, and define

$$M(z(t)) = M_1(z(t)) + M_2(z(t)), \quad t \in [0, T],$$

where

$$\begin{aligned} M_1(z(t)) &= \int_0^t \dot{F}_{N,\mu}(\nu, Z_{N,\nu}) d\nu + \int_0^\nu \dot{F}_{N,\sigma}(\nu, Z_{N,\nu}) dW_\nu, \\ M_2(z(t)) &= \int_0^t g(\nu)[(\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r)(X_{N,\nu})] d\nu. \end{aligned}$$

We can then conclude the well-posedness by following the proof of Proposition 4.1 for $M_1(z(t))$, and deriving similar conclusions for $M_2(z(t))$ by invoking Proposition D.1. \square

The next two results help us with the computation of the stochastic differential of $\mathcal{V}(Z_{N,t})$ in the proof of Lemma 4.5. **Proposition D.4** Let $Z_{N,t}$ be the strong solution of (99), and let $\tau(t)$ be the stopping time defined in (104), Lemma 4.5. Then,

$$\begin{aligned} &\int_0^{\tau(t)} e^{2\lambda \nu} \nabla \mathcal{V}(Z_{N,\nu})^\top J_\sigma(\nu, Z_{N,\nu}) dW_\nu \\ &+ \int_0^{\tau(t)} e^{2\lambda \nu} \left(\nabla \mathcal{V}(Z_{N,\nu})^\top J_\mu(\nu, Z_{N,\nu}) + \frac{1}{2} \text{Tr}[K_\sigma(\nu, Z_{N,\nu}) \nabla^2 \mathcal{V}(Z_{N,\nu})] \right) d\nu \\ &\leq -2\lambda \int_0^{\tau(t)} e^{2\lambda \nu} \mathcal{V}(Z_{N,\nu}) d\nu + \int_0^{\tau(t)} e^{2\lambda \nu} \left([\phi_\mu(\nu, Z_{N,\nu}) + \tilde{\mathcal{U}}(\nu, Z_{N,\nu})] d\nu + \phi_\sigma(\nu, Z_{N,\nu}) dW_\nu \right) \\ &\quad + \int_0^{\tau(t)} e^{2\lambda \nu} ([\phi_{\mu^\parallel}(\nu, Z_{N,\nu}) + \phi_U(\nu, Z_{N,\nu})] d\nu + \phi_{\sigma^\parallel}(\nu, Z_{N,\nu}) dW_\nu), \end{aligned} \quad (\text{D.24})$$

for all $t \in \mathbb{R}_{\geq 0}$, where $K_\sigma(\nu, Z_{N,\nu}) = J_\sigma(\nu, Z_{N,\nu}) J_\sigma(\nu, Z_{N,\nu})^\top$, $J_\mu(\nu, Z_{N,\nu})$ and $J_\sigma(\nu, Z_{N,\nu})$ are defined in (97), and ϕ_μ , ϕ_μ , and $\tilde{\mathcal{U}}$ are defined in (108). Additionally, we have defined

$$\begin{aligned}\phi_{\mu^\parallel}(\nu, Z_{N,\nu}) &= \partial\mathcal{V}(Z_{N,\nu})^\top g(\nu) \mathring{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}), & \phi_{\sigma^\parallel}(\nu, Z_{N,\nu}) &= \partial\mathcal{V}(Z_{N,\nu})^\top g(\nu) \mathring{F}_\sigma^\parallel(\nu, Z_{N,\nu}), \\ \phi_U(\nu, Z_{N,\nu}) &= \partial\mathcal{V}(Z_{N,\nu})^\top g(\nu) \mathring{\mathcal{F}}_r(Z_N)(\nu),\end{aligned}$$

where $\partial\mathcal{V}(Z_{N,\nu}) = 2(X_{N,\nu} - X_{N,\nu}^r)$, and

$$\begin{aligned}\mathring{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) &= \Lambda_\mu^\parallel(\nu, X_{N,\nu}) - \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r), & \mathring{F}_\sigma^\parallel(\nu, Z_{N,\nu}) &= F_\sigma^\parallel(\nu, X_{N,\nu}) - F_\sigma^\parallel(\nu, X_{N,\nu}^r), \\ \mathring{\mathcal{F}}_r(Z_N)(\nu) &= \mathcal{F}_r(X_N)(\nu) - \mathcal{F}_r(X_N^r)(\nu).\end{aligned}$$

Proof. Using the definitions of J_μ in (97) and the decomposition of F_μ in (3), we have that

$$\begin{aligned}\nabla\mathcal{V}(Z_{N,\nu})^\top J_\mu(\nu, Z_{N,\nu}) &= 2(X_{N,\nu} - X_{N,\nu}^r)^\top (F_\mu(\nu, X_{N,\nu}, U_{\mathcal{L}_1,\nu}) - F_\mu(\nu, X_{N,\nu}^r, U_\nu^r)) \\ &= 2(X_{N,\nu} - X_{N,\nu}^r)^\top (\bar{F}_\mu(\nu, X_{N,\nu}) - \bar{F}_\mu(\nu, X_{N,\nu}^r)) \\ &\quad + 2(X_{N,\nu} - X_{N,\nu}^r)^\top [g(\nu)(U_{\mathcal{L}_1,\nu} - U_\nu^r) + \Lambda_\mu(\nu, X_{N,\nu}) - \Lambda_\mu(\nu, X_{N,\nu}^r)],\end{aligned}$$

for all $\nu \in [0, \tau(t)]$, which, due to Lemma 2.1 reduces to

$$\begin{aligned}\nabla\mathcal{V}(Z_{N,\nu})^\top J_\mu(\nu, Z_{N,\nu}) &\leq -2\lambda\mathcal{V}(Z_{N,\nu}) + \partial\mathcal{V}(Z_{N,\nu})^\top g(\nu)(U_{\mathcal{L}_1,\nu} - U_\nu^r) \\ &\quad + \partial\mathcal{V}(Z_{N,\nu})^\top (\Lambda_\mu(\nu, X_{N,\nu}) - \Lambda_\mu(\nu, X_{N,\nu}^r)),\end{aligned}\quad (\text{D.25})$$

for all $\nu \in [0, \tau(t)]$, where we have used the definition $\partial\mathcal{V}(Z_{N,\nu}) = 2(X_{N,\nu} - X_{N,\nu}^r)$. We develop the expression further by using (7) in Assumption 4 to conclude that

$$\Lambda_\mu(\nu, \cdot) = [g(\nu) \quad g(\nu)^\perp] \begin{bmatrix} \Lambda_\mu^\parallel(\nu, \cdot) \\ \Lambda_\mu^\perp(\nu, \cdot) \end{bmatrix} = g(\nu) \Lambda_\mu^\parallel(\nu, \cdot) + g(\nu)^\perp \Lambda_\mu^\perp(\nu, \cdot),$$

which allows us to write (D.25) as

$$\begin{aligned}\nabla\mathcal{V}(Z_{N,\nu})^\top J_\mu(\nu, Z_{N,\nu}) &\leq -2\lambda\mathcal{V}(Z_{N,\nu}) + \partial\mathcal{V}(Z_{N,\nu})^\top g(\nu)(U_{\mathcal{L}_1,\nu} - U_\nu^r) \\ &\quad + \partial\mathcal{V}(Z_{N,\nu})^\top g(\nu) (\Lambda_\mu^\parallel(\nu, X_{N,\nu}) - \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)) \\ &\quad + \partial\mathcal{V}(Z_{N,\nu})^\top g(\nu)^\perp (\Lambda_\mu^\perp(\nu, X_{N,\nu}) - \Lambda_\mu^\perp(\nu, X_{N,\nu}^r)), \quad \forall \nu \in [0, \tau(t)].\end{aligned}\quad (\text{D.26})$$

For the truncated processes $X_{N,t}$ and $X_{N,t}^r$, we have from (27) and (91) that

$$U_t^r = \mathcal{F}_r(X_N^r)(t), \quad U_{\mathcal{L}_1,t} = \mathcal{F}_{\mathcal{L}_1}(X_N)(t).$$

Thus, by adding and subtracting $\mathcal{F}_r(X_N)(t)$, we obtain

$$\begin{aligned}U_{\mathcal{L}_1,t} - U_t^r &= \mathcal{F}_{\mathcal{L}_1}(X_N)(t) - \mathcal{F}_r(X_N^r)(t) = [\mathcal{F}_{\mathcal{L}_1}(X_N)(t) - \mathcal{F}_r(X_N)(t)] - \mathcal{F}_r(X_N^r)(t) + \mathcal{F}_r(X_N)(t) \\ &= (\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r)(X_N)(t) + \mathring{\mathcal{F}}_r(Z_N)(\nu).\end{aligned}\quad (\text{D.27})$$

Substituting into (D.26) yields

$$\begin{aligned}\nabla\mathcal{V}(Z_{N,\nu})^\top J_\mu(\nu, Z_{N,\nu}) &\leq -2\lambda\mathcal{V}(Z_{N,\nu}) + \tilde{\mathcal{U}}(\nu, Z_{N,\nu}) + \phi_U(\nu, Z_{N,\nu}) + \phi_{\mu^\parallel}(\nu, Z_{N,\nu}) \\ &\quad + \partial\mathcal{V}(Z_{N,\nu})^\top g(\nu)^\perp (\Lambda_\mu^\perp(\nu, X_{N,\nu}) - \Lambda_\mu^\perp(\nu, X_{N,\nu}^r)),\end{aligned}\quad (\text{D.28})$$

for all $\nu \in [0, \tau(t)]$, where

$$\begin{aligned}\phi_{\mu^\parallel}(\nu, Z_{N,\nu}) &= \partial\mathcal{V}(Z_{N,\nu})^\top g(\nu) [\Lambda_\mu^\parallel(\nu, X_{N,\nu}) - \Lambda_\mu^\parallel(\nu, X_{N,\nu}^r)], \\ \phi_U(\nu, Z_{N,\nu}) &= \partial\mathcal{V}(Z_{N,\nu})^\top g(\nu) [\mathcal{F}_r(X_N)(\nu) - \mathcal{F}_r(X_N^r)(\nu)], \\ \tilde{\mathcal{U}}(\nu, Z_{N,\nu}) &= \partial\mathcal{V}(Z_{N,\nu})^\top g(\nu) (\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r)(X_N)(\nu).\end{aligned}$$

Next, using the definition of J_σ in (97), we have that

$$\frac{1}{2} \text{Tr} [K_\sigma(\nu, Z_{N,\nu}) \nabla^2 \mathcal{V}(Z_{N,\nu})] = \text{Tr} [K_\sigma(\nu, Z_{N,\nu})] = \|J_\sigma(\nu, Z_{N,\nu})\|_F^2$$

$$= \|F_\sigma(\nu, X_{N,\nu}) - F_\sigma(\nu, X_{N,\nu}^r)\|_F^2, \quad (\text{D.29})$$

and

$$\nabla \mathcal{V}(Z_{N,\nu})^\top J_\sigma(\nu, Z_{N,\nu}) = \partial \mathcal{V}(Z_{N,\nu})^\top (F_\sigma(\nu, X_{N,\nu}) - F_\sigma(\nu, X_{N,\nu}^r)). \quad (\text{D.30})$$

As before, (7) in Assumption 4 allows us to write

$$F_\sigma(\nu, \cdot) = g(\nu)^\perp F_\sigma^\perp(\nu, \cdot) + g(\nu) F_\sigma^\parallel(\nu, \cdot),$$

which in turn lets us express (D.30) as

$$\begin{aligned} \nabla \mathcal{V}(Z_{N,\nu})^\top J_\sigma(\nu, Z_{N,\nu}) &= \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) (F_\sigma^\parallel(\nu, X_{N,\nu}) - F_\sigma^\parallel(\nu, X_{N,\nu}^r)) \\ &\quad + \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu)^\perp (F_\sigma^\perp(\nu, X_{N,\nu}) - F_\sigma^\perp(\nu, X_{N,\nu}^r)). \end{aligned} \quad (\text{D.31})$$

The expressions in (D.28), (D.29), and (D.31) then lead to the desired result in (D.24). □

In the subsequent proposition, we derive the effect of reference feedback operator \mathcal{F}_r on the truncated joint process $Z_{N,t}$.

Proposition D.5 *Let $Z_{N,t}$ be the strong solution of (99), and let $\tau(t)$ be the stopping time defined in (104), Lemma 4.5. Then, for the term ϕ_U defined in the statement of Proposition D.4, we have that*

$$\begin{aligned} \int_0^{\tau(t)} e^{2\lambda\nu} \phi_U(\nu, Z_{N,\nu}) d\nu &= \int_0^{\tau(t)} (\hat{\mathcal{U}}_\mu(\tau(t), \nu, Z_N; \omega) d\nu + \hat{\mathcal{U}}_\sigma(\tau(t), \nu, Z_N; \omega) dW_\nu) \\ &\quad + \int_0^{\tau(t)} e^{2\lambda\nu} (\phi_{U_\mu}(\nu, Z_{N,\nu}; \omega) d\nu + \phi_{U_\sigma}(\nu, Z_{N,\nu}; \omega) dW_\nu), \quad t \in \mathbb{R}_{\geq 0}, \end{aligned} \quad (\text{D.32})$$

where

$$\begin{aligned} \hat{\mathcal{U}}_\mu(\tau(t), \nu, Z_N; \omega) &= e^{-\omega\tau(t)} \frac{\omega}{2\lambda - \omega} \left(e^{\omega\tau(t)} \mathcal{P}(\tau(t), \nu) - e^{2\lambda\tau(t)} \partial \mathcal{V}(Z_{N,\tau(t)})^\top g(\tau(t)) \right) e^{\omega\nu} \mathring{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}), \\ \hat{\mathcal{U}}_\sigma(\tau(t), \nu, Z_N; \omega) &= e^{-\omega\tau(t)} \frac{\omega}{2\lambda - \omega} \left(e^{\omega\tau(t)} \mathcal{P}(\tau(t), \nu) - e^{2\lambda\tau(t)} \partial \mathcal{V}(Z_{N,\tau(t)})^\top g(\tau(t)) \right) e^{\omega\nu} \mathring{F}_\sigma^\parallel(\nu, Z_{N,\nu}), \\ \phi_{U_\mu}(\nu, Z_{N,\nu}; \omega) &= \frac{\omega}{2\lambda - \omega} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \mathring{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}), \\ \phi_{U_\sigma}(\nu, Z_{N,\nu}; \omega) &= \frac{\omega}{2\lambda - \omega} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \mathring{F}_\sigma^\parallel(\nu, Z_{N,\nu}), \end{aligned}$$

and where $\mathcal{P}(\tau(t), \nu)$ is defined in (111), Lemma 4.5.

Proof. Using the definition of ϕ_U in the statement of Proposition D.4, we have that

$$\begin{aligned} &\int_0^{\tau(t)} e^{2\lambda\nu} \phi_U(\nu, Z_{N,\nu}) d\nu \\ &= \int_0^{\tau(t)} e^{2\lambda\nu} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \mathring{\mathcal{F}}_r(Z_N)(\nu) d\nu \\ &= \int_0^{\tau(t)} e^{2\lambda\nu} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) (\mathcal{F}_r(X_N)(\nu) - \mathcal{F}_r(X_N^r)(\nu)) d\nu \\ &= \int_0^{\tau(t)} e^{2\lambda\nu} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) (\mathcal{F}_\omega(\Lambda_\mu^\parallel(\cdot, X_N))(\nu) - \mathcal{F}_\omega(\Lambda_\mu^\parallel(\cdot, X_N^r))(\nu)) d\nu \\ &\quad + \int_0^{\tau(t)} e^{2\lambda\nu} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) (\mathcal{F}_{N,\omega}(F_\sigma^\parallel(\cdot, X_N), W)(\nu) - \mathcal{F}_{N,\omega}(F_\sigma^\parallel(\cdot, X_N^r), W)(\nu)) d\nu, \end{aligned}$$

for all $t \geq 0$, where we have used the definition of \mathcal{F}_r in (27). Next, using the definitions of \mathcal{F}_ω and $\mathcal{F}_{N,\omega}(\cdot, W)$ in (15a) and (27), respectively, we can re-write the previous expression as

$$\begin{aligned} \int_0^{\tau(t)} e^{2\lambda\nu} \phi_U(\nu, Z_{N,\nu}) d\nu &= \int_0^{\tau(t)} \int_0^\nu \left(-\omega e^{(2\lambda-\omega)\nu} \right) \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \left(e^{\omega\beta} \mathring{\Lambda}_\mu^\parallel(\beta, Z_{N,\beta}) d\beta \right) d\nu \\ &\quad + \int_0^{\tau(t)} \int_0^\nu \left(-\omega e^{(2\lambda-\omega)\nu} \right) \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \left(e^{\omega\beta} \mathring{F}_\sigma^\parallel(\beta, Z_{N,\beta}) dW_\beta \right) d\nu, \end{aligned}$$

for all $t \in \mathbb{R}_{\geq 0}$. Changing the order of integration in the first integral on the right hand side, and applying Lemma B.1 to the second integral:

$$\begin{aligned} &\int_0^{\tau(t)} e^{2\lambda\nu} \phi_U(\nu, Z_{N,\nu}) d\nu \\ &= \int_0^{\tau(t)} \left(- \int_\nu^{\tau(t)} \omega e^{(2\lambda-\omega)\beta} \partial \mathcal{V}(Z_{N,\beta})^\top g(\beta) d\beta \right) e^{\omega\nu} \mathring{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu \\ &\quad + \int_0^{\tau(t)} \left(- \int_\nu^{\tau(t)} \omega e^{(2\lambda-\omega)\beta} \partial \mathcal{V}(Z_{N,\beta})^\top g(\beta) d\beta \right) e^{\omega\nu} \mathring{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu, \quad (\text{D.33}) \end{aligned}$$

for all $t \in \mathbb{R}_{\geq 0}$, where in the first integral, we switch between the variables β and ν after changing the order of integration. We then complete the proof by using an identical line of reasoning in the proof of Proposition C.2 from (C.12) onwards. \square

Similar to Proposition C.3, we derive an alternative representation for the term $\mathcal{P}(\tau(t), \nu)$ next.

Proposition D.6 Recall the expression for $\mathcal{P}(\tau(t), \nu)$ in (111) in the statement of Lemma 4.5 which we restate below:

$$\mathcal{P}(\tau(t), \nu) = \int_\nu^{\tau(t)} e^{(2\lambda-\omega)\beta} d\beta \left[\partial \mathcal{V}(Z_{N,\beta})^\top g(\beta) \right] \in \mathbb{R}^{1 \times m}, \quad 0 \leq \nu \leq \tau(t), \quad (\text{D.34})$$

where the $\tau(t)$ is defined in (104). Then, $\mathcal{P}(\tau(t), \nu)$ admits the following representation:

$$\mathcal{P}(\tau(t), \nu) = \mathcal{P}_\circ(\tau(t), \nu) + \mathcal{P}_{ad}(\tau(t), \nu) + \tilde{\mathcal{P}}(\tau(t), \nu) \in \mathbb{R}^{1 \times m}, \quad 0 \leq \nu \leq \tau(t), \quad t \in \mathbb{R}_{\geq 0}, \quad (\text{D.35})$$

where

$$\mathcal{P}_\circ(\tau(t), \nu) = \int_\nu^{\tau(t)} e^{(2\lambda-\omega)\beta} (\mathcal{P}_\mu(\beta) d\beta + \mathcal{P}_\sigma(\beta) dW_\beta)^\top, \quad (\text{D.36a})$$

$$\mathcal{P}_{ad}(\tau(t), \nu) = \int_\nu^{\tau(t)} e^{(2\lambda-\omega)\beta} \mathcal{P}_U(\beta)^\top d\beta, \quad \tilde{\mathcal{P}}(\tau(t), \nu) = \int_\nu^{\tau(t)} e^{(2\lambda-\omega)\beta} \tilde{\mathcal{P}}_U(\beta)^\top d\beta, \quad (\text{D.36b})$$

and where $\mathcal{P}_\mu(\beta)$, $\mathcal{P}_U(\beta)$, $\tilde{\mathcal{P}}_U(\beta) \in \mathbb{R}^m$ and $\mathcal{P}_\sigma(\beta) \in \mathbb{R}^{m \times d}$ are defined using the expressions in (103) as

$$\begin{aligned} \mathcal{P}_\mu(\beta) &= \dot{g}(\beta)^\top \partial \mathcal{V}(Z_{N,\beta}) + 2g(\beta)^\top \left(\mathring{f}(\beta, Z_{N,\beta}) + \mathring{\Lambda}_\mu(\beta, Z_{N,\beta}) \right), \\ \mathcal{P}_U(\beta) &= 2g(\beta)^\top g(\beta) \mathring{F}_r(Z_N)(\beta), \quad \tilde{\mathcal{P}}_U(\beta) = 2g(\beta)^\top g(\beta) (\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r)(X_N)(\beta), \\ \mathcal{P}_\sigma(\beta) &= 2g(\beta)^\top \mathring{F}_\sigma(\beta, Z_{N,\beta}). \end{aligned}$$

Proof. We closely follow the proof of Proposition C.3 and begin by writing

$$\begin{aligned} \partial \mathcal{V}(Z_{N,\beta})^\top g(\beta) &= [\partial \mathcal{V}(Z_{N,\beta})^\top g_{\cdot,1}(\beta) \quad \cdots \quad \partial \mathcal{V}(Z_{N,\beta})^\top g_{\cdot,m}(\beta)] \\ &= [\sum_{i=1}^n \partial_i \mathcal{V}(Z_{N,\beta}) g_{i,1}(\beta) \quad \cdots \quad \sum_{i=1}^n \partial_i \mathcal{V}(Z_{N,\beta}) g_{i,m}(\beta)] \in \mathbb{R}^{1 \times m}, \quad (\text{D.37}) \end{aligned}$$

where $g_{\cdot,j}(\beta) \in \mathbb{R}^n$ is the j -th column of $g(\beta)$, and

$$\partial_i \mathcal{V}(Z_{N,\beta}) \doteq \frac{\partial}{\partial (X_N - X_N^r)_i} \mathcal{V}(Z_{N,\beta}) = 2(X_{N,\beta} - X_{N,\beta}^r)_i \in \mathbb{R}, \quad i \in \{1, \dots, n\}.$$

Applying Itô's lemma to $\partial_i \mathcal{V}(Z_{N,\beta}) g_{i,j}(\beta) \in \mathbb{R}$, $(i, j) \in \{1, \dots, n\} \times \{1, \dots, m\}$, and using the truncated dynamics in (99) we get

$$d_\beta [\partial_i \mathcal{V}(Z_{N,\beta}) g_{i,j}(\beta)] = \left(\partial_i \mathcal{V}(Z_{N,\beta}) \dot{g}_{i,j}(\beta) + \nabla \partial_i \mathcal{V}(Z_{N,\beta})^\top J_\mu(\beta, Z_{N,\beta}) g_{i,j}(\beta) \right) d\beta \\ + \nabla \partial_i \mathcal{V}(Z_{N,\beta})^\top J_\sigma(\beta, Z_{N,\beta}) g_{i,j}(\beta) dW_\beta, \quad (\text{D.38})$$

where we have used the straightforward fact that $\nabla^2 \partial_i \mathcal{V}(Z_{N,\beta}) = 0 \in \mathbb{S}^n$. We have further replaced $J_{N,\mu}$ and $J_{N,\sigma}$ with J_μ and J_σ because from Proposition 4.3, $Z_{N,\beta}$ is also a strong solution of the joint process (97) for all $\beta \in [\nu, \tau(t)] \subseteq [0, \tau^*] \subseteq [0, \tau_N]$. Observe from (D.37) that

$$d_\beta \left[\partial \mathcal{V}(Z_{N,\beta})^\top g_{\cdot,j}(\beta) \right] = \sum_{i=1}^n d_\beta [\partial_i \mathcal{V}(Z_{N,\beta}) g_{i,j}(\beta)] \in \mathbb{R}, \quad j \in \{1, \dots, m\}.$$

It therefore follows from the expression in (D.38)

$$d_\beta \left[\partial \mathcal{V}(Z_{N,\beta})^\top g_{\cdot,j}(\beta) \right] = \sum_{i=1}^n \left(\partial_i \mathcal{V}(Z_{N,\beta}) \dot{g}_{i,j}(\beta) + \nabla \partial_i \mathcal{V}(Z_{N,\beta})^\top J_\mu(\beta, Z_{N,\beta}) g_{i,j}(\beta) \right) d\beta \\ + \sum_{i=1}^n \nabla \partial_i \mathcal{V}(Z_{N,\beta})^\top J_\sigma(\beta, Z_{N,\beta}) g_{i,j}(\beta) dW_\beta,$$

which can be further written as

$$d_\beta \left[\partial \mathcal{V}(Z_{N,\beta})^\top g_{\cdot,j}(\beta) \right] = \left(\partial \mathcal{V}(Z_{N,\beta})^\top \dot{g}_{\cdot,j}(\beta) + g_{\cdot,j}(\beta)^\top \nabla \partial \mathcal{V}(Z_{N,\beta})^\top J_\mu(\beta, Y_{N,\beta}) \right) d\beta \\ + g_{\cdot,j}(\beta)^\top \nabla \partial \mathcal{V}(Z_{N,\beta})^\top J_\sigma(\beta, Z_{N,\beta}) dW_\beta \in \mathbb{R}, \quad (\text{D.39})$$

for $j \in \{1, \dots, m\}$. Once again, (D.37) implies that

$$d_\beta \left[\partial \mathcal{V}(Z_{N,\beta})^\top g(\beta) \right] = \left[d_\beta \left[\partial \mathcal{V}(Z_{N,\beta})^\top g_{\cdot,1}(\beta) \right] \quad \dots \quad d_\beta \left[\partial \mathcal{V}(Z_{N,\beta})^\top g_{\cdot,m}(\beta) \right] \right] \in \mathbb{R}^{1 \times m}.$$

Thus, we use the expression in (D.39) to write

$$d_\beta \left[\partial \mathcal{V}(Z_{N,\beta})^\top g(\beta) \right] = \left(\partial \mathcal{V}(Z_{N,\beta})^\top \dot{g}(\beta) + \left(g(\beta)^\top \nabla \partial \mathcal{V}(Z_{N,\beta})^\top J_\mu(\beta, Z_{N,\beta}) \right)^\top \right) d\beta \\ + \left(g(\beta)^\top \nabla \partial \mathcal{V}(Z_{N,\beta})^\top J_\sigma(\beta, Z_{N,\beta}) dW_\beta \right)^\top \in \mathbb{R}^{1 \times m}. \quad (\text{D.40})$$

Now, since $Z_N = X_N - X_N^r$, using the definition of the gradient of vector valued functions in Sec. 1.4, we observe that

$$\nabla \partial \mathcal{V}(Z_{N,\beta}) = 2 \left[\nabla \left(X_{N,\beta} - X_{N,\beta}^r \right)_1 \quad \dots \quad \nabla \left(X_{N,\beta} - X_{N,\beta}^r \right)_n \right] = 2 \left[\nabla (Z_{N,\beta})_1 \quad \dots \quad \nabla (Z_{N,\beta})_n \right] \\ = 2 \mathbb{I}_n \in \mathbb{S}^n.$$

Thus, it follows from the definitions of the vector fields in (97) that

$$\nabla \partial \mathcal{V}(Z_{N,\beta})^\top J_\mu(\beta, Z_{N,\beta}) = 2 \left(F_\mu(\beta, X_{N,\beta}, U_{\mathcal{L}_1,\beta}) - F_\mu(\beta, X_{N,\beta}^r, U_\beta^r) \right) \\ \stackrel{(\star)}{=} 2 \hat{f}(\beta, Z_{N,\beta}) + 2g(\beta) (\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r)(X_N)(\beta) + 2g(\beta) \hat{\mathcal{F}}_r(Z_N)(\beta) \\ + 2 \hat{\Lambda}_\mu(\beta, Z_{N,\beta}) \in \mathbb{R}^n, \\ \nabla \partial \mathcal{V}(Z_{N,\beta})^\top J_\sigma(\beta, Z_{N,\beta}) = 2 \left(F_\sigma(\beta, X_{N,\beta}) - F_\sigma(\beta, X_{N,\beta}^r) \right) = 2 \hat{F}_\sigma(\beta, Z_{N,\beta}) \in \mathbb{R}^{n \times d},$$

where we obtain (\star) by using the decomposition from (3) in Definition 1 to obtain

$$F_\mu(\beta, X_{N,\beta}, U_{\mathcal{L}_1,\beta}) - F_\mu(\beta, X_{N,\beta}^r, U_\beta^r) \\ = \bar{F}_\mu(\beta, X_{N,\beta}) - \bar{F}_\mu(\beta, X_{N,\beta}^r) + g(\beta) (U_{\mathcal{L}_1,\beta} - U_\beta^r) + \Lambda_\mu(\beta, X_{N,\beta}) - \Lambda_\mu(\beta, X_{N,\beta}^r) \\ = \hat{f}(\beta, Z_{N,\beta}) + g(\beta) (\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r)(X_N)(\beta) + g(\beta) \hat{\mathcal{F}}_r(Z_N)(\beta) + \hat{\Lambda}_\mu(\beta, Z_{N,\beta}),$$

and where we have also used (D.27) in the proof of Proposition D.4. Substituting the above identities into (D.40) produces

$$d_\beta \left[\partial \mathcal{V}(Z_{N,\beta})^\top g(\beta) \right] = \left[\mathcal{P}_\mu(\beta) + \mathcal{P}_\mathcal{U}(\beta) + \tilde{\mathcal{P}}_\mathcal{U}(\beta) \right]^\top d\beta + [\mathcal{P}_\sigma(\beta) dW_\beta]^\top \in \mathbb{R}^{1 \times m}.$$

Then, (D.35) is established by substituting the above into (D.34). □

The next result establishes the bounds for the pertinent entities derived in the last proposition.

Proposition D.7 Consider the functions $\mathcal{P}_\mu(t) \in \mathbb{R}^m$ and $\mathcal{P}_\sigma(t) \in \mathbb{R}^{m \times d}$ defined in the statement of Proposition D.6. If the stopping time τ^* , defined in (104), Lemma 4.5, satisfies $\tau^* = t^*$, then

$$\sup_{\nu \in [0, t^*]} \|\mathcal{P}_\mu(\nu)\|_{L_{2p}} \leq 2\Delta_g \Delta_{\mathcal{P}_\mu}(\mathbf{p}) + \Delta_{\dot{g}} \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}}, \quad (\text{D.41a})$$

$$\sup_{\nu \in [0, t^*]} \|\mathcal{P}_\sigma(\nu)\|_{L_q} \leq 2\Delta_g \Delta_{\mathcal{P}_\sigma}(\mathbf{p}), \quad (\text{D.41b})$$

where

$$\Delta_{\mathcal{P}_\mu}(\mathbf{p}) = \sup_{\nu \in [0, t^*]} \left(\left\| \dot{f}(\nu, Z_{N,\nu}) \right\|_{L_{2p}} + \left\| \dot{\Lambda}_\mu(\nu, Z_{N,\nu}) \right\|_{L_{2p}} \right), \quad \Delta_{\mathcal{P}_\sigma}(\mathbf{p}) = \sup_{\nu \in [0, t^*]} \left\| \dot{F}_\sigma(\nu, Z_{N,\nu}) \right\|_{L_{2p}}.$$

Proof. We have the following bound for the term \mathcal{P}_μ defined in (D.36a):

$$\begin{aligned} \|\mathcal{P}_\mu(\nu)\| &\leq \|\dot{g}(\nu)\|_F \|\partial \mathcal{V}(Z_{N,\nu})\| + 2\|g(\nu)\|_F \left(\left\| \dot{f}(\nu, Z_{N,\nu}) \right\| + \left\| \dot{\Lambda}_\mu(\nu, Z_{N,\nu}) \right\| \right) \\ &\leq \Delta_{\dot{g}} \|\partial \mathcal{V}(Z_{N,\nu})\| + 2\Delta_g \left(\left\| \dot{f}(\nu, Z_{N,\nu}) \right\| + \left\| \dot{\Lambda}_\mu(\nu, Z_{N,\nu}) \right\| \right), \quad \forall \nu \in [0, t^*], \end{aligned}$$

where we have used the bounds on $g(t)$ and $\dot{g}(t)$ in Assumption 1. It then follows from the Minkowski's inequality that

$$\|\mathcal{P}_\mu(\nu)\|_{L_{2p}} \leq \Delta_{\dot{g}} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} + 2\Delta_g \left(\left\| \dot{f}(\nu, Z_{N,\nu}) \right\|_{L_{2p}} + \left\| \dot{\Lambda}_\mu(\nu, Z_{N,\nu}) \right\|_{L_{2p}} \right), \quad \forall \nu \in [0, t^*],$$

which in turn implies (D.41a). The bound in (D.41b) follows trivially. □

The next lemma establishes the bound on Ξ .

Lemma D.1 Consider the term $\Xi(\tau(t), Z_N)$ defined in (107a), Lemma 4.5. If the stopping time τ^* , defined in (104), Lemma 4.5, satisfies $\tau^* = t^*$, then:

$$\mathbb{E} \left[e^{-2\lambda t} \Xi(t, Z_N) \right] \leq \frac{\Delta_{\Xi_1}(1, t^*)}{2\lambda}, \quad (\text{D.42a})$$

$$\|e^{-2\lambda t} \Xi(t, Z_N)\|_{L_p} \leq \frac{\Delta_{\Xi_1}(\mathbf{p}, t^*)}{2\lambda} + \frac{\Delta_{\Xi_2}(\mathbf{p}, t^*)}{2\sqrt{\lambda}}, \quad (\text{D.42b})$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 2}$, where

$$\begin{aligned} \Delta_{\Xi_1}(\mathbf{p}, t^*) &= \sup_{\nu \in [0, t^*]} \left(\Delta_g^\perp \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \left\| \dot{\Lambda}_\mu^\perp(\nu, Z_{N,\nu}) \right\|_{L_{2p}} + \left(\left\| \dot{F}_\sigma(\nu, Z_{N,\nu}) \right\|_{L_{2p}} \right)^2 \right), \\ \Delta_{\Xi_2}(\mathbf{p}, t^*) &= \Delta_g^\perp \mathbf{p}(\mathbf{p}) \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \left\| \dot{F}_\sigma^\perp(\nu, Z_{N,\nu}) \right\|_{L_{2p}}. \end{aligned}$$

Proof. Note that $\tau(t) \in [0, t^*]$ since $\tau^* = t^*$. Thus, in the following, we replace $\tau(t)$ with t with the condition that $t \in [0, t^*]$. We then begin by writing

$$e^{-2\lambda t} \Xi(t, Z_N) = e^{-2\lambda t} \int_0^t e^{2\lambda \nu} \phi_\mu(\nu, Z_{N,\nu}) d\nu + e^{-2\lambda t} \int_0^t e^{2\lambda \nu} \phi_\sigma(\nu, Z_{N,\nu}) dW_\nu, \quad t \in [0, t^*]. \quad (\text{D.43})$$

The regularity assumptions and the t -continuity of the strong solutions $X_{N,t}$ and $X_{N,t}^r$ imply that $\phi_\sigma(\nu, Z_{N,\nu}) \in \mathcal{M}_2^{loc}(\mathbb{R}^{1 \times d} | \mathfrak{M}_t)$. Hence, taking the expectation of both sides of (D.43) and using [91, Thm. 1.5.21] produces

$$\begin{aligned} \mathbb{E} [e^{-2\lambda t} \Xi(t, Z_N)] &= e^{-2\lambda t} \mathbb{E} \left[\int_0^t e^{2\lambda \nu} \phi_\mu(\nu, Z_{N,\nu}) d\nu \right] \\ &\leq e^{-2\lambda t} \mathbb{E} \left[\int_0^t e^{2\lambda \nu} |\phi_\mu(\nu, Z_{N,\nu})| d\nu \right], \quad \forall t \in [0, t^*], \end{aligned} \quad (\text{D.44})$$

where we have used the fact that since $\tau^* = t^*$, and t^* is a constant,

$$\mathbb{E} [e^{-2\lambda t}] = e^{-2\lambda t}, \quad \forall t \in [0, t^*].$$

Using the fact above, we further apply the Minkowski's inequality to (D.43) and obtain

$$\begin{aligned} \|e^{-2\lambda t} \Xi(t, Z_N)\|_{L_p} &\leq e^{-2\lambda t} \left\| \int_0^t e^{2\lambda \nu} \phi_\mu(\nu, Z_{N,\nu}) d\nu \right\|_{L_p} \\ &\quad + e^{-2\lambda t} \left\| \int_0^t e^{2\lambda \nu} \phi_\sigma(\nu, Z_{N,\nu}) dW_\nu \right\|_{L_p}, \quad \forall (t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 2}. \end{aligned} \quad (\text{D.45})$$

The regularity assumptions on the vector fields and the t -continuity of the strong solutions $X_{N,t}$ and $X_{N,t}^r$ over $[0, t^*] \subseteq [0, \tau_N]$ imply that $\phi_\mu(\nu, Z_{N,\nu}) \in \mathcal{M}_2^{loc}(\mathbb{R} | \mathfrak{M}_t)$. Thus, we invoke Proposition B.1 for $\{\theta, \xi, n_s, S\} = \{2\lambda, 0, 1, \phi_\mu(\cdot, Z_N)\}$ to obtain

$$\left\| \int_0^t e^{2\lambda \nu} |\phi_\mu(\nu, Z_{N,\nu})| d\nu \right\|_{L_p} \leq \frac{e^{2\lambda t}}{2\lambda} \sup_{\nu \in [0, t^*]} \|\phi_\mu(\nu, Z_{N,\nu})\|_{L_p}, \quad \forall (t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 1},$$

where we have bounded $e^{2\lambda t} - 1$ by $e^{2\lambda t}$. Using the definition of ϕ_μ in (108), we can further bound the above by invoking the Minkowski's and Cauchy-Schwarz inequalities:

$$\begin{aligned} \left\| \int_0^t e^{2\lambda \nu} |\phi_\mu(\nu, Z_{N,\nu})| d\nu \right\|_{L_p} &\leq \frac{e^{2\lambda t}}{2\lambda} \sup_{\nu \in [0, t^*]} \left(\Delta_g^\perp \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \left\| \mathring{\Lambda}_\mu^\perp(\nu, Z_{N,\nu}) \right\|_{L_{2p}} + \left(\left\| \mathring{F}_\sigma^\perp(\nu, Z_{N,\nu}) \right\|_{L_{2p}} \right)^2 \right), \end{aligned} \quad (\text{D.46})$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 1}$, where we have also used the bound on $g(\nu)^\perp$ in Assumption 1.

Next, we have the following from the definition of ϕ_σ in (108):

$$\int_0^t e^{2\lambda \nu} \phi_\sigma(\nu, Z_{N,\nu}) dW_\nu = \int_0^t e^{2\lambda \nu} \left(\partial \mathcal{V}(Z_{N,\nu})^\top g(\nu)^\perp \mathring{F}_\sigma^\perp(\nu, Z_{N,\nu}) \right) dW_\nu.$$

We bound the above by using Lemma B.2 with $\{\theta, \xi, n_l, n_q, Q\} = \{2\lambda, 0, 1, d, W\}$, and

$$L(\nu) = \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu)^\perp \mathring{F}_\sigma^\perp(\nu, Z_{N,\nu}),$$

to obtain

$$\begin{aligned} \left\| \int_0^t e^{2\lambda \nu} \phi_\sigma(\nu, Z_{N,\nu}) dW_\nu \right\|_{L_p} &\leq \mathbf{p}(\mathbf{p}) \frac{e^{2\lambda t}}{2\sqrt{\lambda}} \sup_{\nu \in [0, t^*]} \left\| \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu)^\perp \mathring{F}_\sigma^\perp(\nu, Z_{N,\nu}) \right\|_{L_p} \\ &\leq \mathbf{p}(\mathbf{p}) \Delta_g^\perp \frac{e^{2\lambda t}}{2\sqrt{\lambda}} \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \left\| \mathring{F}_\sigma^\perp(\nu, Z_{N,\nu}) \right\|_{L_{2p}}, \quad \forall (t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 2}, \end{aligned} \quad (\text{D.47})$$

where we have bounded $(e^{4\lambda t} - 1)^{\frac{1}{2}}$ by $e^{2\lambda t}$ and used the Cauchy-Schwarz inequality to obtain the last inequality.

Substituting (D.46) into (D.44) leads to

$$\begin{aligned} & \mathbb{E} [e^{-2\lambda t} \Xi(t, Z_N)] \\ & \leq \frac{1}{2\lambda} \sup_{\nu \in [0, t^*]} \left(\Delta_g^\perp \|\partial \mathcal{V}(Z_{N, \nu})\|_{L_2} \left\| \dot{\Lambda}_\mu^\perp(\nu, Z_{N, \nu}) \right\|_{L_2} + \left(\left\| \dot{F}_\sigma^\perp(\nu, Z_{N, \nu}) \right\|_{L_2} \right)^2 \right), \quad \forall t \in [0, t^*]. \end{aligned} \quad (\text{D.48})$$

Similarly, substituting (D.46) and (D.47) into (D.45) produces

$$\begin{aligned} & \|e^{-2\lambda t} \Xi(t, Z_N)\|_{L_p} \\ & \leq \Delta_g^\perp \frac{\mathfrak{p}(\mathfrak{p})}{2\sqrt{\lambda}} \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N, \nu})\|_{L_{2\mathfrak{p}}} \left\| \dot{F}_\sigma^\perp(\nu, Z_{N, \nu}) \right\|_{L_{2\mathfrak{p}}} \\ & \quad + \frac{1}{2\lambda} \sup_{\nu \in [0, t^*]} \left(\Delta_g^\perp \|\partial \mathcal{V}(Z_{N, \nu})\|_{L_{2\mathfrak{p}}} \left\| \dot{\Lambda}_\mu^\perp(\nu, Z_{N, \nu}) \right\|_{L_{2\mathfrak{p}}} + \left(\left\| \dot{F}_\sigma^\perp(\nu, Z_{N, \nu}) \right\|_{L_{2\mathfrak{p}}} \right)^2 \right), \end{aligned} \quad (\text{D.49})$$

for all $(t, \mathfrak{p}) \in [0, t^*] \times \mathbb{N}_{\geq 2}$. We have thus established the bounds in (D.42). □

Next, we provide a result that is essential to the proof of the main results of the section.

Proposition D.8 Consider the following scalar processes:

$$\begin{aligned} N_1(t) &= \int_0^t e^{\omega \nu} M_\mu(t, \nu) \dot{\Lambda}_\mu^\parallel(\nu, Z_{N, \nu}) d\nu, & N_2(t) &= \int_0^t e^{\omega \nu} M_\mu(t, \nu) \dot{F}_\sigma^\parallel(\nu, Z_{N, \nu}) dW_\nu, \\ N_3(t) &= \int_0^t e^{\omega \nu} M_\sigma(t, \nu) \dot{\Lambda}_\mu^\parallel(\nu, Z_{N, \nu}) d\nu, & N_4(t) &= \int_0^t e^{\omega \nu} M_\sigma(t, \nu) \dot{F}_\sigma^\parallel(\nu, Z_{N, \nu}) dW_\nu, \end{aligned} \quad (\text{D.50})$$

where

$$M_\mu(t, \nu) = \int_\nu^t e^{(2\lambda - \omega)\beta} \mathcal{P}_\mu(\beta)^\top d\beta, \quad M_\sigma(t, \nu) = \int_\nu^t e^{(2\lambda - \omega)\beta} (\mathcal{P}_\sigma(\beta) dW_\beta)^\top,$$

for $0 \leq \nu \leq t \leq T$, and where $\mathcal{P}_\mu(\beta) \in \mathbb{R}^m$ and $\mathcal{P}_\sigma(\beta) \in \mathbb{R}^{m \times d}$ are defined in the statement of Proposition D.6.

If the stopping time τ^* , defined in (34), Lemma 4.1, satisfies $\tau^* = t^*$, then we have the following bounds:

$$\begin{aligned} \sum_{i=1}^4 \mathbb{E} [N_i(t)] & \leq \frac{e^{2\lambda t}}{\sqrt{\lambda \omega}} \Delta_{\mathcal{P}_1}(1, t^*) \sup_{\nu \in [0, t^*]} \left\| \dot{\Lambda}_\mu^\parallel(\nu, Z_{N, \nu}) \right\|_{L_2} \\ & \quad + \left(\frac{e^{2\lambda t}}{\sqrt{\lambda \omega}} \Delta_{\mathcal{P}_2}(1, t^*) + \frac{e^{2\lambda t}}{\lambda} \Delta_{\mathcal{P}_3}(1, t^*) \right) \sup_{\nu \in [0, t^*]} \left\| \dot{F}_\sigma^\parallel(\nu, Z_{N, \nu}) \right\|_{L_2}, \end{aligned} \quad (\text{D.51})$$

for all $t \in [0, t^*]$, and

$$\begin{aligned} \sum_{i=1}^4 \|N_i(t)\|_{L_p} & \leq \frac{e^{2\lambda t}}{\sqrt{\lambda \omega}} \Delta_{\mathcal{P}_1}(\mathfrak{p}, t^*) \sup_{\nu \in [0, t^*]} \left\| \dot{\Lambda}_\mu^\parallel(\nu, Z_{N, \nu}) \right\|_{L_{2\mathfrak{p}}} \\ & \quad + \left(\frac{e^{2\lambda t}}{\sqrt{\lambda \omega}} \Delta_{\mathcal{P}_2}(\mathfrak{p}, t^*) + \frac{e^{2\lambda t}}{\lambda} \Delta_{\mathcal{P}_3}(\mathfrak{p}, t^*) \right) \sup_{\nu \in [0, t^*]} \left\| \dot{F}_\sigma^\parallel(\nu, Z_{N, \nu}) \right\|_{L_{2\mathfrak{p}}}, \end{aligned} \quad (\text{D.52})$$

for all $(t, \mathfrak{p}) \in [0, t^*] \times \mathbb{N}_{\geq 2}$, where

$$\Delta_{\mathcal{P}_1}(\mathfrak{p}, t^*) = \frac{1}{2\sqrt{\lambda}} \sup_{\nu \in [0, t^*]} \|\mathcal{P}_\mu(\nu)\|_{L_{2\mathfrak{p}}} + \frac{\mathfrak{p}(\mathfrak{p})}{2} \sup_{\nu \in [0, t^*]} \|\mathcal{P}_\sigma(\nu)\|_{L_{2\mathfrak{p}}}, \quad \Delta_{\mathcal{P}_2}(\mathfrak{p}, t^*) = \mathfrak{p}'(\mathfrak{p}) \Delta_{\mathcal{P}_1}(\mathfrak{p}, t^*),$$

$$\Delta_{\mathcal{P}_3}(\mathfrak{p}, t^*) = \frac{\sqrt{m}}{2} \sup_{\nu \in [0, t^*]} \|\mathcal{P}_\sigma(\nu)\|_{L_{2\mathfrak{p}}},$$

and where the constants $\mathfrak{p}(\mathfrak{p})$ and $\mathfrak{p}'(\mathfrak{p})$ are defined in (A.1a).

Proof. We omit the details of the proof since it follows *mutatis mutandis* the proof of Proposition C.5. □

The following result establishes the effect of the reference feedback operator \mathcal{F}_r on the truncated joint process in Definition 9.

Proposition D.9 Consider the following scalar process:

$$N_{\mathcal{U}}(t) = \int_0^t e^{\omega \nu} M_{\mathcal{U}}(t, \nu) \left[\dot{\Lambda}_{\mu}^{\parallel}(\nu, Z_{N, \nu}) d\nu + \dot{F}_{\sigma}^{\parallel}(\nu, Z_{N, \nu}) dW_{\nu} \right], \quad (\text{D.53})$$

where

$$M_{\mathcal{U}}(t, \nu) = \int_{\nu}^t e^{(2\lambda - \omega)\beta} \mathcal{P}_{\mathcal{U}}(\beta)^{\top} d\beta,$$

for $0 \leq \nu \leq t \leq T$, and where $\mathcal{P}_{\mathcal{U}}(\beta) \in \mathbb{R}^m$ is defined in the statement of Proposition D.6. If the stopping time τ^* , defined in (104), Lemma 4.5, satisfies $\tau^* = t^*$, then

$$\|N_{\mathcal{U}}(t)\|_{L_p} \leq \Delta_g^2 \frac{e^{2\lambda t}}{\lambda} \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^*]} \left\| \dot{\Lambda}_{\mu}^{\parallel}(\nu, Z_{N, \nu}) \right\|_{L_{2p}} + p'(\mathbf{p}) \sup_{\nu \in [0, t^*]} \left\| \dot{F}_{\sigma}^{\parallel}(\nu, Z_{N, \nu}) \right\|_{L_{2p}} \right)^2, \quad (\text{D.54})$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 1}$.

Proof. We follow the derivation of (C.39) in the proof of Proposition C.6 to obtain the following representation of the process $N_{\mathcal{U}}(t)$:

$$N_{\mathcal{U}}(t) = \int_0^t e^{(2\lambda - \omega)\nu} \mathcal{P}_{\mathcal{U}}(\nu)^{\top} \left(\int_0^{\nu} e^{\omega \beta} \left[\dot{\Lambda}_{\mu}^{\parallel}(\beta, Z_{N, \beta}) d\beta + \dot{F}_{\sigma}^{\parallel}(\beta, Z_{N, \beta}) dW_{\beta} \right] \right) d\nu. \quad (\text{D.55})$$

Using the definition of $\mathcal{P}_{\mathcal{U}}$ in Proposition D.6, followed by the definition of $\dot{\mathcal{F}}_r$ in (103), we get:

$$\mathcal{P}_{\mathcal{U}}(\nu) = 2g(\nu)^{\top} g(\nu) \dot{\mathcal{F}}_r(Z_N)(\nu) = 2g(\nu)^{\top} g(\nu) (\mathcal{F}_r(X_N)(\nu) - \mathcal{F}_r(X_N^r)(\nu)).$$

We next incorporate the definition of \mathcal{F}_r in (27) for the truncated process (99) to obtain:

$$\begin{aligned} \mathcal{P}_{\mathcal{U}}(\nu) &= -2\omega e^{-\omega \nu} g(\nu)^{\top} g(\nu) \int_0^{\nu} e^{\omega \beta} (\dot{\Lambda}_{\mu}^{\parallel}(\beta, X_{N, \beta}) - \dot{\Lambda}_{\mu}^{\parallel}(\beta, X_{N, \beta}^r)) d\beta \\ &\quad - 2\omega e^{-\omega \nu} g(\nu)^{\top} g(\nu) \int_0^{\nu} e^{\omega \beta} (F_{\sigma}^{\parallel}(\beta, X_{N, \beta}) - F_{\sigma}^{\parallel}(\beta, X_{N, \beta}^r)) dW_{\beta} \\ &= -2\omega e^{-\omega \nu} g(\nu)^{\top} g(\nu) \int_0^{\nu} e^{\omega \beta} \left[\dot{\Lambda}_{\mu}^{\parallel}(\beta, Z_{N, \beta}) d\beta + \dot{F}_{\sigma}^{\parallel}(\beta, Z_{N, \beta}) dW_{\beta} \right], \end{aligned}$$

where we have used the definitions of $\dot{\Lambda}_{\mu}^{\parallel}(\beta, Z_{N, \beta})$ and $\dot{F}_{\sigma}^{\parallel}(\beta, Z_{N, \beta})$ in (103). Substituting the above expression into (D.55) produces

$$\begin{aligned} N_{\mathcal{U}}(t) &= -2\omega \int_0^t e^{2(\lambda - \omega)\nu} \left(\int_0^{\nu} e^{\omega \beta} \left[\dot{\Lambda}_{\mu}^{\parallel}(\beta, Z_{N, \beta}) d\beta + \dot{F}_{\sigma}^{\parallel}(\beta, Z_{N, \beta}) dW_{\beta} \right] \right)^{\top} g(\nu)^{\top} g(\nu) \\ &\quad \times \left(\int_0^{\nu} e^{\omega \beta} \left[\dot{\Lambda}_{\mu}^{\parallel}(\beta, Z_{N, \beta}) d\beta + \dot{F}_{\sigma}^{\parallel}(\beta, Z_{N, \beta}) dW_{\beta} \right] \right) d\nu. \quad (\text{D.56}) \end{aligned}$$

Observe that $N_{\mathcal{U}}(t)$ can be expressed as the process $N(t)$ in the statement of Lemma B.6 by setting

$$\begin{aligned} Q_t &= W_t \in \mathbb{R}^d (n_q = d), \quad \mathfrak{F}_t = \mathfrak{W}_t, \quad \theta_1 = \lambda, \quad \theta_2 = \omega, \quad \xi = 0, \\ R(t) &= -2\omega g(t)^{\top} g(t) \in \mathcal{M}_2^{loc}(\mathbb{S}^m | \mathfrak{W}_t), \\ S_1(t) &= S_2(t) = \dot{\Lambda}_{\mu}^{\parallel}(t, Z_{N, t}) \in \mathcal{M}_2^{loc}(\mathbb{R}^m | \mathfrak{W}_t), \\ L_1(t) &= L_2(t) = \dot{F}_{\sigma}^{\parallel}(t, Z_{N, t}) \in \mathcal{M}_2^{loc}(\mathbb{R}^{m \times d} | \mathfrak{W}_t). \end{aligned}$$

Furthermore, as a consequence of Assumption 1, we may set Δ_R in the hypothesis of Lemma B.6 as $\Delta_R = 2\Delta_g^2 \omega$. Hence, applying Lemma B.6 to the process $N_{\mathcal{U}}(t)$ in (D.56) leads to

$$\|N_{\mathcal{U}}(t)\|_{L_p} \leq \Delta_g^2 \omega \frac{e^{2\lambda t} - 1}{\lambda} \left(\frac{1 - e^{-\omega t}}{\omega} \sup_{\nu \in [0, t^*]} \left\| \dot{\Lambda}_{\mu}^{\parallel}(\nu, Z_{N, \nu}) \right\|_{L_{2p}} \right)$$

$$+ \left(\mathfrak{p} \frac{2\mathfrak{p} - 1}{2} \right)^{\frac{1}{2}} \left(\frac{1 - e^{-2\omega t}}{\omega} \right)^{\frac{1}{2}} \sup_{\nu \in [0, t^*]} \left\| \mathring{F}_\sigma^\parallel(\nu, Z_{N, \nu}) \right\|_{L_{2\mathfrak{p}}} \right)^2,$$

for all $(t, \mathfrak{p}) \in [0, t^*] \times \mathbb{N}_{\geq 1}$. Then, bounding $e^{2\lambda t} - 1$ by $e^{2\lambda t}$, and $1 - e^{-\omega t}$ and $1 - e^{-2\omega t}$ by 1 produces desired result in (D.54). \square

The subsequent results quantify the effect of the feedback error $\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r$ on the truncated joint process $Z_{N, t}$. We begin by deriving the moment bounds over the sampling interval \mathbf{T}_s for the truncated true (uncertain) \mathcal{L}_1 -DRAC process $X_{N, t}$.

Proposition D.10 *Consider the truncated true (uncertain) \mathcal{L}_1 -DRAC process $X_{N, t}$ presented in Definition 9. Then, for any $k \in \mathbb{N}_{\geq 1}$, for which $k\mathbf{T}_s < \tau^*$, we have the following bound*

$$\|X_{N, k\mathbf{T}_s} - X_{N, t}\|_{L_{2\mathfrak{p}}} \leq \mathbf{T}_s \tilde{\delta}_1(\mathbf{T}_s, \mathfrak{p}, \tau^*) + \mathbf{T}_s^{\frac{1}{2}} \tilde{\delta}_2(\mathfrak{p}, \tau^*), \quad \forall (t, \mathfrak{p}) \in [(k-1)\mathbf{T}_s, k\mathbf{T}_s] \times \mathbb{N}_{\geq 1}, \quad (\text{D.57})$$

where the stopping time τ^* is defined in Proposition 4.3, and

$$\begin{aligned} \tilde{\delta}_1(\mathbf{T}_s, \mathfrak{p}, \tau^*) &= 2 \sup_{\nu \in [0, \tau^*]} \left\| \bar{F}_\mu(\nu, X_{N, \nu}) \right\|_{L_{2\mathfrak{p}}} + 2(1 + \Delta_g \Delta_\Theta e^{-\lambda_s \mathbf{T}_s}) \sup_{\nu \in [0, \tau^*]} \left\| \Lambda_\mu(\nu, X_{N, \nu}) \right\|_{L_{2\mathfrak{p}}} \\ &\quad + 2\mathfrak{p}'(\mathfrak{p}) \Delta_g \Delta_\Theta e^{-\lambda_s \mathbf{T}_s} \gamma_1(\mathbf{T}_s) \sup_{\nu \in [0, \tau^*]} \left\| F_\sigma(\nu, X_{N, \nu}) \right\|_{L_{2\mathfrak{p}}}, \\ \tilde{\delta}_2(\mathfrak{p}, \tau^*) &= 2\sqrt{2}\mathfrak{p}'(\mathfrak{p}) \sup_{\nu \in [0, \tau^*]} \left\| F_\sigma(\nu, X_{N, \nu}) \right\|_{L_{2\mathfrak{p}}}, \\ \gamma_1(\mathbf{T}_s) &\doteq \left(\lambda_s \frac{e^{\lambda_s \mathbf{T}_s} + 1}{e^{\lambda_s \mathbf{T}_s} - 1} \right)^{\frac{1}{2}} = \left(\lambda_s \frac{1 + e^{-\lambda_s \mathbf{T}_s}}{1 - e^{-\lambda_s \mathbf{T}_s}} \right)^{\frac{1}{2}}. \end{aligned}$$

Proof. From Proposition 4.3, $Z_{N, t}$ is a unique strong solution of (97), for all $t \in [0, \tau_N]$. Moreover, $\tau^* \leq \tau_N$ by definition, therefore implying that $X_{N, t}$ is a unique strong solution of (90) in Definition 7, for all $t \in [0, \tau^*]$. That is,

$$dX_{N, t} = F_\mu(t, X_{N, t}, U_{\mathcal{L}_1, t}) dt + F_\sigma(t, X_{N, t}) dW_t, \quad t \in [0, \tau^*], \quad U_{\mathcal{L}_1, t} \doteq \mathcal{F}_{\mathcal{L}_1}(X_N)(t),$$

with $X_{N, 0} = x_0 \sim \xi_0$. Hence,

$$X_{N, t} = X_{N, 0} + \int_0^t F_\mu(\nu, X_{N, \nu}, U_{\mathcal{L}_1, \nu}) d\nu + \int_0^t F_\sigma(\nu, X_{N, \nu}) dW_\nu, \quad t \in [0, \tau^*], \quad U_{\mathcal{L}_1, t} \doteq \mathcal{F}_{\mathcal{L}_1}(X_N)(t).$$

Since $(k-1)\mathbf{T}_s$ is a constant, and hence a stopping time, it then follows that for any $t \in [(k-1)\mathbf{T}_s, k\mathbf{T}_s]$, we have

$$\begin{aligned} X_{N, t} - X_{N, (k-1)\mathbf{T}_s} &= \int_{(k-1)\mathbf{T}_s}^t F_\mu(\nu, X_{N, \nu}, U_{\mathcal{L}_1, \nu}) d\nu + \int_{(k-1)\mathbf{T}_s}^t F_\sigma(\nu, X_{N, \nu}) dW_\nu, \\ U_{\mathcal{L}_1, t} &\doteq \mathcal{F}_{\mathcal{L}_1}(X_N)(t), \end{aligned}$$

which yields the following due to the Minkowski's inequality:

$$\begin{aligned} \|X_{N, t} - X_{N, (k-1)\mathbf{T}_s}\|_{L_{2\mathfrak{p}}} &\leq \left\| \int_{(k-1)\mathbf{T}_s}^t F_\mu(\nu, X_{N, \nu}, U_{\mathcal{L}_1, \nu}) d\nu \right\|_{L_{2\mathfrak{p}}} \\ &\quad + \left\| \int_{(k-1)\mathbf{T}_s}^t F_\sigma(\nu, X_{N, \nu}) dW_\nu \right\|_{L_{2\mathfrak{p}}}, \quad (\text{D.58}) \end{aligned}$$

for all $(t, \mathfrak{p}) \in [(k-1)\mathbf{T}_s, k\mathbf{T}_s] \times \mathbb{N}_{\geq 1}$, where $U_{\mathcal{L}_1, t} \doteq \mathcal{F}_{\mathcal{L}_1}(X_N)(t)$.

Now, using Corollary B.1 with $\xi = (k-1)\mathbf{T}_s$ gives us

$$\begin{aligned} \left\| \int_{(k-1)\mathbf{T}_s}^t F_\mu(\nu, X_{N, \nu}, U_{\mathcal{L}_1, \nu}) d\nu \right\|_{L_{2\mathfrak{p}}} &\leq (t - (k-1)\mathbf{T}_s) \sup_{\nu \in [(k-1)\mathbf{T}_s, t]} \|F_\mu(\nu, X_{N, \nu}, U_{\mathcal{L}_1, \nu})\|_{L_{2\mathfrak{p}}} \\ &\leq \mathbf{T}_s \sup_{\nu \in [(k-1)\mathbf{T}_s, k\mathbf{T}_s]} \|F_\mu(\nu, X_{N, \nu}, U_{\mathcal{L}_1, \nu})\|_{L_{2\mathfrak{p}}}, \end{aligned}$$

for all $t \in [(k-1)\mathbf{T}_s, k\mathbf{T}_s]$. Using (3) in Definition 1 and the Minkowski's inequality, we can further develop the above bound as

$$\begin{aligned} & \left\| \int_{(k-1)\mathbf{T}_s}^t \|F_\mu(\nu, X_N, \nu, U_{\mathcal{L}_1, \nu})\| d\nu \right\|_{L_{2p}} \\ & \leq \mathbf{T}_s \sup_{\nu \in [(k-1)\mathbf{T}_s, k\mathbf{T}_s]} \left(\|\bar{F}_\mu(\nu, X_N, \nu)\|_{L_{2p}} + \Delta_g \|U_{\mathcal{L}_1, \nu}\|_{L_{2p}} + \|\Lambda_\mu(\nu, X_N, \nu)\|_{L_{2p}} \right), \end{aligned} \quad (\text{D.59})$$

for all $t \in [(k-1)\mathbf{T}_s, k\mathbf{T}_s]$, where we have used the uniform bound on $\|g(t)\|_F$ in Assumption 1. Next, from (93) and (94), we have

$$U_{\mathcal{L}_1, t} = \mathcal{F}_{\mathcal{L}_1}(X_N)(t) = \mathcal{F}_\omega(\hat{\Lambda}^\parallel)(t) = -\omega \int_0^t e^{-\omega(t-\nu)} \hat{\Lambda}^\parallel(\nu) d\nu, \quad t \in [(k-1)\mathbf{T}_s, k\mathbf{T}_s],$$

and thus

$$\|U_{\mathcal{L}_1, t}\|_{L_{2p}} \leq \omega e^{-\omega t} \left\| \int_0^t e^{\omega \nu} \hat{\Lambda}^\parallel(\nu) d\nu \right\|_{L_{2p}}, \quad \forall t \in [(k-1)\mathbf{T}_s, k\mathbf{T}_s].$$

It then follows from Proposition B.1 that

$$\|U_{\mathcal{L}_1, t}\|_{L_{2p}} \leq (1 - e^{-\omega t}) \sup_{\nu \in [0, t]} \|\hat{\Lambda}^\parallel(\nu)\|_{L_{2p}} \leq \sup_{\nu \in [0, t]} \|\hat{\Lambda}^\parallel(\nu)\|_{L_{2p}} \leq \sup_{\nu \in [0, k\mathbf{T}_s]} \|\hat{\Lambda}^\parallel(\nu)\|_{L_{2p}},$$

for all $t \in [(k-1)\mathbf{T}_s, k\mathbf{T}_s]$. The above further implies

$$\sup_{\nu \in [(k-1)\mathbf{T}_s, k\mathbf{T}_s]} \|U_{\mathcal{L}_1, \nu}\|_{L_{2p}} \leq \sup_{\nu \in [0, k\mathbf{T}_s]} \|\hat{\Lambda}^\parallel(\nu)\|_{L_{2p}}. \quad (\text{D.60})$$

Now, we obtain the following definition of $\hat{\Lambda}^\parallel(\nu)$ from (D.3), Proposition D.1, with $\tau = k\mathbf{T}_s$ and $z = X_N$:

$$\begin{aligned} \hat{\Lambda}^\parallel(\nu) &= 0_m \mathbb{1}_{[0, \mathbf{T}_s)}(\nu) \\ &\quad - \sum_{l=1}^k \left(\lambda_s (1 - e^{\lambda_s \mathbf{T}_s})^{-1} \Theta_{ad}(l\mathbf{T}_s) \int_{(l-1)\mathbf{T}_s}^{l\mathbf{T}_s} e^{-\lambda_s(l\mathbf{T}_s - \beta) \Sigma_\beta(X_N)} \mathbb{1}_{[l\mathbf{T}_s, (l+1)\mathbf{T}_s)}(\nu) d\beta, \quad \nu \in [0, k\mathbf{T}_s], \right. \end{aligned}$$

with $k \in \mathbb{N}_{\geq 1}$, where we have the following formal definition:

$$\Sigma_t(X_N) = \Lambda_\mu(t, X_N, t) dt + F_\sigma(t, X_N, t) dW_t, \quad t \in [0, k\mathbf{T}_s]. \quad (\text{D.61})$$

Therefore,

$$\|\hat{\Lambda}^\parallel(\nu)\| \leq 0 \cdot \mathbb{1}_{[0, \mathbf{T}_s)}(\nu) + \sum_{l=1}^k \left(\lambda_s (e^{\lambda_s \mathbf{T}_s} - 1)^{-1} \Delta_\Theta \left\| \int_{(l-1)\mathbf{T}_s}^{l\mathbf{T}_s} e^{-\lambda_s(l\mathbf{T}_s - \beta) \Sigma_\beta(X_N)} d\beta \right\|_{L_{2p}} \right) \mathbb{1}_{[l\mathbf{T}_s, (l+1)\mathbf{T}_s)}(\nu),$$

for all $\nu \in [0, k\mathbf{T}_s]$, where we have used the uniform bound on $\Theta_{ad}(t)$ in Assumption 1. The last inequality allows us to bound the $2p^{\text{th}}$ moment as follows:

$$\begin{aligned} \|\hat{\Lambda}^\parallel(\nu)\|_{L_{2p}} &\leq 0 \cdot \mathbb{1}_{[0, \mathbf{T}_s)}(\nu) \\ &\quad + \sum_{l=1}^k \left(\lambda_s (e^{\lambda_s \mathbf{T}_s} - 1)^{-1} \Delta_\Theta \left\| \int_{(l-1)\mathbf{T}_s}^{l\mathbf{T}_s} e^{-\lambda_s(l\mathbf{T}_s - \beta) \Sigma_\beta(X_N)} d\beta \right\|_{L_{2p}} \right) \mathbb{1}_{[l\mathbf{T}_s, (l+1)\mathbf{T}_s)}(\nu), \end{aligned}$$

for all $\nu \in [0, k\mathbf{T}_s]$. Thus, over each piecewise temporal interval $[l\mathbf{T}_s, (l+1)\mathbf{T}_s]$, $l \in \{1, \dots, k\}$, we have

$$\begin{aligned} \|\hat{\Lambda}^\parallel(\nu)\|_{L_{2p}} &\leq \lambda_s (e^{\lambda_s \mathbf{T}_s} - 1)^{-1} \Delta_\Theta \left\| \int_{(l-1)\mathbf{T}_s}^{l\mathbf{T}_s} e^{-\lambda_s(l\mathbf{T}_s - \beta) \Sigma_\beta(X_N)} d\beta \right\|_{L_{2p}} \\ &\leq \lambda_s (e^{\lambda_s \mathbf{T}_s} - 1)^{-1} \Delta_\Theta e^{-\lambda_s l \mathbf{T}_s} \left\| \int_{(l-1)\mathbf{T}_s}^{l\mathbf{T}_s} e^{\lambda_s \beta \Sigma_\beta(X_N)} d\beta \right\|_{L_{2p}}, \quad \forall \nu \in [l\mathbf{T}_s, (l+1)\mathbf{T}_s], \end{aligned}$$

where the last inequality is due to both l and \mathbf{T}_s being constants. Substituting the expression in (D.61) into the above and using the Minkowski's inequality leads to

$$\begin{aligned} \left\| \hat{\Lambda}^\parallel(\nu) \right\|_{L_{2p}} &\leq \lambda_s (e^{\lambda_s \mathbf{T}_s} - 1)^{-1} \Delta_\Theta e^{-\lambda_s l \mathbf{T}_s} \left\| \int_{(l-1)\mathbf{T}_s}^{l\mathbf{T}_s} e^{\lambda_s \beta} \Lambda_\mu(\beta, X_{N,\beta}) d\beta \right\|_{L_{2p}} \\ &\quad + \lambda_s (e^{\lambda_s \mathbf{T}_s} - 1)^{-1} \Delta_\Theta e^{-\lambda_s l \mathbf{T}_s} \left\| \int_{(l-1)\mathbf{T}_s}^{l\mathbf{T}_s} e^{\lambda_s \beta} F_\sigma(\beta, X_{N,\beta}) dW_\beta \right\|_{L_{2p}}, \end{aligned}$$

for all $\nu \in [l\mathbf{T}_s, (l+1)\mathbf{T}_s]$, $l \in \{1, \dots, k\}$. We now use Proposition B.1 and Lemma B.2 to bound the terms on the right-hand side of the above expression

$$\begin{aligned} \left\| \hat{\Lambda}^\parallel(\nu) \right\|_{L_{2p}} &\leq \Delta_\Theta e^{-\lambda_s \mathbf{T}_s} \sup_{\nu \in [(l-1)\mathbf{T}_s, l\mathbf{T}_s]} \left\| \Lambda_\mu(\nu, X_{N,\nu}) \right\|_{L_{2p}} \\ &\quad + \Delta_\Theta \mathbf{p}'(\mathbf{p}) e^{-\lambda_s \mathbf{T}_s} \left(\lambda_s \frac{e^{\lambda_s \mathbf{T}_s} + 1}{e^{\lambda_s \mathbf{T}_s} - 1} \right)^{\frac{1}{2}} \sup_{\nu \in [(l-1)\mathbf{T}_s, l\mathbf{T}_s]} \left\| F_\sigma(\nu, X_{N,\nu}) \right\|_{L_{2p}}, \end{aligned}$$

for all $\nu \in [l\mathbf{T}_s, (l+1)\mathbf{T}_s]$, $l \in \{1, \dots, k\}$, where we have used the following:

$$(e^{\lambda_s \mathbf{T}_s} - 1)^{-1} (e^{2\lambda_s \mathbf{T}_s} - 1)^{\frac{1}{2}} = (e^{\lambda_s \mathbf{T}_s} - 1)^{-1} (e^{\lambda_s \mathbf{T}_s} - 1)^{\frac{1}{2}} (e^{\lambda_s \mathbf{T}_s} + 1)^{\frac{1}{2}} = \left(\frac{e^{\lambda_s \mathbf{T}_s} + 1}{e^{\lambda_s \mathbf{T}_s} - 1} \right)^{\frac{1}{2}}.$$

Since $k \in \mathbb{N}_{\geq 1}$ satisfies $k\mathbf{T}_s < \tau^*$, we have that $[(l-1)\mathbf{T}_s, l\mathbf{T}_s] \subset [0, \tau^*]$, for all $l \in \{1, \dots, k\}$. Thus, we can further develop the above inequality into

$$\begin{aligned} \left\| \hat{\Lambda}^\parallel(\nu) \right\|_{L_{2p}} &\leq \Delta_\Theta e^{-\lambda_s \mathbf{T}_s} \sup_{\nu \in [0, \tau^*]} \left\| \Lambda_\mu(\nu, X_{N,\nu}) \right\|_{L_{2p}} \\ &\quad + \Delta_\Theta \mathbf{p}'(\mathbf{p}) e^{-\lambda_s \mathbf{T}_s} \left(\lambda_s \frac{e^{\lambda_s \mathbf{T}_s} + 1}{e^{\lambda_s \mathbf{T}_s} - 1} \right)^{\frac{1}{2}} \sup_{\nu \in [0, \tau^*]} \left\| F_\sigma(\nu, X_{N,\nu}) \right\|_{L_{2p}}, \end{aligned}$$

for all $\nu \in [l\mathbf{T}_s, (l+1)\mathbf{T}_s]$, $l \in \{1, \dots, k\}$. Since

$$\sup_{\nu \in [0, k\mathbf{T}_s]} \left\| \hat{\Lambda}^\parallel(\nu) \right\|_{L_{2p}} \leq \max_{l \in \{1, \dots, k\}} \sup_{\nu \in [(l-1)\mathbf{T}_s, l\mathbf{T}_s]} \left\| \hat{\Lambda}^\parallel(\nu) \right\|_{L_{2p}},$$

the previous bound implies

$$\begin{aligned} \sup_{\nu \in [0, k\mathbf{T}_s]} \left\| \hat{\Lambda}^\parallel(\nu) \right\|_{L_{2p}} &\leq \Delta_\Theta e^{-\lambda_s \mathbf{T}_s} \sup_{\nu \in [0, \tau^*]} \left\| \Lambda_\mu(\nu, X_{N,\nu}) \right\|_{L_{2p}} \\ &\quad + \Delta_\Theta \mathbf{p}'(\mathbf{p}) e^{-\lambda_s \mathbf{T}_s} \left(\lambda_s \frac{e^{\lambda_s \mathbf{T}_s} + 1}{e^{\lambda_s \mathbf{T}_s} - 1} \right)^{\frac{1}{2}} \sup_{\nu \in [0, \tau^*]} \left\| F_\sigma(\nu, X_{N,\nu}) \right\|_{L_{2p}}. \end{aligned}$$

Substituting this bound into (D.60):

$$\begin{aligned} \sup_{\nu \in [(k-1)\mathbf{T}_s, k\mathbf{T}_s]} \left\| U_{\mathcal{L}_1, \nu} \right\|_{L_{2p}} &\leq \Delta_\Theta e^{-\lambda_s \mathbf{T}_s} \sup_{\nu \in [0, \tau^*]} \left\| \Lambda_\mu(\nu, X_{N,\nu}) \right\|_{L_{2p}} \\ &\quad + \Delta_\Theta \mathbf{p}'(\mathbf{p}) e^{-\lambda_s \mathbf{T}_s} \left(\lambda_s \frac{e^{\lambda_s \mathbf{T}_s} + 1}{e^{\lambda_s \mathbf{T}_s} - 1} \right)^{\frac{1}{2}} \sup_{\nu \in [0, \tau^*]} \left\| F_\sigma(\nu, X_{N,\nu}) \right\|_{L_{2p}}, \end{aligned}$$

which we further substitute into (D.59) and get

$$\begin{aligned} &\left\| \int_{(k-1)\mathbf{T}_s}^t \left\| F_\mu(\nu, X_{N,\nu}, U_{\mathcal{L}_1, \nu}) \right\| d\nu \right\|_{L_{2p}} \\ &\leq \mathbf{T}_s \sup_{\nu \in [0, \tau^*]} \left\| \bar{F}_\mu(\nu, X_{N,\nu}) \right\|_{L_{2p}} + \mathbf{T}_s (1 + \Delta_g \Delta_\Theta e^{-\lambda_s \mathbf{T}_s}) \sup_{\nu \in [0, \tau^*]} \left\| \Lambda_\mu(\nu, X_{N,\nu}) \right\|_{L_{2p}} \end{aligned}$$

$$+ \mathbf{p}'(\mathbf{p}) \mathbf{T}_s \Delta_g \Delta_\Theta e^{-\lambda_s \mathbf{T}_s} \left(\lambda_s \frac{e^{\lambda_s \mathbf{T}_s} + 1}{e^{\lambda_s \mathbf{T}_s} - 1} \right)^{\frac{1}{2}} \sup_{\nu \in [0, \tau^*]} \|F_\sigma(\nu, X_{N,\nu})\|_{L_{2\mathbf{p}}}, \quad (\text{D.62})$$

for all $t \in [(k-1)\mathbf{T}_s, k\mathbf{T}_s]$, where, similar as above, we have generalized the upper bound to $[0, \tau^*]$.

Next, we invoke [91, Thm. 1.7.1] (as used in [91, Thm. 2.4.3]) to obtain

$$\begin{aligned} & \mathbb{E} \left[\left\| \int_{(k-1)\mathbf{T}_s}^t F_\sigma(\nu, X_{N,\nu}) dW_\nu \right\|^{2\mathbf{p}} \right] \\ & \leq (\mathbf{p}(2\mathbf{p}-1))^\mathbf{p} (t - (k-1)\mathbf{T}_s)^{\mathbf{p}-1} \mathbb{E} \left[\int_{(k-1)\mathbf{T}_s}^t \|F_\sigma(\nu, X_{N,\nu})\|^{2\mathbf{p}} d\nu \right], \quad \forall t \in [(k-1)\mathbf{T}_s, k\mathbf{T}_s]. \end{aligned}$$

Hence, it follows from the Fubini's theorem that

$$\begin{aligned} & \mathbb{E} \left[\left\| \int_{(k-1)\mathbf{T}_s}^t F_\sigma(\nu, X_{N,\nu}) dW_\nu \right\|^{2\mathbf{p}} \right] \\ & \leq (\mathbf{p}(2\mathbf{p}-1))^\mathbf{p} (t - (k-1)\mathbf{T}_s)^{\mathbf{p}-1} \int_{(k-1)\mathbf{T}_s}^t \mathbb{E} [\|F_\sigma(\nu, X_{N,\nu})\|^{2\mathbf{p}}] d\nu \\ & \leq (\mathbf{p}(2\mathbf{p}-1))^\mathbf{p} (t - (k-1)\mathbf{T}_s)^\mathbf{p} \sup_{\nu \in [t, (k-1)\mathbf{T}_s]} \mathbb{E} [\|F_\sigma(\nu, X_{N,\nu})\|^{2\mathbf{p}}], \end{aligned}$$

for all $t \in [(k-1)\mathbf{T}_s, k\mathbf{T}_s]$, which further implies

$$\begin{aligned} & \mathbb{E} \left[\left\| \int_{(k-1)\mathbf{T}_s}^t F_\sigma(\nu, X_{N,\nu}) dW_\nu \right\|^{2\mathbf{p}} \right] \leq (\mathbf{p}(2\mathbf{p}-1))^\mathbf{p} \mathbf{T}_s^\mathbf{p} \sup_{\nu \in [k\mathbf{T}_s, (k-1)\mathbf{T}_s]} \mathbb{E} [\|F_\sigma(\nu, X_{N,\nu})\|^{2\mathbf{p}}] \\ & \leq (\mathbf{p}(2\mathbf{p}-1))^\mathbf{p} \mathbf{T}_s^\mathbf{p} \sup_{\nu \in [0, \tau^*]} \mathbb{E} [\|F_\sigma(\nu, X_{N,\nu})\|^{2\mathbf{p}}], \end{aligned}$$

for all $t \in [(k-1)\mathbf{T}_s, k\mathbf{T}_s]$. Hence,

$$\begin{aligned} & \left\| \int_{(k-1)\mathbf{T}_s}^t F_\sigma(\nu, X_{N,\nu}) dW_\nu \right\|_{L_{2\mathbf{p}}} \leq (\mathbf{p}(2\mathbf{p}-1))^{\frac{1}{2}} \mathbf{T}_s^{\frac{1}{2}} \sup_{\nu \in [0, \tau^*]} \|F_\sigma(\nu, X_{N,\nu})\|_{L_{2\mathbf{p}}} \\ & = \sqrt{2\mathbf{p}'}(\mathbf{p}) \mathbf{T}_s^{\frac{1}{2}} \sup_{\nu \in [0, \tau^*]} \|F_\sigma(\nu, X_{N,\nu})\|_{L_{2\mathbf{p}}}, \quad (\text{D.63}) \end{aligned}$$

for all $t \in [(k-1)\mathbf{T}_s, k\mathbf{T}_s]$. Substituting (D.62) and (D.63) into (D.58) produces

$$\begin{aligned} & \|X_{N,t} - X_{N,(k-1)\mathbf{T}_s}\|_{L_{2\mathbf{p}}} \\ & \leq \mathbf{T}_s \sup_{\nu \in [0, \tau^*]} \|\bar{F}_\mu(\nu, X_{N,\nu})\|_{L_{2\mathbf{p}}} + \mathbf{T}_s (1 + \Delta_g \Delta_\Theta e^{-\lambda_s \mathbf{T}_s}) \sup_{\nu \in [0, \tau^*]} \|\Lambda_\mu(\nu, X_{N,\nu})\|_{L_{2\mathbf{p}}} \\ & \quad + \mathbf{p}'(\mathbf{p}) \mathbf{T}_s \Delta_g \Delta_\Theta e^{-\lambda_s \mathbf{T}_s} \left(\lambda_s \frac{e^{\lambda_s \mathbf{T}_s} + 1}{e^{\lambda_s \mathbf{T}_s} - 1} \right)^{\frac{1}{2}} \sup_{\nu \in [0, \tau^*]} \|F_\sigma(\nu, X_{N,\nu})\|_{L_{2\mathbf{p}}} \\ & \quad + \sqrt{2\mathbf{p}'}(\mathbf{p}) \mathbf{T}_s^{\frac{1}{2}} \sup_{\nu \in [0, \tau^*]} \|F_\sigma(\nu, X_{N,\nu})\|_{L_{2\mathbf{p}}}, \end{aligned}$$

for all $(t, \mathbf{p}) \in [(k-1)\mathbf{T}_s, k\mathbf{T}_s] \times \mathbb{N}_{\geq 1}$. The proof is then concluded by using this bound into the following consequence of the triangle inequality:

$$\|X_{N,k\mathbf{T}_s} - X_{N,t}\|_{L_{2\mathbf{p}}} \leq \|X_{N,t} - X_{N,(k-1)\mathbf{T}_s}\|_{L_{2\mathbf{p}}} + \|X_{N,k\mathbf{T}_s} - X_{N,(k-1)\mathbf{T}_s}\|_{L_{2\mathbf{p}}},$$

for all $(t, \mathbf{p}) \in [(k-1)\mathbf{T}_s, k\mathbf{T}_s] \times \mathbb{N}_{\geq 1}$.

□

The next result establishes uniform bounds on pertinent integrands that define the feedback error operator $\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_\tau$.

Proposition D.11 Consider the terms $\tilde{\mathcal{H}}_{\{\mu, \sigma\}}^{\{\cdot, \parallel\}}$, \mathcal{G}_μ , and $\tilde{\mathcal{G}}_{\{\sigma_1, \sigma_2\}}$, defined in the statement of Proposition D.3. Then, if the stopping time τ^* , defined in (104), Lemma 4.5, satisfies $\tau^* = t^*$ and $t^* \geq \mathbf{T}_s$, we have the following bounds for all $(t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 1}$:

$$\left\| \tilde{\mathcal{H}}_\mu^\parallel(t, X_N) + e^{(\lambda_s - \omega)t} \tilde{\mathcal{H}}_\mu(t, X_N) \right\|_{L_{2\mathbf{p}}} \leq \Delta_{\tilde{\mathcal{H}}_\mu}(\mathbf{T}_s, \mathbf{p}, t^*), \quad (\text{D.64a})$$

$$\left\| \tilde{\mathcal{H}}_\sigma^\parallel(t, X_N) + e^{(\lambda_s - \omega)t} \tilde{\mathcal{H}}_\sigma(t, X_N) \right\|_{L_{2\mathbf{p}}} \leq \Delta_{\tilde{\mathcal{H}}_\sigma}(\mathbf{T}_s, \mathbf{p}, t^*), \quad (\text{D.64b})$$

$$\left\| \tilde{\mathcal{G}}_{\sigma_1}(t, X_N) + e^{(\lambda_s - \omega)t} \tilde{\mathcal{G}}_{\sigma_2}(t, X_N) \right\|_{L_{2\mathbf{p}}} \leq \Delta_{\tilde{\mathcal{G}}_\sigma}(\mathbf{T}_s, \mathbf{p}, t^*), \quad (\text{D.64c})$$

$$\|\mathcal{G}_\mu(t, X_N)\|_{L_{2\mathbf{p}}} \leq \Delta_{\mathcal{G}_\mu}(\mathbf{T}_s, t^*), \quad (\text{D.64d})$$

where

$$\begin{aligned} \Delta_{\tilde{\mathcal{H}}_\mu}(\mathbf{T}_s, \mathbf{p}, t^*) &= \sqrt{\mathbf{T}_s} \gamma_\mu(\omega, \mathbf{T}_s) \tilde{\delta}_2(\mathbf{p}, t^*) + \mathbf{T}_s \left(\hat{\gamma}_\mu(\omega, \mathbf{T}_s) + \gamma_\mu(\omega, \mathbf{T}_s) \tilde{\delta}_1(\mathbf{T}_s, \mathbf{p}, t^*) \right), \\ \Delta_{\tilde{\mathcal{H}}_\sigma}(\mathbf{T}_s, \mathbf{p}, t^*) &= \gamma_\sigma(\omega, \mathbf{T}_s) \left(\mathbf{T}_s \tilde{\delta}_1(\mathbf{T}_s, \mathbf{p}, t^*) + \mathbf{T}_s^{\frac{1}{2}} \tilde{\delta}_2(\mathbf{p}, t^*) \right)^{\frac{1}{2}} + \mathbf{T}_s \hat{\gamma}_\sigma(\omega, \mathbf{T}_s), \\ \Delta_{\tilde{\mathcal{G}}_\sigma}(\mathbf{T}_s, \mathbf{p}, t^*) &= (1 + \gamma_2(\omega, \mathbf{T}_s)) \\ &\quad \times \left[(L_p^\parallel + L_\sigma^\parallel) \left(\mathbf{T}_s \tilde{\delta}_1(\mathbf{T}_s, \mathbf{p}, t^*) + \mathbf{T}_s^{\frac{1}{2}} \tilde{\delta}_2(\mathbf{p}, t^*) \right)^{\frac{1}{2}} + \mathbf{T}_s (\hat{L}_p^\parallel + \hat{L}_\sigma^\parallel) \right], \\ \Delta_{\mathcal{G}_\mu}(\mathbf{T}_s, t^*) &= (1 - e^{-\lambda_s \mathbf{T}_s}) \sup_{t \in [0, t^*]} \|\Lambda_\mu^\parallel(t, X_{N,t})\|_{L_{2\mathbf{p}}}, \end{aligned}$$

and where $\tilde{\delta}_1(\mathbf{T}_s, \mathbf{p}, t^*)$, and $\tilde{\delta}_2(\mathbf{p}, t^*)$ are defined in the statement of Proposition D.10, $\gamma_2(\omega, \mathbf{T}_s)$ is defined in (A.13), and $\gamma_{\{\mu, \sigma\}}(\omega, \mathbf{T}_s)$ and $\hat{\gamma}_{\{\mu, \sigma\}}(\omega, \mathbf{T}_s)$ are defined in (A.14).

Proof. We first begin by recalling the definitions in (D.11b)-(D.11c) with $z = X_N$ and for any $i \in \mathbb{N}_{\geq 1}$:

$$\begin{aligned} \tilde{\Lambda}_\mu^{\{\cdot, \parallel\}}(t, i\mathbf{T}_s, X_N) &= \Lambda_\mu^{\{\cdot, \parallel, \perp\}}(t, X_{N,t}) - \Lambda_\mu^{\{\cdot, \parallel, \perp\}}(i\mathbf{T}_s, X_{N,i\mathbf{T}_s}), \\ \tilde{F}_\sigma^{\{\cdot, \parallel\}}(t, i\mathbf{T}_s, X_N) &= F_\sigma^{\{\cdot, \parallel, \perp\}}(t, X_{N,t}) - F_\sigma^{\{\cdot, \parallel, \perp\}}(i\mathbf{T}_s, X_{N,i\mathbf{T}_s}) \\ &= \left(p^{\{\cdot, \parallel\}}(t, X_{N,t}) - p^{\{\cdot, \parallel\}}(i\mathbf{T}_s, X_{N,i\mathbf{T}_s}) \right) + \left(\Lambda_\sigma^{\{\cdot, \parallel\}}(t, X_{N,t}) - \Lambda_\sigma^{\{\cdot, \parallel\}}(i\mathbf{T}_s, X_{N,i\mathbf{T}_s}) \right), \end{aligned}$$

where we have used the decompositions $F_\sigma(\nu, \cdot) = p(\nu, \cdot) + \Lambda_\sigma(\nu, \cdot)$ and $F_\sigma^\parallel(\nu, \cdot) = p^\parallel(\nu, \cdot) + \Lambda_\sigma^\parallel(\nu, \cdot)$ from Definitions 1 and 3, respectively.

It then follows from Assumptions 1, 5, 6, and 8, and the Minkowski's inequality that

$$\begin{aligned} \left\| \tilde{\Lambda}_\mu^{\{\cdot, \parallel\}}(t, i\mathbf{T}_s, X_N) \right\|_{L_{2\mathbf{p}}} &\leq \hat{L}_\mu^{\{\cdot, \parallel\}} \mathbf{T}_s + L_\mu^{\{\cdot, \parallel\}} \|X_{N,t} - X_{N,i\mathbf{T}_s}\|_{L_{2\mathbf{p}}}, \\ \left\| \tilde{F}_\sigma^{\{\cdot, \parallel\}}(t, i\mathbf{T}_s, X_N) \right\|_{L_{2\mathbf{p}}} &\leq \left(\hat{L}_p^{\{\cdot, \parallel\}} + \hat{L}_\sigma^{\{\cdot, \parallel\}} \right) \mathbf{T}_s + \left(L_p^{\{\cdot, \parallel\}} + L_\sigma^{\{\cdot, \parallel\}} \right) \left(\|X_{N,t} - X_{N,i\mathbf{T}_s}\|_{L_{2\mathbf{p}}} \right)^{\frac{1}{2}}, \end{aligned}$$

for all $(i, t, \mathbf{p}) \in \mathbb{N}_{\geq 1} \times [(i-1)\mathbf{T}_s, i\mathbf{T}_s] \times \mathbb{N}_{\geq 1}$. With $\tau^* = t^*$ and $k = i$, we use the bound for $\|X_{N,t} - X_{N,i\mathbf{T}_s}\|_{L_{2\mathbf{p}}}$ from Proposition D.10 and obtain

$$\left\| \tilde{\Lambda}_\mu^{\{\cdot, \parallel\}}(t, i\mathbf{T}_s, X_N) \right\|_{L_{2\mathbf{p}}} \leq \sqrt{\mathbf{T}_s} L_\mu^{\{\cdot, \parallel\}} \tilde{\delta}_2(\mathbf{p}, t^*) + \mathbf{T}_s \left(\hat{L}_\mu^{\{\cdot, \parallel\}} + L_\mu^{\{\cdot, \parallel\}} \tilde{\delta}_1(\mathbf{T}_s, \mathbf{p}, t^*) \right), \quad (\text{D.65a})$$

$$\begin{aligned} \left\| \tilde{F}_\sigma^{\{\cdot, \parallel\}}(t, i\mathbf{T}_s, X_N) \right\|_{L_{2\mathbf{p}}} &\leq \mathbf{T}_s \left(\hat{L}_p^{\{\cdot, \parallel\}} + \hat{L}_\sigma^{\{\cdot, \parallel\}} \right) \\ &\quad + \left(L_p^{\{\cdot, \parallel\}} + L_\sigma^{\{\cdot, \parallel\}} \right) \left(\mathbf{T}_s \tilde{\delta}_1(\mathbf{T}_s, \mathbf{p}, t^*) + \mathbf{T}_s^{\frac{1}{2}} \tilde{\delta}_2(\mathbf{p}, t^*) \right)^{\frac{1}{2}}, \end{aligned} \quad (\text{D.65b})$$

for all $(i, t, \mathbf{p}) \in \mathbb{N}_{\geq 1} \times [(i-1)\mathbf{T}_s, i\mathbf{T}_s] \times \mathbb{N}_{\geq 1}$.

Now, with $z = X_N$ and $\tau = t^*$, we use the definition of $\tilde{\mathcal{H}}_\mu^\parallel$ and the Minkowski's inequality to conclude that

$$\left\| \tilde{\mathcal{H}}_\mu^\parallel(t, X_N) \right\|_{L_{2\mathbf{p}}}$$

$$\begin{aligned} &\leq \mathbb{1}_{<T_s}(t) 0 + \sum_{i=1}^{i^*(t^*)-1} \mathbb{1}_{[iT_s, (i+1)T_s)}(t) \left\| \tilde{\Lambda}_\mu^\parallel(t, iT_s, X_N) \right\|_{L_{2p}} + \mathbb{1}_{\geq i^*(t^*)T_s}(t) \left\| \tilde{\Lambda}_\mu^\parallel(t, i^*(t^*)T_s, X_N) \right\|_{L_{2p}} \\ &\leq \max_{i \in \{1, \dots, i^*(t^*)-1\}} \left\{ \sup_{t \in [iT_s, (i+1)T_s)} \left\| \tilde{\Lambda}_\mu^\parallel(t, iT_s, X_N) \right\|_{L_{2p}}, \sup_{t \in [i^*(t^*)T_s, t^*]} \left\| \tilde{\Lambda}_\mu^\parallel(t, i^*(t^*)T_s, X_N) \right\|_{L_{2p}} \right\}, \end{aligned}$$

for all $(t, p) \in [0, t^*] \times \mathbb{N}_{\geq 1}$. Similarly, using the definition of $\tilde{\mathcal{H}}_\mu$ from Proposition D.3, with $z = X_N$ and $\tau = t^*$, we get

$$\begin{aligned} &\left\| e^{(\lambda_s - \omega)t} \tilde{\mathcal{H}}_\mu(t, X_N) \right\|_{L_{2p}} \\ &\leq \Delta_\Theta \max_{i \in \{1, \dots, i^*(t^*)-1\}} \left\{ \sup_{t \in [(i-1)T_s, iT_s]} \left(\left| e^{(\lambda_s - \omega)t} \tilde{\gamma}_i(T_s) \right| \left\| \tilde{\Lambda}_\mu(t, iT_s, X_N) \right\|_{L_{2p}} \right), \right. \\ &\quad \left. \sup_{t \in [(i^*(t^*)-1)T_s, i^*(t^*)T_s]} \left(\left| e^{(\lambda_s - \omega)t} \tilde{\gamma}^*(t^*, T_s) \right| \left\| \tilde{\Lambda}_\mu(t, i^*(t^*)T_s, X_N) \right\|_{L_{2p}} \right) \right\}, \end{aligned}$$

for all $(t, p) \in [0, t^*] \times \mathbb{N}_{\geq 1}$, where we have used the uniform bound on $\|\Theta_{ad}(t)\|_F$ from Assumption 1. It then follows from the bounds above and the Minkowski's inequality that

$$\left\| \tilde{\mathcal{H}}_\mu(t, X_N) + e^{(\lambda_s - \omega)t} \tilde{\mathcal{H}}_\mu(t, X_N) \right\|_{L_{2p}} \leq \Psi_{\mathcal{H}_\mu}^\parallel + \Psi_{\mathcal{H}_\mu}, \quad \forall (t, p) \in [0, t^*] \times \mathbb{N}_{\geq 1}, \quad (\text{D.66})$$

where

$$\begin{aligned} \Psi_{\mathcal{H}_\mu}^\parallel &= \max_{i \in \{1, \dots, i^*(t^*)-1\}} \left\{ \sup_{t \in [iT_s, (i+1)T_s)} \left\| \tilde{\Lambda}_\mu^\parallel(t, iT_s, X_N) \right\|_{L_{2p}}, \sup_{t \in [i^*(t^*)T_s, t^*]} \left\| \tilde{\Lambda}_\mu^\parallel(t, i^*(t^*)T_s, X_N) \right\|_{L_{2p}} \right\}, \\ \Psi_{\mathcal{H}_\mu} &= \Delta_\Theta \max_{i \in \{1, \dots, i^*(t^*)-1\}} \left\{ \sup_{t \in [(i-1)T_s, iT_s]} \left(\left| e^{(\lambda_s - \omega)t} \tilde{\gamma}_i(T_s) \right| \left\| \tilde{\Lambda}_\mu(t, iT_s, X_N) \right\|_{L_{2p}} \right), \right. \\ &\quad \left. \sup_{t \in [(i^*(t^*)-1)T_s, i^*(t^*)T_s]} \left(\left| e^{(\lambda_s - \omega)t} \tilde{\gamma}^*(t^*, T_s) \right| \left\| \tilde{\Lambda}_\mu(t, i^*(t^*)T_s, X_N) \right\|_{L_{2p}} \right) \right\}. \end{aligned}$$

It is straightforward to see from (D.65a) that

$$\Psi_{\mathcal{H}_\mu}^\parallel \leq \sqrt{T_s} L_\mu^\parallel \tilde{\delta}_2(p, t^*) + T_s \left(\hat{L}_\mu^\parallel + L_\mu^\parallel \tilde{\delta}_1(T_s, p, t^*) \right). \quad (\text{D.67})$$

Next, using the definition of $\tilde{\gamma}_i$ from Proposition D.3, one sees that

$$e^{(\lambda_s - \omega)t} \tilde{\gamma}_i(T_s) = e^{(\lambda_s - \omega)(t - iT_s)} \frac{\lambda_s}{\omega} \frac{e^{\omega T_s} - 1}{e^{\lambda_s T_s} - 1} \leq \gamma_2(\omega, T_s), \quad (\text{D.68})$$

for all $t \in [(i-1)T_s, iT_s]$. Similarly, using the definition of $\tilde{\gamma}^*$ from Proposition D.3, we have that

$$e^{(\lambda_s - \omega)t} \tilde{\gamma}^*(t^*, T_s) = e^{(\lambda_s - \omega)(t - i^*(t^*)T_s)} \frac{\lambda_s}{\omega} \frac{e^{\omega(t^* - i^*(t^*)T_s)} - 1}{e^{\lambda_s T_s} - 1} \leq \gamma_2(\omega, T_s), \quad (\text{D.69})$$

for all $t \in [(i^*(t^*)-1)T_s, i^*(t^*)T_s]$, where we have used the fact that $i^*(t^*) \doteq \lfloor t^*/T_s \rfloor$, and thus $(t^* - i^*(t^*)T_s) \leq T_s$. Then, (D.65a), along with (D.68) and (D.69), allow us to obtain the following bound:

$$\Psi_{\mathcal{H}_\mu} \leq \sqrt{T_s} L_\mu \gamma_2(\omega, T_s) \Delta_\Theta \tilde{\delta}_2(p, t^*) + T_s \gamma_2(\omega, T_s) \Delta_\Theta \left(\hat{L}_\mu + L_\mu \tilde{\delta}_1(T_s, p, t^*) \right). \quad (\text{D.70})$$

Substituting (D.67) and (D.70) into (D.66) leads to (D.64a).

Next, with $z = X_N$ and $\tau = t^*$, we use the definition of $\tilde{\mathcal{H}}_\sigma^\parallel$ and the Minkowski's inequality to conclude that

$$\begin{aligned} &\left\| \tilde{\mathcal{H}}_\mu(t, X_N) \right\|_{L_{2p}} \\ &\leq \mathbb{1}_{<T_s}(t) 0 + \sum_{i=1}^{i^*(t^*)-1} \mathbb{1}_{[iT_s, (i+1)T_s)}(t) \left\| \tilde{F}_\sigma^\parallel(t, iT_s, X_N) \right\|_{L_{2p}} + \mathbb{1}_{\geq i^*(t^*)T_s}(t) \left\| \tilde{F}_\sigma^\parallel(t, i^*(t^*)T_s, X_N) \right\|_{L_{2p}} \end{aligned}$$

$$\leq \max_{i \in \{1, \dots, i^*(t^*)-1\}} \left\{ \sup_{t \in [i\mathbf{T}_s, (i+1)\mathbf{T}_s]} \left\| \tilde{F}_\sigma^\parallel(t, i\mathbf{T}_s, X_N) \right\|_{L_{2p}}, \sup_{t \in [i^*(t^*)\mathbf{T}_s, t^*]} \left\| \tilde{F}_\sigma^\parallel(t, i^*(t^*)\mathbf{T}_s, X_N) \right\|_{L_{2p}} \right\},$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 1}$. Similarly, using the definition of $\tilde{\mathcal{H}}_\mu$ from Proposition D.3, with $z = X_N$ and $\tau = t^*$, we get

$$\begin{aligned} & \left\| e^{(\lambda_s - \omega)t} \tilde{\mathcal{H}}_\mu(t, X_N) \right\|_{L_{2p}} \\ & \leq \Delta_\Theta \max_{i \in \{1, \dots, i^*(t^*)-1\}} \left\{ \sup_{t \in [(i-1)\mathbf{T}_s, i\mathbf{T}_s]} \left(\left| e^{(\lambda_s - \omega)t} \tilde{\gamma}_i(\mathbf{T}_s) \right| \left\| \tilde{F}_\sigma(t, i\mathbf{T}_s, X_N) \right\|_{L_{2p}} \right), \right. \\ & \quad \left. \sup_{t \in [(i^*(t^*)-1)\mathbf{T}_s, i^*(t^*)\mathbf{T}_s]} \left(\left| e^{(\lambda_s - \omega)t} \tilde{\gamma}^*(t^*, \mathbf{T}_s) \right| \left\| \tilde{F}_\sigma(t, i^*(t^*)\mathbf{T}_s, X_N) \right\|_{L_{2p}} \right) \right\}, \end{aligned}$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 1}$, where we have used the uniform bound on $\|\Theta_{ad}(t)\|_F$ from Assumption 1. It then follows from the bounds above and the Minkowski's inequality that

$$\left\| \tilde{\mathcal{H}}_\sigma^\parallel(t, X_N) + e^{(\lambda_s - \omega)t} \tilde{\mathcal{H}}_\sigma(t, X_N) \right\|_{L_{2p}} \leq \Psi_{\mathcal{H}_\sigma}^\parallel + \Psi_{\mathcal{H}_\sigma}, \quad \forall (t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 1}, \quad (\text{D.71})$$

where

$$\begin{aligned} \Psi_{\mathcal{H}_\sigma}^\parallel &= \max_{i \in \{1, \dots, i^*(t^*)-1\}} \left\{ \sup_{t \in [i\mathbf{T}_s, (i+1)\mathbf{T}_s]} \left\| \tilde{F}_\sigma^\parallel(t, i\mathbf{T}_s, X_N) \right\|_{L_{2p}}, \sup_{t \in [i^*(t^*)\mathbf{T}_s, t^*]} \left\| \tilde{F}_\sigma^\parallel(t, i^*(t^*)\mathbf{T}_s, X_N) \right\|_{L_{2p}} \right\}, \\ \Psi_{\mathcal{H}_\sigma} &= \Delta_\Theta \max_{i \in \{1, \dots, i^*(t^*)-1\}} \left\{ \sup_{t \in [(i-1)\mathbf{T}_s, i\mathbf{T}_s]} \left(\left| e^{(\lambda_s - \omega)t} \tilde{\gamma}_i(\mathbf{T}_s) \right| \left\| \tilde{F}_\sigma(t, i\mathbf{T}_s, X_N) \right\|_{L_{2p}} \right), \right. \\ & \quad \left. \sup_{t \in [(i^*(t^*)-1)\mathbf{T}_s, i^*(t^*)\mathbf{T}_s]} \left(\left| e^{(\lambda_s - \omega)t} \tilde{\gamma}^*(t^*, \mathbf{T}_s) \right| \left\| \tilde{F}_\sigma(t, i^*(t^*)\mathbf{T}_s, X_N) \right\|_{L_{2p}} \right) \right\}. \end{aligned}$$

Then, using (D.65b), along with (D.68) and (D.69), we obtain the following bounds:

$$\begin{aligned} \Psi_{\mathcal{H}_\sigma}^\parallel &\leq 2\Delta_p^\parallel (1 - \mathbb{1}_{\text{Lip}}(p)) + \mathbf{T}_s \left(\hat{L}_p^\parallel \mathbb{1}_{\text{Lip}}(p) + \hat{L}_\sigma^\parallel \right) \\ &\quad + (L_p^\parallel \mathbb{1}_{\text{Lip}}(p) + L_\sigma^\parallel) \left(\mathbf{T}_s \tilde{\delta}_1(\mathbf{T}_s, \mathbf{p}, t^*) + \mathbf{T}_s^{\frac{1}{2}} \tilde{\delta}_2(\mathbf{p}, t^*) \right)^{\frac{1}{2}}, \\ \Psi_{\mathcal{H}_\sigma} &\leq 2\Delta_p \gamma_2(\omega, \mathbf{T}_s) \Delta_\Theta (1 - \mathbb{1}_{\text{Lip}}(p)) + \mathbf{T}_s \gamma_2(\omega, \mathbf{T}_s) \Delta_\Theta \left(\hat{L}_p \mathbb{1}_{\text{Lip}}(p) + \hat{L}_\sigma \right) \\ &\quad + \gamma_2(\omega, \mathbf{T}_s) \Delta_\Theta (L_p \mathbb{1}_{\text{Lip}}(p) + L_\sigma) \left(\mathbf{T}_s \tilde{\delta}_1(\mathbf{T}_s, \mathbf{p}, t^*) + \mathbf{T}_s^{\frac{1}{2}} \tilde{\delta}_2(\mathbf{p}, t^*) \right)^{\frac{1}{2}}. \end{aligned}$$

Substituting the above bounds into (D.71) leads to (D.64b).

Next, with $z = X_N$ and $\tau = t^*$, we use the definition of $\tilde{\mathcal{G}}_{\sigma_1}$ and $\tilde{\mathcal{G}}_{\sigma_2}$ and the Minkowski's inequality to conclude that

$$\begin{aligned} \left\| \tilde{\mathcal{G}}_{\sigma_1}(t, X_N) \right\|_{L_{2p}} &\leq \mathbb{1}_{<\mathbf{T}_s}(t) \left\| \tilde{F}_\sigma^\parallel(t, 0, X_N) \right\|_{L_{2p}} + \mathbb{1}_{\geq \mathbf{T}_s}(t) 0 \\ &\leq \sup_{t \in [0, \mathbf{T}_s]} \left\| \tilde{F}_\sigma^\parallel(t, 0, X_N) \right\|_{L_{2p}}, \\ \left\| e^{(\lambda_s - \omega)t} \tilde{\mathcal{G}}_{\sigma_2}(t, X_N) \right\|_{L_{2p}} &\leq \sum_{i=1}^{i^*(t^*)-1} \mathbb{1}_{[(i-1)\mathbf{T}_s, i\mathbf{T}_s]}(t) \left| e^{(\lambda_s - \omega)t} \tilde{\gamma}_i(\mathbf{T}_s) \right| \left\| \tilde{F}_\sigma^\parallel(i\mathbf{T}_s, (i-1)\mathbf{T}_s, X_N) \right\|_{L_{2p}} \\ &\quad + \mathbb{1}_{\geq (i^*(t)-1)\mathbf{T}_s}(t) 0 \\ &\leq \max_{i \in \{1, \dots, i^*(t^*)-1\}} \sup_{t \in [(i-1)\mathbf{T}_s, i\mathbf{T}_s]} \left(\left| e^{(\lambda_s - \omega)t} \tilde{\gamma}_i(\mathbf{T}_s) \right| \left\| \tilde{F}_\sigma^\parallel(i\mathbf{T}_s, (i-1)\mathbf{T}_s, X_N) \right\|_{L_{2p}} \right) \end{aligned}$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 1}$. Hence, it follows from the Minkowski's inequality that

$$\left\| \tilde{\mathcal{G}}_{\sigma_1}(t, X_N) + e^{(\lambda_s - \omega)t} \tilde{\mathcal{G}}_{\sigma_2}(t, X_N) \right\|_{L_{2p}}$$

$$\leq \sup_{t \in [0, \mathbf{T}_s]} \left\| \tilde{F}_\sigma^\parallel(t, 0, X_N) \right\|_{L_{2p}} + \max_{i \in \{1, \dots, i^*(t^*)-1\}} \sup_{t \in [(i-1)\mathbf{T}_s, i\mathbf{T}_s]} \left(\left| e^{(\lambda_s - \omega)t} \tilde{\gamma}_i(\mathbf{T}_s) \right| \left\| \tilde{F}_\sigma^\parallel(i\mathbf{T}_s, (i-1)\mathbf{T}_s, X_N) \right\|_{L_{2p}} \right),$$

for all $(t, p) \in [0, t^*] \times \mathbb{N}_{\geq 1}$. Using (D.65b) and (D.68) leads to

$$\begin{aligned} & \left\| \tilde{\mathcal{G}}_{\sigma_1}(t, X_N) + e^{(\lambda_s - \omega)t} \tilde{\mathcal{G}}_{\sigma_2}(t, X_N) \right\|_{L_{2p}} \\ & \leq (1 + \gamma_2(\omega, \mathbf{T}_s)) \left[2\Delta_p^\parallel (1 - \mathbb{1}_{\text{Lip}}(p)) + (L_p^\parallel \mathbb{1}_{\text{Lip}}(p) + L_\sigma^\parallel) \left(\mathbf{T}_s \tilde{\delta}_1(\mathbf{T}_s, p, t^*) + \mathbf{T}_s^{\frac{1}{2}} \tilde{\delta}_2(p, t^*) \right)^{\frac{1}{2}} \right. \\ & \quad \left. + \mathbf{T}_s \left(\hat{L}_p^\parallel \mathbb{1}_{\text{Lip}}(p) + \hat{L}_\sigma^\parallel \right) \right], \end{aligned}$$

for all $(t, p) \in [0, t^*] \times \mathbb{N}_{\geq 1}$, thus establishing (D.64c).

Finally, using the definition of \mathcal{G}_μ , we deduce that

$$\begin{aligned} \left\| \mathcal{G}_\mu(t, X_N) \right\|_{L_{2p}} & \leq (1 - e^{-\lambda_s \mathbf{T}_s}) \\ & \quad \times \max_{i \in \{1, \dots, i^*(t^*)-1\}} \left\{ \sup_{t \in [i\mathbf{T}_s, (i+1)\mathbf{T}_s]} \left\| \Lambda_\mu^\parallel(t, X_{N,t}) \right\|_{L_{2p}}, \sup_{t \in [i^*(t^*)\mathbf{T}_s, t^*]} \left\| \Lambda_\mu^\parallel(t, X_{N,t}) \right\|_{L_{2p}} \right\}, \end{aligned}$$

for all $t \in [0, t^*]$. Therefore,

$$\left\| \mathcal{G}_\mu(t, X_N) \right\|_{L_{2p}} \leq (1 - e^{-\lambda_s \mathbf{T}_s}) \sup_{t \in [0, t^*]} \left\| \Lambda_\mu^\parallel(t, X_{N,t}) \right\|_{L_{2p}},$$

for all $t \in [0, t^*]$, thus establishing (D.64d). \square

The next result bounds the moments of a few pertinent processes that we need later.

Proposition D.12 *Let*

$$\begin{aligned} \mathfrak{Z}_j(t) &= \int_{\mathbf{T}_s}^t e^{2(\lambda - \omega)\nu} \left(\int_0^\nu e^{\omega\beta} \mathcal{J}_j(\beta, X_N) dW_\beta \right)^\top g(\nu)^\top g(\nu) \\ & \quad \times \left(\int_0^\nu e^{\omega\beta} \left[\dot{\Lambda}_\mu^\parallel(\beta, Z_{N,\beta}) d\beta + \dot{F}_\sigma^\parallel(\beta, Z_{N,\beta}) dW_\beta \right] \right) d\nu, \quad (j, t) \in \{1, \dots, 4\} \times [\mathbf{T}_s, t^*], \end{aligned} \quad (\text{D.72})$$

where the piecewise functions \mathcal{J}_j are defined as follows for $\beta \in [0, \nu]$, $\nu \in [\mathbf{T}_s, t]$, $t \geq \mathbf{T}_s$:

$$\mathcal{J}_1(\beta, X_N) = \sum_{i=1}^{i^*(\nu)-1} \mathbb{1}_{[(i-1)\mathbf{T}_s, i\mathbf{T}_s]}(\beta) F_\sigma^\parallel((i-1)\mathbf{T}_s, X_{N,(i-1)\mathbf{T}_s}) \left(1 - \tilde{\gamma}_i(\mathbf{T}_s) e^{(\lambda_s - \omega)\beta} \right), \quad (\text{D.73a})$$

$$\mathcal{J}_2(\beta, X_N) = \mathbb{1}_{[i^*(\nu)-1)\mathbf{T}_s, i^*(\nu)\mathbf{T}_s]}(\beta) F_\sigma^\parallel((i^*(\nu)-1)\mathbf{T}_s, X_{N,(i^*(\nu)-1)\mathbf{T}_s}), \quad (\text{D.73b})$$

$$\mathcal{J}_3(\beta, X_N) = -\mathbb{1}_{[i^*(\nu)-1)\mathbf{T}_s, i^*(\nu)\mathbf{T}_s]}(\beta) \tilde{\gamma}^*(\nu, \mathbf{T}_s) e^{(\lambda_s - \omega)\beta} F_\sigma^\parallel(i^*(\nu)\mathbf{T}_s, X_{N,i^*(\nu)\mathbf{T}_s}), \quad (\text{D.73c})$$

$$\mathcal{J}_4(\beta, X_N) = \mathbb{1}_{\geq i^*(\nu)\mathbf{T}_s}(\beta) F_\sigma^\parallel(i^*(\nu)\mathbf{T}_s, X_{N,i^*(\nu)\mathbf{T}_s}), \quad (\text{D.73d})$$

and the entities $i^*(\nu)$, $\tilde{\gamma}_i(\mathbf{T}_s)$, and $\tilde{\gamma}^*(\nu, \mathbf{T}_s)$ are defined in Proposition D.3 as:

$$i^*(\nu) \doteq \max \{i \in \mathbb{N}_{\geq 1} : i\mathbf{T}_s \leq \nu\},$$

$$\tilde{\gamma}_i(\mathbf{T}_s) \doteq e^{(\omega - \lambda_s)i\mathbf{T}_s} \frac{\lambda_s}{\omega} \frac{e^{\omega\mathbf{T}_s} - 1}{e^{\lambda_s\mathbf{T}_s} - 1}, \quad \tilde{\gamma}^*(\nu, \mathbf{T}_s) \doteq e^{(\omega - \lambda_s)i^*(\nu)\mathbf{T}_s} \frac{\lambda_s}{\omega} \frac{e^{\omega(\nu - i^*(\nu)\mathbf{T}_s)} - 1}{e^{\lambda_s\mathbf{T}_s} - 1}.$$

If the stopping time τ^* , defined in (104), Lemma 4.5, satisfies $\tau^* = t^*$ and $t^* \geq \mathbf{T}_s$, then

$$\begin{aligned} \sum_{j=1}^4 \left\| \mathcal{Z}_j(t) \right\|_{L_p} & \leq \Delta_g^2 \frac{e^{2\lambda t}}{2\lambda\omega} \gamma''(p, \omega, \mathbf{T}_s) \sup_{\nu \in [0, t^*]} \left\| F_\sigma^\parallel(\nu, X_{N,\nu}) \right\|_{L_{4p}} \\ & \quad \times \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^*]} \left\| \dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) \right\|_{L_{2p}} + p'(p) \sup_{\nu \in [0, t^*]} \left\| \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) \right\|_{L_{2p}} \right), \end{aligned} \quad (\text{D.74})$$

for all $(t, p) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$, where $\gamma''(p, \omega, \mathbf{T}_s)$ is defined in (A.13d).

Proof. We begin with the bound for $\mathfrak{Z}_1(t)$.

Bound for $\mathfrak{Z}_1(t)$: Invoking Lemma B.6 by setting

$$\begin{aligned}\xi &= \mathbf{T}_s, \quad \theta_1 = \lambda, \quad \theta_2 = \omega, \quad Q_t = W_t(n_q = d), \\ R(t) &= g(t)^\top g(t) \in \mathbb{R}^{m \times m} (n_1 = n_2 = m), \\ S_1(t) &= 0_m, \quad L_1(t) = \mathcal{J}_1(\beta, X_N), \quad S_2(t) = \mathring{\Lambda}_\mu^\parallel(t, Z_{N,t}), \quad L_2(t) = \mathring{F}_\sigma^\parallel(t, Z_{N,t}),\end{aligned}$$

we obtain

$$\begin{aligned}\|\mathfrak{Z}_1(t)\|_{L_p} &\leq \Delta_g^2 \frac{e^{2\lambda t}}{2\lambda\omega} \mathfrak{p}'(\mathfrak{p}) \sup_{\nu \in [0, t^*]} \|\mathcal{J}_1(\nu, X_N)\|_{L_{2p}} \\ &\quad \times \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^*]} \|\mathring{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu})\|_{L_{2p}} + \mathfrak{p}'(\mathfrak{p}) \sup_{\nu \in [0, t^*]} \|\mathring{F}_\sigma^\parallel(\nu, Z_{N,\nu})\|_{L_{2p}} \right), \quad (\text{D.75})\end{aligned}$$

for all $(t, \mathfrak{p}) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$, where we have used the bound $\Delta_R = \Delta_g^2 \geq \|g(t)\|_F^2 = \|R(t)\|_F$ due to Assumption 1. Using the definition of \mathcal{J}_1 and the Minkowski's inequality leads to

$$\|\mathcal{J}_1(\nu, X_N)\|_{L_{2p}} \leq \sum_{i=1}^{i^*(t)-1} \mathbf{1}_{[(i-1)\mathbf{T}_s, i\mathbf{T}_s]}(\nu) \left| 1 - \tilde{\gamma}_i(\mathbf{T}_s) e^{(\lambda_s - \omega)\nu} \right| \|F_\sigma^\parallel((i-1)\mathbf{T}_s, X_{N, (i-1)\mathbf{T}_s})\|_{L_{2p}},$$

for all $\nu \in [0, t]$ and $(t, \mathfrak{p}) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$. Hence,

$$\begin{aligned}\|\mathcal{J}_1(\nu, X_N)\|_{L_{2p}} &\leq \max_{i \in \{1, \dots, i^*(t^*)-1\}} \sup_{t \in [(i-1)\mathbf{T}_s, i\mathbf{T}_s]} \left(\left| 1 - \tilde{\gamma}_i(\mathbf{T}_s) e^{(\lambda_s - \omega)t} \right| \|F_\sigma^\parallel((i-1)\mathbf{T}_s, X_{N, (i-1)\mathbf{T}_s})\|_{L_{2p}} \right) \\ &\leq \left(\max_{i \in \{1, \dots, i^*(t^*)-1\}} \sup_{t \in [(i-1)\mathbf{T}_s, i\mathbf{T}_s]} \left| 1 - \tilde{\gamma}_i(\mathbf{T}_s) e^{(\lambda_s - \omega)t} \right| \right) \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{2p}} \\ &\leq \gamma'_2(\omega, \mathbf{T}_s) \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{2p}}, \quad (\text{D.76})\end{aligned}$$

for all $(t, \mathfrak{p}) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$, where, we have used the definition of $\tilde{\gamma}_i$ from Proposition D.3 to conclude that

$$\sup_{t \in [(i-1)\mathbf{T}_s, i\mathbf{T}_s]} \left| 1 - \tilde{\gamma}_i(\mathbf{T}_s) e^{(\lambda_s - \omega)t} \right| = \sup_{t \in [(i-1)\mathbf{T}_s, i\mathbf{T}_s]} \left| 1 - e^{(\omega - \lambda_s)(i\mathbf{T}_s - t)} \frac{\lambda_s}{\omega} \frac{e^{\omega\mathbf{T}_s} - 1}{e^{\lambda_s\mathbf{T}_s} - 1} \right| \leq \gamma'_2(\omega, \mathbf{T}_s),$$

for all $i \in \{1, \dots, i^*(t^*) - 1\}$. Substituting (D.76) into (D.75) leads to

$$\begin{aligned}\|\mathfrak{Z}_1(t)\|_{L_p} &\leq \Delta_g^2 \frac{e^{2\lambda t}}{2\lambda\omega} \mathfrak{p}'(\mathfrak{p}) \gamma'_2(\omega, \mathbf{T}_s) \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{4p}} \\ &\quad \times \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^*]} \|\mathring{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu})\|_{L_{2p}} + \mathfrak{p}'(\mathfrak{p}) \sup_{\nu \in [0, t^*]} \|\mathring{F}_\sigma^\parallel(\nu, Z_{N,\nu})\|_{L_{2p}} \right), \quad (\text{D.77})\end{aligned}$$

for all $(t, \mathfrak{p}) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$. Note that we have also used the bound $\|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{2p}} \leq \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{4p}}$ which is a consequence of the Jensen's inequality.

Bound for $\mathfrak{Z}_2(t)$: The definition of \mathcal{J}_2 allows us to write

$$\begin{aligned}\mathfrak{Z}_2(t) &= \int_{\mathbf{T}_s}^t e^{2(\lambda - \omega)\nu} \left(\int_0^\nu e^{\omega\beta} \mathcal{J}_2(\beta, X_N) dW_\beta \right)^\top g(\nu)^\top g(\nu) \\ &\quad \times \left(\int_0^\nu e^{\omega\beta} \left[\mathring{\Lambda}_\mu^\parallel(\beta, Z_{N,\beta}) d\beta + \mathring{F}_\sigma^\parallel(\beta, Z_{N,\beta}) dW_\beta \right] \right) d\nu \\ &= \int_{\mathbf{T}_s}^t e^{2(\lambda - \omega)\nu} \left(F_\sigma^\parallel((i^*(\nu) - 1)\mathbf{T}_s, X_{N, (i^*(\nu) - 1)\mathbf{T}_s}) \int_{(i^*(\nu) - 1)\mathbf{T}_s}^{i^*(\nu)\mathbf{T}_s} e^{\omega\beta} dW_\beta \right)^\top g(\nu)^\top g(\nu) \\ &\quad \times \left(\int_0^\nu e^{\omega\beta} \left[\mathring{\Lambda}_\mu^\parallel(\beta, Z_{N,\beta}) d\beta + \mathring{F}_\sigma^\parallel(\beta, Z_{N,\beta}) dW_\beta \right] \right) d\nu,\end{aligned}$$

for $t \in [\mathbf{T}_s, t^*]$, This process can be cast as $N_a(t)$ in Corollary B.5 by setting

$$\begin{aligned} \theta_1 &= \lambda, \quad \theta_2 = \omega, \quad Q_t = W_t(n_q = d), \\ R(t) &= g(t)^\top g(t) \in \mathbb{R}^{m \times m} (n_1 = n_2 = m), \\ L_a(\nu) &= F_\sigma^\parallel((i^*(\nu) - 1)\mathbf{T}_s, X_{N, (i^*(\nu) - 1)\mathbf{T}_s}) \in \mathbb{R}^{m \times d}, \quad \varsigma_1(\nu) = \varsigma_2(\beta) = 1, \\ S(t) &= \mathring{A}_\mu^\parallel(t, Z_{N,t}), \quad L(t) = \mathring{F}_\sigma^\parallel(t, Z_{N,t}). \end{aligned}$$

Hence, by (B.51) in Corollary B.5, we have

$$\begin{aligned} \|Z_2(t)\|_{L_p} &\leq \Delta_g^2 \mathbf{p}''(\mathbf{p}) \frac{e^{2\lambda t}}{2\lambda\omega} (1 - e^{-2\omega\mathbf{T}_s})^{\frac{1}{2}} \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{4p}} \\ &\quad \times \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^*]} \|\mathring{A}_\mu^\parallel(\nu, Z_{N,\nu})\|_{L_{2p}} + \mathbf{p}'(\mathbf{p}) \sup_{\nu \in [0, t^*]} \|\mathring{F}_\sigma^\parallel(\nu, Z_{N,\nu})\|_{L_{2p}} \right), \quad (\text{D.78}) \end{aligned}$$

for all $(t, \mathbf{p}) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$, where we have used the fact that

$$\begin{aligned} &\sup_{\nu \in [\mathbf{T}_s, t]} \|F_\sigma^\parallel((i^*(\nu) - 1)\mathbf{T}_s, X_{N, (i^*(\nu) - 1)\mathbf{T}_s})\|_{L_{4p}} \\ &\leq \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel((i^*(\nu) - 1)\mathbf{T}_s, X_{N, (i^*(\nu) - 1)\mathbf{T}_s})\|_{L_{4p}} \leq \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{4p}}, \quad \forall t \in [\mathbf{T}_s, t^*]. \end{aligned}$$

Bound for $\mathfrak{Z}_3(t)$: Similar to the above, from the definition of \mathcal{J}_3 , we have:

$$\begin{aligned} \mathfrak{Z}_3(t) &= \int_{\mathbf{T}_s}^t e^{2(\lambda - \omega)\nu} \left(\int_0^\nu e^{\omega\beta} \mathcal{J}_3(\beta, X_N) dW_\beta \right)^\top g(\nu)^\top g(\nu) \\ &\quad \times \left(\int_0^\nu e^{\omega\beta} \left[\mathring{A}_\mu^\parallel(\beta, Z_{N,\beta}) d\beta + \mathring{F}_\sigma^\parallel(\beta, Z_{N,\beta}) dW_\beta \right] \right) d\nu \\ &= \int_{\mathbf{T}_s}^t e^{2(\lambda - \omega)\nu} \left(-F_\sigma^\parallel(i^*(\nu)\mathbf{T}_s, X_{N, i^*(\nu)\mathbf{T}_s}) \tilde{\gamma}^*(\nu, \mathbf{T}_s) \int_{(i^*(\nu) - 1)\mathbf{T}_s}^{i^*(\nu)\mathbf{T}_s} e^{\omega\beta} e^{(\lambda_s - \omega)\beta} dW_\beta \right)^\top g(\nu)^\top g(\nu) \\ &\quad \times \left(\int_0^\nu e^{\omega\beta} \left[\mathring{A}_\mu^\parallel(\beta, Z_{N,\beta}) d\beta + \mathring{F}_\sigma^\parallel(\beta, Z_{N,\beta}) dW_\beta \right] \right) d\nu, \end{aligned}$$

for all $t \in [\mathbf{T}_s, t^*]$. We can cast the above process as $N_a(t)$ in Corollary B.5 by setting

$$\begin{aligned} \theta_1 &= \lambda, \quad \theta_2 = \omega, \quad Q_t = W_t(n_q = d), \\ R(t) &= g(t)^\top g(t) \in \mathbb{R}^{m \times m} (n_1 = n_2 = m), \\ L_a(\nu) &= -F_\sigma^\parallel(i^*(\nu)\mathbf{T}_s, X_{N, i^*(\nu)\mathbf{T}_s}) \in \mathbb{R}^{m \times d}, \quad \varsigma_1(\nu) = \tilde{\gamma}^*(\nu, \mathbf{T}_s), \quad \varsigma_2(\beta) = e^{(\lambda_s - \omega)\beta}, \\ S(t) &= \mathring{A}_\mu^\parallel(t, Z_{N,t}), \quad L(t) = \mathring{F}_\sigma^\parallel(t, Z_{N,t}). \end{aligned}$$

Hence, we use (B.51) in Corollary B.5 to obtain

$$\begin{aligned} \|\mathfrak{Z}_3(t)\|_{L_p} &\leq \Delta_g^2 \mathbf{p}''(\mathbf{p}) \frac{e^{2\lambda t}}{2\lambda\omega} (1 - e^{-2\omega\mathbf{T}_s})^{\frac{1}{2}} \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{4p}} \\ &\quad \times \sup_{\nu \in [\mathbf{T}_s, t]} \left(|\tilde{\gamma}^*(\nu, \mathbf{T}_s)| \sup_{\beta \in [(i^*(\nu) - 1)\mathbf{T}_s, i^*(\nu)\mathbf{T}_s]} e^{(\lambda_s - \omega)\beta} \right) \\ &\quad \times \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^*]} \|\mathring{A}_\mu^\parallel(\nu, Z_{N,\nu})\|_{L_{2p}} + \mathbf{p}'(\mathbf{p}) \sup_{\nu \in [0, t^*]} \|\mathring{F}_\sigma^\parallel(\nu, Z_{N,\nu})\|_{L_{2p}} \right), \quad (\text{D.79}) \end{aligned}$$

for all $(t, \mathbf{p}) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$, where we have used the fact that

$$\sup_{\nu \in [\mathbf{T}_s, t]} \|F_\sigma^\parallel(i^*(\nu)\mathbf{T}_s, X_{N, i^*(\nu)\mathbf{T}_s})\|_{L_{4p}}$$

$$\leq \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(i^*(\nu) \mathbf{T}_s, X_{N, i^*(\nu) \mathbf{T}_s})\|_{L_{4p}} \leq \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N, \nu})\|_{L_{4p}}, \quad \forall t \in [\mathbf{T}_s, t^*]. \quad (\text{D.80})$$

We now use the definition of $\tilde{\gamma}^*$ to obtain

$$\begin{aligned} |\tilde{\gamma}^*(\nu, \mathbf{T}_s)| &= \sup_{\beta \in [(i^*(\nu)-1)\mathbf{T}_s, i^*(\nu)\mathbf{T}_s]} e^{(\lambda_s - \omega)\beta} \\ &= e^{(\omega - \lambda_s)i^*(\nu)\mathbf{T}_s} \frac{\lambda_s}{\omega} \frac{e^{\omega(\nu - i^*(\nu)\mathbf{T}_s)} - 1}{e^{\lambda_s \mathbf{T}_s} - 1} \sup_{\beta \in [(i^*(\nu)-1)\mathbf{T}_s, i^*(\nu)\mathbf{T}_s]} e^{(\lambda_s - \omega)\beta} \\ &\leq e^{(\omega - \lambda_s)i^*(\nu)\mathbf{T}_s} \frac{\lambda_s}{\omega} \frac{e^{\omega \mathbf{T}_s} - 1}{e^{\lambda_s \mathbf{T}_s} - 1} \sup_{\beta \in [(i^*(\nu)-1)\mathbf{T}_s, i^*(\nu)\mathbf{T}_s]} e^{(\lambda_s - \omega)\beta}, \quad \forall \nu \in [\mathbf{T}_s, t^*], \end{aligned}$$

where the last inequality is due to $\nu - i^*(\nu) \mathbf{T}_s \leq \mathbf{T}_s, \forall \nu \geq 0$. Hence, if $\omega \geq \lambda_s$, then the inequality above can be further generalized into

$$\begin{aligned} |\tilde{\gamma}^*(\nu, \mathbf{T}_s)| &\leq \sup_{\beta \in [(i^*(\nu)-1)\mathbf{T}_s, i^*(\nu)\mathbf{T}_s]} e^{(\lambda_s - \omega)\beta} \leq e^{(\omega - \lambda_s)i^*(\nu)\mathbf{T}_s} \frac{\lambda_s}{\omega} \frac{e^{\omega \mathbf{T}_s} - 1}{e^{\lambda_s \mathbf{T}_s} - 1} e^{(\lambda_s - \omega)(i^*(\nu)-1)\mathbf{T}_s} \\ &= e^{(\omega - \lambda_s)\mathbf{T}_s} \frac{\lambda_s}{\omega} \frac{e^{\omega \mathbf{T}_s} - 1}{e^{\lambda_s \mathbf{T}_s} - 1}, \quad \forall \nu \in [\mathbf{T}_s, t^*]. \end{aligned}$$

If instead $\omega < \lambda_s$, then

$$\begin{aligned} |\tilde{\gamma}^*(\nu, \mathbf{T}_s)| &\leq \sup_{\beta \in [(i^*(\nu)-1)\mathbf{T}_s, i^*(\nu)\mathbf{T}_s]} e^{(\lambda_s - \omega)\beta} \leq e^{(\omega - \lambda_s)i^*(\nu)\mathbf{T}_s} \frac{\lambda_s}{\omega} \frac{e^{\omega \mathbf{T}_s} - 1}{e^{\lambda_s \mathbf{T}_s} - 1} e^{(\lambda_s - \omega)i^*(\nu)\mathbf{T}_s} \\ &= \frac{\lambda_s}{\omega} \frac{e^{\omega \mathbf{T}_s} - 1}{e^{\lambda_s \mathbf{T}_s} - 1}, \quad \forall \nu \in [\mathbf{T}_s, t^*]. \end{aligned}$$

We therefore conclude that

$$|\tilde{\gamma}^*(\nu, \mathbf{T}_s)| \leq \sup_{\beta \in [(i^*(\nu)-1)\mathbf{T}_s, i^*(\nu)\mathbf{T}_s]} e^{(\lambda_s - \omega)\beta} \leq \max\{e^{(\omega - \lambda_s)\mathbf{T}_s}, 1\} \frac{\lambda_s}{\omega} \frac{e^{\omega \mathbf{T}_s} - 1}{e^{\lambda_s \mathbf{T}_s} - 1} \doteq \gamma_2(\omega, \mathbf{T}_s), \quad (\text{D.81})$$

for all $\nu \in [\mathbf{T}_s, t^*]$. Substituting the above into (D.79) then leads to

$$\begin{aligned} \|Z_3(t)\|_{L_p} &\leq \Delta_g^2 \mathbf{p}''(\mathbf{p}) \frac{e^{2\lambda t}}{2\lambda \omega} \gamma_2(\omega, \mathbf{T}_s) (1 - e^{-2\omega \mathbf{T}_s})^{\frac{1}{2}} \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N, \nu})\|_{L_{4p}} \\ &\quad \times \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^*]} \|\hat{\Lambda}_\mu^\parallel(\nu, Z_{N, \nu})\|_{L_{2p}} + \mathbf{p}'(\mathbf{p}) \sup_{\nu \in [0, t^*]} \|\hat{F}_\sigma^\parallel(\nu, Z_{N, \nu})\|_{L_{2p}} \right), \quad (\text{D.82}) \end{aligned}$$

for all $(t, \mathbf{p}) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$.

Bound for $\mathfrak{Z}_4(t)$: Finally, we have

$$\begin{aligned} \mathfrak{Z}_4(t) &= \int_{\mathbf{T}_s}^t e^{2(\lambda - \omega)\nu} \left(\int_0^\nu e^{\omega\beta} \mathcal{J}_4(\beta, X_N) dW_\beta \right)^\top g(\nu)^\top g(\nu) \\ &\quad \times \left(\int_0^\nu e^{\omega\beta} \left[\hat{\Lambda}_\mu^\parallel(\beta, Z_{N, \beta}) d\beta + \hat{F}_\sigma^\parallel(\beta, Z_{N, \beta}) dW_\beta \right] \right) d\nu \\ &= \int_{\mathbf{T}_s}^t e^{2(\lambda - \omega)\nu} \left(F_\sigma^\parallel(i^*(\nu) \mathbf{T}_s, X_{N, i^*(\nu) \mathbf{T}_s}) \int_{i^*(\nu) \mathbf{T}_s}^\nu e^{\omega\beta} dW_\beta \right)^\top g(\nu)^\top g(\nu) \\ &\quad \times \left(\int_0^\nu e^{\omega\beta} \left[\hat{\Lambda}_\mu^\parallel(\beta, Z_{N, \beta}) d\beta + \hat{F}_\sigma^\parallel(\beta, Z_{N, \beta}) dW_\beta \right] \right) d\nu, \end{aligned}$$

for all $t \in [\mathbf{T}_s, t^*]$, which we can use to cast the above process as $N_b(t)$ in Corollary B.5 by setting

$$\begin{aligned} \theta_1 &= \lambda, \quad \theta_2 = \omega, \quad Q_t = W_t(n_q = d), \\ R(t) &= g(t)^\top g(t) \in \mathbb{R}^{m \times m} (n_1 = n_2 = m), \end{aligned}$$

$$L_b(\nu) = F_\sigma^\parallel(i^*(\nu) \mathbf{T}_s, X_{N,i^*(\nu)} \mathbf{T}_s) \in \mathbb{R}^{m \times d}, \quad S(t) = \mathring{\Lambda}_\mu^\parallel(t, Z_{N,t}), \quad L(t) = \mathring{F}_\sigma^\parallel(t, Z_{N,t}).$$

Hence, we use (B.52) in Corollary B.5, and the bound in (D.80) to obtain

$$\begin{aligned} \|Z_4(t)\|_{L_p} &\leq \Delta_g^2 \mathbf{p}''(\mathbf{p}) \frac{e^{2\lambda t}}{2\lambda \omega} (1 - e^{-2\omega \mathbf{T}_s})^{\frac{1}{2}} \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{4p}} \\ &\quad \times \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^*]} \|\mathring{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu})\|_{L_{2p}} + \mathbf{p}'(\mathbf{p}) \sup_{\nu \in [0, t^*]} \|\mathring{F}_\sigma^\parallel(\nu, Z_{N,\nu})\|_{L_{2p}} \right), \end{aligned} \quad (\text{D.83})$$

for all $(t, \mathbf{p}) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$. The proof is then concluded by adding together the bounds in (D.77), (D.78), (D.82), and (D.83). \square

Using the bounds established in the last proposition, the following result establishes the effect of the feedback error operator $\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r$ on the truncated joint process in Definition 9.

Lemma D.2 Consider the following scalar process:

$$N_{\mathcal{U}}(t) = \int_0^t e^{\omega \nu} \widetilde{M}_{\mathcal{U}}(t, \nu) \left[\mathring{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \mathring{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right] \in \mathbb{R}, \quad (\text{D.84})$$

where

$$\widetilde{M}_{\mathcal{U}}(t, \nu) = \int_\nu^t e^{(2\lambda - \omega)\beta} \widetilde{\mathcal{P}}_{\mathcal{U}}(\beta)^\top d\beta,$$

for $0 \leq \nu \leq t \leq T$, and where $\widetilde{\mathcal{P}}_{\mathcal{U}}(\beta) \in \mathbb{R}^m$ is defined in the statement of Proposition D.6. If the stopping time τ^* , defined in (104), Lemma 4.5, satisfies $\tau^* = t^*$, then

$$\|N_{\mathcal{U}}(t)\|_{L_p} \leq \frac{e^{2\lambda t}}{\lambda} \Delta_{\mathcal{U}}(\mathbf{p}, t^*, \omega, \mathbf{T}_s), \quad \forall (t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 1}, \quad (\text{D.85})$$

where

$$\Delta_{\mathcal{U}}(\mathbf{p}, t^*, \omega, \mathbf{T}_s) \doteq \Delta^-(\mathbf{p}, t^*, \omega, \mathbf{T}_s) + \mathbb{1}_{\geq \mathbf{T}_s}(t^*) \sum_{k=1}^3 \Delta_k^+(\mathbf{p}, t^*, \omega, \mathbf{T}_s),$$

and where

$$\begin{aligned} \Delta^-(\mathbf{p}, t^*, \omega, \mathbf{T}_s) &= \mathring{\Delta}(\mathbf{p}, t^*, \omega) \gamma_1'(\omega, \mathbf{T}_s) \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^*]} \|\mathring{\Lambda}_\mu^\parallel(\nu, X_{N,\nu})\|_{L_{2p}} + 2\mathbf{p}'(\mathbf{p}) \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{2p}} \right), \\ \Delta_1^+(\mathbf{p}, t^*, \omega, \mathbf{T}_s) &= \frac{1}{\sqrt{\omega}} \mathring{\Delta}(\mathbf{p}, t^*, \omega) \left(\Delta_{\widetilde{\mathcal{H}}_\mu}(\mathbf{T}_s, \mathbf{p}, t^*) + \Delta_{\mathcal{G}_\mu}(\mathbf{T}_s, t^*) + \sqrt{\omega} \mathbf{p}'(\mathbf{p}) \left[\Delta_{\widetilde{\mathcal{H}}_\sigma}(\mathbf{T}_s, \mathbf{p}, t^*) + \Delta_{\widetilde{\mathcal{G}}_\sigma}(\mathbf{T}_s, \mathbf{p}, t^*) \right] \right), \\ \Delta_2^+(\mathbf{p}, t^*, \omega, \mathbf{T}_s) &= \mathring{\Delta}(\mathbf{p}, t^*, \omega) \gamma''(\mathbf{p}, \omega, \mathbf{T}_s) \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{4p}}, \\ \Delta_3^+(\mathbf{p}, t^*, \omega, \mathbf{T}_s) &= \frac{e^{\omega \mathbf{T}_s} - 1}{\sqrt{\omega}} \mathring{\Delta}(\mathbf{p}, t^*, \omega) \sup_{\nu \in [0, t^*]} \|\mathring{\Lambda}_\mu^\parallel(\nu, X_{N,\nu})\|_{L_{2p}}. \end{aligned}$$

In the above, the terms $\Delta_{\widetilde{\mathcal{H}}_\mu}$, $\Delta_{\widetilde{\mathcal{H}}_\sigma}$, $\Delta_{\widetilde{\mathcal{G}}_\sigma}$, and $\Delta_{\mathcal{G}_\mu}$ are defined in Proposition D.11, the terms $\mathbf{p}'(\mathbf{p})$, $\mathbf{p}''(\mathbf{p})$, γ_1' , γ_2' , and γ'' are defined in Appendix A, and

$$\mathring{\Delta}(\mathbf{p}, t^*, \omega) = \Delta_g^2 \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^*]} \|\mathring{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu})\|_{L_{2p}} + \mathbf{p}'(\mathbf{p}) \sup_{\nu \in [0, t^*]} \|\mathring{F}_\sigma^\parallel(\nu, Z_{N,\nu})\|_{L_{2p}} \right).$$

Proof. Assume w.l.o.g. that $\mathbf{T}_s < T$. Then, as in (D.55) in the proof of Proposition D.9, we obtain

$$N_{\mathcal{U}}(t) = \int_0^t e^{(2\lambda - \omega)\nu} \widetilde{\mathcal{P}}_{\mathcal{U}}(\nu)^\top \left(\int_0^\nu e^{\omega \beta} \left[\mathring{\Lambda}_\mu^\parallel(\beta, Z_{N,\beta}) d\beta + \mathring{F}_\sigma^\parallel(\beta, Z_{N,\beta}) dW_\beta \right] \right) d\nu, \quad t \in [0, T].$$

Using the definition of $\tilde{\mathcal{P}}_{\mathcal{U}}$ in Proposition D.6, we get

$$\begin{aligned} N_{\mathcal{U}}(t) &= 2 \int_0^t e^{(2\lambda - \omega)\nu} [(\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r)(X_N)(\nu)]^\top g(\nu)^\top g(\nu) \left(\int_0^\nu e^{\omega\beta} [\dot{\Lambda}_\mu^\parallel(\beta, Z_{N,\beta}) d\beta + \dot{F}_\sigma^\parallel(\beta, Z_{N,\beta}) dW_\beta] \right) d\nu \\ &= 2\omega \int_0^t e^{2(\lambda - \omega)\nu} \tilde{\mathcal{F}}(X_N, \nu)^\top g(\nu)^\top g(\nu) \left(\int_0^\nu e^{\omega\beta} [\dot{\Lambda}_\mu^\parallel(\beta, Z_{N,\beta}) d\beta + \dot{F}_\sigma^\parallel(\beta, Z_{N,\beta}) dW_\beta] \right) d\nu, \end{aligned} \quad (\text{D.86})$$

for $t \in [0, T]$, where the last line is due to the conclusion in Proposition D.1 that $(\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r)(X_N)(\nu) = \omega e^{-\omega\nu} \tilde{\mathcal{F}}(X_N, \nu)$.

To obtain the expression for $\tilde{\mathcal{F}}(X_N, \nu)$, we invoke Proposition D.3 with $z = X_N$ to obtain:

$$\tilde{\mathcal{F}}(X_N, \nu) = \int_0^\nu e^{\omega\beta} \Sigma_\beta^\parallel(X_N), \quad \nu < T_s,$$

and for $\nu \geq T_s$,

$$\begin{aligned} \tilde{\mathcal{F}}(X_N, \nu) &= \int_0^{T_s} e^{\omega\beta} \Lambda_\mu^\parallel(\beta, X_{N,\beta}) d\beta \\ &\quad + \int_0^\nu e^{\omega\beta} \left[\left(\tilde{\mathcal{H}}_\mu^\parallel(\beta, X_N) + e^{(\lambda_s - \omega)\beta} \tilde{\mathcal{H}}_\mu(\beta, X_N) \right) d\beta + \left(\tilde{\mathcal{H}}_\sigma^\parallel(\beta, X_N) + e^{(\lambda_s - \omega)\beta} \tilde{\mathcal{H}}_\sigma(\beta, X_N) \right) dW_\beta \right] \\ &\quad + \int_0^\nu e^{\omega\beta} \left(\tilde{\mathcal{G}}_{\sigma_1}(\beta, X_N) + e^{(\lambda_s - \omega)\beta} \tilde{\mathcal{G}}_{\sigma_2}(\beta, X_N) \right) dW_\beta \\ &\quad + \int_0^\nu e^{\omega\beta} [\mathcal{G}_\mu(\beta, X_N) d\beta + (\mathcal{G}_{\sigma_1}(\beta, X_N) + \mathcal{G}_{\sigma_2}(\beta, X_N)) dW_\beta]. \end{aligned}$$

Substituting the above expression for $\tilde{\mathcal{F}}(X_N, \nu)$ into (D.86):

$$N_{\mathcal{U}}(t) = \begin{cases} 2\omega \int_0^t e^{2(\lambda - \omega)\nu} \mathcal{N}_*(\nu)^\top g(\nu)^\top g(\nu) \dot{\mathcal{N}}(\nu) d\nu, & \text{if } t < T_s, \\ 2\omega \int_0^{T_s} e^{2(\lambda - \omega)\nu} \mathcal{N}_*(\nu)^\top g(\nu)^\top g(\nu) \dot{\mathcal{N}}(\nu) d\nu \\ + 2\omega \int_{T_s}^t e^{2(\lambda - \omega)\nu} (\mathcal{N}_1(\nu) + \mathcal{N}_2(\nu) + \mathcal{N}_3(\nu))^\top \\ \times g(\nu)^\top g(\nu) \dot{\mathcal{N}}(\nu) d\nu & , \text{ if otherwise } t \geq T_s, \end{cases} \quad (\text{D.87})$$

where

$$\begin{aligned} \mathcal{N}_*(\nu) &= \int_0^\nu e^{\omega\beta} \Sigma_\beta^\parallel(X_N) = \int_0^\nu e^{\omega\beta} [\Lambda_\mu^\parallel(\beta, X_{N,\beta}) d\beta + F_\sigma^\parallel(\beta, X_{N,\beta}) dW_\beta], \\ \mathcal{N}_1(\nu) &= \int_0^\nu e^{\omega\beta} \left[\left(\tilde{\mathcal{H}}_\mu^\parallel(\beta, X_N) + e^{(\lambda_s - \omega)\beta} \tilde{\mathcal{H}}_\mu(\beta, X_N) + \mathcal{G}_\mu(\beta, X_N) \right) d\beta \right. \\ &\quad \left. + \left(\tilde{\mathcal{H}}_\sigma^\parallel(\beta, X_N) + e^{(\lambda_s - \omega)\beta} \tilde{\mathcal{H}}_\sigma(\beta, X_N) + \tilde{\mathcal{G}}_{\sigma_1}(\beta, X_N) + e^{(\lambda_s - \omega)\beta} \tilde{\mathcal{G}}_{\sigma_2}(\beta, X_N) \right) dW_\beta \right], \\ \mathcal{N}_2(\nu) &= \int_0^\nu e^{\omega\beta} (\mathcal{G}_{\sigma_1}(\beta, X_N) + \mathcal{G}_{\sigma_2}(\beta, X_N)) dW_\beta, \\ \mathcal{N}_3 &= \int_0^{T_s} e^{\omega\beta} \Lambda_\mu^\parallel(\beta, X_{N,\beta}) d\beta, \\ \dot{\mathcal{N}}(\nu) &= \int_0^\nu e^{\omega\beta} [\dot{\Lambda}_\mu^\parallel(\beta, Z_{N,\beta}) d\beta + \dot{F}_\sigma^\parallel(\beta, Z_{N,\beta}) dW_\beta]. \end{aligned}$$

We can further write (D.87) as

$$N_{\mathcal{U}}(t) = \begin{cases} 2\omega \mathfrak{N}_*(t), & \text{if } t < T_s, \\ 2\omega \left(\mathfrak{N}_*(T_s) + \sum_{k=1}^3 \mathfrak{N}_k(t) \right), & \text{if otherwise } t \geq T_s, \end{cases} \quad (\text{D.88})$$

where we have defined

$$\begin{aligned}\mathfrak{N}_\star(t) &= \int_0^t e^{2(\lambda-\omega)\nu} \mathcal{N}_\star(\nu)^\top g(\nu)^\top g(\nu) \dot{\mathcal{N}}(\nu) d\nu, \\ \mathfrak{N}_k(t) &= \int_{\mathbf{T}_s}^t e^{2(\lambda-\omega)\nu} \mathcal{N}_k(\nu)^\top g(\nu)^\top g(\nu) \dot{\mathcal{N}}(\nu) d\nu, \quad k \in \{1, 2, 3\}.\end{aligned}$$

Using the Minkowski's inequality along with the fact that t^\star and \mathbf{T}_s are constants leads to the following bounds for all $\mathbf{p} \in \mathbb{N}_{\geq 1}$:

$$\|N_{\mathcal{U}}(t)\|_{L_{\mathbf{p}}} \leq \begin{cases} 2\omega \|\mathfrak{N}_\star(t)\|_{L_{\mathbf{p}}}, & \forall t \in [0, t^\star] \cap [0, \mathbf{T}_s], \\ 2\omega \left(\|\mathfrak{N}_\star(\mathbf{T}_s)\|_{L_{\mathbf{p}}} + \sum_{k=1}^3 \|\mathfrak{N}_k(t)\|_{L_{\mathbf{p}}} \right), & \forall t \in [0, t^\star] \cap [\mathbf{T}_s, T]. \end{cases} \quad (\text{D.89})$$

We now bound the processes $\mathfrak{N}_\star(t)$ and $\mathfrak{N}_k(t)$, $k \in \{1, 2, 3\}$.

Bound for $\mathfrak{N}_\star(t)$: Using the definitions in (D.87)-(D.88), we have

$$\begin{aligned}\mathfrak{N}_\star(t) &= \int_0^t e^{2(\lambda-\omega)\nu} \mathcal{N}_\star(\nu)^\top g(\nu)^\top g(\nu) \dot{\mathcal{N}}(\nu) d\nu \\ &= \int_0^t e^{2(\lambda-\omega)\nu} \left(\int_0^\nu e^{\omega\beta} \left[\Lambda_\mu^\parallel(\beta, X_{N,\beta}) d\beta + F_\sigma^\parallel(\beta, X_{N,\beta}) dW_\beta \right] \right)^\top g(\nu)^\top g(\nu) \\ &\quad \times \left(\int_0^\nu e^{\omega\beta} \left[\dot{\Lambda}_\mu^\parallel(\beta, Z_{N,\beta}) d\beta + \dot{F}_\sigma^\parallel(\beta, Z_{N,\beta}) dW_\beta \right] \right) d\nu,\end{aligned}$$

for all $(t, \mathbf{p}) \in [0, t^\star \wedge \mathbf{T}_s] \times \mathbb{N}_{\geq 1}$. Therefore, we can cast the process above as $N(t)$ in Lemma B.6 by setting

$$\begin{aligned}\xi &= 0, \quad \theta_1 = \lambda, \quad \theta_2 = \omega, \quad Q_t = W_t(n_q = d), \\ R(t) &= g(t)^\top g(t) \in \mathbb{R}^{m \times m} (n_1 = n_2 = m), \\ S_1(t) &= \Lambda_\mu^\parallel(t, X_{N,t}), \quad L_1(t) = F_\sigma^\parallel(t, X_{N,t}), \quad S_2(t) = \dot{\Lambda}_\mu^\parallel(t, Z_{N,t}), \quad L_2(t) = \dot{F}_\sigma^\parallel(t, Z_{N,t}).\end{aligned}$$

Thus, we apply Lemma B.6 and obtain the following upon re-arranging terms:

$$\begin{aligned}\|\mathfrak{N}_\star(t)\|_{L_{\mathbf{p}}} &\leq \Delta_g^2 \frac{e^{2\lambda t}}{2\lambda\omega} \left(1 - e^{-2\lambda(t^\star \wedge \mathbf{T}_s)} \right) \left(1 - e^{-\omega(t^\star \wedge \mathbf{T}_s)} \right) \\ &\quad \times \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^\star \wedge \mathbf{T}_s]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu})\|_{L_{2\mathbf{p}}} + 2\mathbf{p}'(\mathbf{p}) \sup_{\nu \in [0, t^\star \wedge \mathbf{T}_s]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{2\mathbf{p}}} \right) \\ &\quad \times \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^\star \wedge \mathbf{T}_s]} \|\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu})\|_{L_{2\mathbf{p}}} + 2\mathbf{p}'(\mathbf{p}) \sup_{\nu \in [0, t^\star \wedge \mathbf{T}_s]} \|\dot{F}_\sigma^\parallel(\nu, Z_{N,\nu})\|_{L_{2\mathbf{p}}} \right),\end{aligned}$$

for all $(t, \mathbf{p}) \in [0, t^\star \wedge \mathbf{T}_s] \times \mathbb{N}_{\geq 1}$, where we have used Assumption 1 to set $\Delta_R = \Delta_g^2$ in the hypothesis of Proposition B.6. We have further used the facts that $1 - e^{-\omega(t^\star \wedge \mathbf{T}_s)} \leq 1$ and $1 + e^{-\omega(t^\star \wedge \mathbf{T}_s)} \leq 2$, $\forall t^\star \wedge \mathbf{T}_s \geq 0$, after first using the following trivial manipulation:

$$\begin{aligned}\left(\frac{1 - e^{-2\omega(t^\star \wedge \mathbf{T}_s)}}{\omega} \right)^{\frac{1}{2}} &= \left(\frac{(1 - e^{-\omega(t^\star \wedge \mathbf{T}_s)}) (1 + e^{-\omega(t^\star \wedge \mathbf{T}_s)})}{\omega} \right)^{\frac{1}{2}} \\ &= \left(\frac{1 - e^{-\omega(t^\star \wedge \mathbf{T}_s)}}{\omega} \right)^{\frac{1}{2}} \left(1 + e^{-\omega(t^\star \wedge \mathbf{T}_s)} \right)^{\frac{1}{2}}. \quad (\text{D.90})\end{aligned}$$

Note that the facts $t^\star \wedge \mathbf{T}_s \leq t^\star$ and $t^\star \wedge \mathbf{T}_s \leq \mathbf{T}_s$ imply that $[0, t^\star \wedge \mathbf{T}_s] \subseteq [0, t^\star]$ and

$$\left(1 - e^{-2\lambda(t^\star \wedge \mathbf{T}_s)} \right) \left(1 - e^{-\omega(t^\star \wedge \mathbf{T}_s)} \right) \leq \left(1 - e^{-2\lambda\mathbf{T}_s} \right) \left(1 - e^{-\omega\mathbf{T}_s} \right) \doteq \gamma_1'(\omega, \mathbf{T}_s).$$

Hence, we can develop the previous bound into

$$\begin{aligned} \|\mathbf{N}_\star(t)\|_{L_p} &\leq \Delta_g^2 \frac{e^{2\lambda t}}{2\lambda \omega} \gamma'_1(\omega, \mathbf{T}_s) \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^\star]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu})\|_{L_{2p}} + 2\mathbf{p}'(\mathbf{p}) \sup_{\nu \in [0, t^\star]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{2p}} \right) \\ &\quad \times \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^\star]} \|\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu})\|_{L_{2p}} + 2\mathbf{p}'(\mathbf{p}) \sup_{\nu \in [0, t^\star]} \|\dot{F}_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{2p}} \right), \end{aligned}$$

which further implies

$$2\omega \|\mathfrak{N}_\star(t)\|_{L_p} \leq \frac{e^{2\lambda t}}{\lambda} \Delta^-(\mathbf{p}, t^\star, \omega, \mathbf{T}_s), \quad \forall (t, \mathbf{p}) \in [0, t^\star \wedge \mathbf{T}_s] \times \mathbb{N}_{\geq 1}. \quad (\text{D.91})$$

Bound for $\mathfrak{N}_1(t)$: The process \mathfrak{N}_1 in (D.88) can be written as

$$\begin{aligned} \tilde{\mathfrak{N}}_1(t) &= \int_{\mathbf{T}_s}^t e^{2(\lambda - \omega)\nu} \mathcal{N}_1(\nu)^\top g(\nu)^\top g(\nu) \dot{\mathcal{N}}(\nu) d\nu \\ &= \int_{\mathbf{T}_s}^t e^{2(\lambda - \omega)\nu} \left(\int_0^\nu e^{\omega\beta} [\mathcal{K}_\mu(\beta, X_N) d\beta + \mathcal{K}_\sigma(\beta, X_N) dW_\beta] \right)^\top g(\nu)^\top g(\nu) \\ &\quad \times \left(\int_0^\nu e^{\omega\beta} [\dot{\Lambda}_\mu^\parallel(\beta, Z_{N,\beta}) d\beta + \dot{F}_\sigma^\parallel(\beta, Z_{N,\beta}) dW_\beta] \right) d\nu, \end{aligned}$$

where we have defined

$$\mathcal{K}_\mu(\beta, X_N) = \tilde{\mathcal{H}}_\mu^\parallel(\beta, X_N) + e^{(\lambda_s - \omega)\beta} \tilde{\mathcal{H}}_\mu(\beta, X_N) + \mathcal{G}_\mu(\beta, X_N), \quad (\text{D.92a})$$

$$\mathcal{K}_\sigma(\beta, X_N) = \tilde{\mathcal{H}}_\sigma^\parallel(\beta, X_N) + e^{(\lambda_s - \omega)\beta} \tilde{\mathcal{H}}_\sigma(\beta, X_N) + \tilde{\mathcal{G}}_{\sigma_1}(\beta, X_N) + e^{(\lambda_s - \omega)\beta} \tilde{\mathcal{G}}_{\sigma_2}(\beta, X_N). \quad (\text{D.92b})$$

We apply Lemma B.6 by setting

$$\begin{aligned} \xi &= \mathbf{T}_s, \quad \theta_1 = \lambda, \quad \theta_2 = \omega, \quad Q_t = W_t(n_q = d), \\ R(t) &= g(t)^\top g(t) \in \mathbb{R}^{m \times m} (n_1 = n_2 = m), \end{aligned} \quad (\text{D.93})$$

$$S_1(t) = \mathcal{K}_\mu(t, X_N), \quad L_1(t) = \mathcal{K}_\sigma(t, X_N), \quad S_2(t) = \dot{\Lambda}_\mu^\parallel(t, Z_{N,t}), \quad L_2(t) = \dot{F}_\sigma^\parallel(t, Z_{N,t}),$$

to obtain the following bound:

$$\begin{aligned} \|\mathfrak{N}_1(t)\|_{L_p} &\leq \Delta_g^2 \frac{e^{2\lambda t}}{2\lambda \omega} \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^\star]} \|\mathcal{K}_\mu(\nu, X_{N,\nu})\|_{L_{2p}} + \mathbf{p}'(\mathbf{p}) \sup_{\nu \in [0, t^\star]} \|\mathcal{K}_\sigma(\nu, X_{N,\nu})\|_{L_{2p}} \right) \\ &\quad \times \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^\star]} \|\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu})\|_{L_{2p}} + \mathbf{p}'(\mathbf{p}) \sup_{\nu \in [0, t^\star]} \|\dot{F}_\sigma^\parallel(\nu, Z_{N,\nu})\|_{L_{2p}} \right), \end{aligned} \quad (\text{D.94})$$

for all $(t, \mathbf{p}) \in [\mathbf{T}_s, t^\star] \times \mathbb{N}_{\geq 1}$, where we have used the bound $\Delta_R = \Delta_g^2 \geq \|R(t)\|_F$ as above. Using the definitions of \mathcal{K}_μ and \mathcal{K}_σ in (D.92), we use the Minkowski's inequality and the bounds derived in Proposition D.11 to obtain the following:

$$\begin{aligned} \|\mathcal{K}_\mu(\nu, X_N)\|_{L_{2p}} &\leq \left\| \tilde{\mathcal{H}}_\mu^\parallel(\nu, X_N) + e^{(\lambda_s - \omega)\nu} \tilde{\mathcal{H}}_\mu(\nu, X_N) \right\|_{L_{2p}} + \|\mathcal{G}_\mu(\nu, X_N)\|_{L_{2p}} \\ &\leq \Delta_{\tilde{\mathcal{H}}_\mu}(\mathbf{T}_s, \mathbf{p}, t^\star) + \Delta_{\mathcal{G}_\mu}(\mathbf{T}_s, t^\star), \end{aligned}$$

and

$$\begin{aligned} \|\mathcal{K}_\sigma(\nu, X_N)\|_{L_{2p}} &\leq \left\| \tilde{\mathcal{H}}_\sigma^\parallel(\nu, X_N) + e^{(\lambda_s - \omega)\nu} \tilde{\mathcal{H}}_\sigma(\nu, X_N) \right\|_{L_{2p}} + \left\| \tilde{\mathcal{G}}_{\sigma_1}(\nu, X_N) + e^{(\lambda_s - \omega)\nu} \tilde{\mathcal{G}}_{\sigma_2}(\nu, X_N) \right\|_{L_{2p}} \\ &\leq \Delta_{\tilde{\mathcal{H}}_\sigma}(\mathbf{T}_s, \mathbf{p}, t^\star) + \Delta_{\tilde{\mathcal{G}}_\sigma}(\mathbf{T}_s, \mathbf{p}, t^\star), \end{aligned}$$

for all $(\nu, \mathbf{p}) \in [0, t^\star] \times \mathbb{N}_{\geq 1}$. Substituting the above bounds into (D.94) then leads to

$$2\omega \|\mathfrak{N}_1(t)\|_{L_p} \leq \frac{e^{2\lambda t}}{\lambda} \Delta_1^+(\mathbf{p}, t^\star, \omega, \mathbf{T}_s), \quad \forall (t, \mathbf{p}) \in [\mathbf{T}_s, t^\star] \times \mathbb{N}_{\geq 1}. \quad (\text{D.95})$$

Bound for $\mathfrak{N}_2(t)$: Recall the definition of the process \mathfrak{N}_2 in (D.88):

$$\begin{aligned}\mathfrak{N}_2(t) &= \int_{\mathbf{T}_s}^t e^{2(\lambda - \omega)\nu} \mathcal{N}_2(\nu)^\top g(\nu)^\top g(\nu) \dot{\mathcal{N}}(\nu) d\nu \\ &= \int_{\mathbf{T}_s}^t e^{2(\lambda - \omega)\nu} \left(\int_0^\nu e^{\omega\beta} (\mathcal{G}_{\sigma_1}(\beta, X_N) + \mathcal{G}_{\sigma_2}(\beta, X_N)) dW_\beta \right)^\top g(\nu)^\top g(\nu) \\ &\quad \times \left(\int_0^\nu e^{\omega\beta} \left[\mathring{\Lambda}_\mu^\parallel(\beta, Z_{N,\beta}) d\beta + \mathring{F}_\sigma^\parallel(\beta, Z_{N,\beta}) dW_\beta \right] \right) d\nu, \quad t \in [\mathbf{T}_s, t^*],\end{aligned}$$

where, we have the following due to the definitions of \mathcal{G}_{σ_1} and \mathcal{G}_{σ_2} in Proposition D.3:

$$\begin{aligned}\int_0^\nu e^{\omega\beta} \mathcal{G}_{\sigma_1}(\beta, X_N) dW_\beta &= \int_0^\nu e^{\omega\beta} \left(\sum_{i=1}^{i^*(\nu)-1} \mathbb{1}_{[(i-1)\mathbf{T}_s, i\mathbf{T}_s)}(\beta) F_\sigma^\parallel((i-1)\mathbf{T}_s, X_{N,(i-1)\mathbf{T}_s}) \left(1 - \tilde{\gamma}_i(\mathbf{T}_s) e^{(\lambda_s - \omega)\beta}\right) \right) dW_\beta \\ &\quad + \int_0^\nu e^{\omega\beta} \mathbb{1}_{[i^*(\nu)-1)\mathbf{T}_s, i^*(\nu)\mathbf{T}_s)}(\beta) F_\sigma^\parallel((i^*(\nu)-1)\mathbf{T}_s, X_{N,(i^*(\nu)-1)\mathbf{T}_s}) dW_\beta,\end{aligned}$$

and

$$\begin{aligned}\int_0^\nu e^{\omega\beta} \mathcal{G}_{\sigma_2}(\beta, X_N) dW_\beta &= - \int_0^\nu e^{\omega\beta} F_\sigma^\parallel(i^*(\nu)\mathbf{T}_s, X_{N,i^*(\nu)\mathbf{T}_s}) \mathbb{1}_{[i^*(\nu)-1)\mathbf{T}_s, i^*(\nu)\mathbf{T}_s)}(\beta) \tilde{\gamma}^*(\nu, \mathbf{T}_s) e^{(\lambda_s - \omega)\beta} dW_\beta \\ &\quad + \int_0^\nu e^{\omega\beta} F_\sigma^\parallel(i^*(\nu)\mathbf{T}_s, X_{N,i^*(\nu)\mathbf{T}_s}) \mathbb{1}_{\geq i^*(\nu)\mathbf{T}_s}(\beta) dW_\beta.\end{aligned}$$

Using the above, it is straightforward to establish that

$$\mathfrak{N}_2(t) = \sum_{j=1}^4 \mathfrak{Z}_j(t), \quad t \in [\mathbf{T}_s, t^*],$$

where the processes \mathfrak{Z}_j are defined in (D.72) in the statement of Proposition D.12. Hence, it follows from the Minkowski's inequality and Proposition D.12:

$$\begin{aligned}\|\mathfrak{N}_2(t)\|_{L_p} &\leq \sum_{j=1}^4 \|\mathfrak{Z}_j(t)\|_{L_p} \\ &\leq \Delta_g^2 \frac{e^{2\lambda t}}{2\lambda\omega} \left(\mathfrak{p}'(\mathfrak{p}) \gamma_2'(\omega, \mathbf{T}_s) + \mathfrak{p}''(\mathfrak{p}) (1 - e^{-2\omega\mathbf{T}_s})^{\frac{1}{2}} (2 + \gamma_2(\omega, \mathbf{T}_s)) \right) \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{4p}} \\ &\quad \times \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^*]} \|\mathring{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu})\|_{L_{2p}} + \mathfrak{p}'(\mathfrak{p}) \sup_{\nu \in [0, t^*]} \|\mathring{F}_\sigma^\parallel(\nu, Z_{N,\nu})\|_{L_{2p}} \right),\end{aligned}$$

for all $(t, p) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$. The above implies that

$$2\omega \|\mathfrak{N}_2(t)\|_{L_p} \leq \frac{e^{2\lambda t}}{\lambda} \Delta_2^+(\mathfrak{p}, t^*, \omega, \mathbf{T}_s), \quad \forall (t, \mathfrak{p}) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}. \quad (\text{D.96})$$

Bound for $\mathfrak{N}_3(t)$: Finally, using the definitions in (D.87)-(D.88), we have

$$\begin{aligned}\mathfrak{N}_3(t) &= \int_{\mathbf{T}_s}^t e^{2(\lambda - \omega)\nu} \mathcal{N}_3(\nu)^\top g(\nu)^\top g(\nu) \dot{\mathcal{N}}(\nu) d\nu \\ &= \int_{\mathbf{T}_s}^t e^{2(\lambda - \omega)\nu} \left(\int_0^{\mathbf{T}_s} e^{\omega\beta} \Lambda_\mu^\parallel(\beta, X_{N,\beta}) d\beta \right)^\top g(\nu)^\top g(\nu)\end{aligned}$$

$$\times \left(\int_0^\nu e^{\omega\beta} \left[\dot{\Lambda}_\mu^\parallel(\beta, Z_{N,\beta}) d\beta + \dot{F}_\sigma^\parallel(\beta, Z_{N,\beta}) dW_\beta \right] \right) d\nu,$$

for $(t, \mathbf{p}) \in [0, t^* \wedge T_s] \times \mathbb{N}_{\geq 1}$, which we can cast as $N_a(t)$ in Corollary B.4 by setting

$$\begin{aligned} \theta_1 &= \lambda, \quad \theta_2 = \omega, \quad \xi = T_s, \quad Q_t = W_t(n_q = d), \quad S_1(\beta) = \Lambda_\mu^\parallel(\beta, X_{N,\beta}) \in \mathbb{R}^m(n_1 = m), \\ R(\nu) &= g(\nu)^\top g(\nu) \in \mathbb{R}^{m \times m}(n_1 = n_2 = m), \\ S_2(\beta) &= \dot{\Lambda}_\mu^\parallel(\beta, Z_{N,\beta}) \in \mathbb{R}^m, \quad L_2(\beta) = \dot{F}_\sigma^\parallel(\beta, Z_{N,\beta}) \in \mathbb{R}^{m \times d}. \end{aligned}$$

Thus, we use (B.48) in Corollary B.4 and use the uniform bound on $g(t)$ from Assumption 1 to obtain

$$\begin{aligned} \|\mathfrak{N}_3(t)\|_{L_p} &\leq \Delta_g^2 \frac{e^{2\lambda t} e^{\omega T_s} - 1}{2\lambda \omega \sqrt{\omega}} \sup_{\nu \in [0, t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu})\|_{L_{2p}} \\ &\quad \times \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^*]} \|\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu})\|_{L_{2p}} + \mathbf{p}'(\mathbf{p}) \sup_{\nu \in [0, t^*]} \|\dot{F}_\sigma^\parallel(\nu, Z_{N,\nu})\|_{L_{2p}} \right), \end{aligned}$$

for all $(t, \mathbf{p}) \in [T_s, t^*] \times \mathbb{N}_{\geq 1}$, which allows us to write

$$2\omega \|\mathfrak{N}_3(t)\|_{L_p} \leq \frac{e^{2\lambda t}}{\lambda} \Delta_3^+(\mathbf{p}, t^*, \omega, T_s), \quad \forall (t, \mathbf{p}) \in [T_s, t^*] \times \mathbb{N}_{\geq 1}. \quad (\text{D.97})$$

Substituting the bounds (D.91), (D.95), (D.96) and (D.97) into (D.89) yields:

$$\|N_{\mathcal{U}}(t)\|_{L_p} \leq \frac{e^{2\lambda t}}{\lambda} \times \begin{cases} \Delta^-(\mathbf{p}, \omega, T_s), & \forall t \in [0, t^*] \cap [0, T_s], \\ \Delta^-(\mathbf{p}, \omega, T_s) + \sum_{k=1}^3 \Delta_k^+(\mathbf{p}, t^*, \omega, T_s), & \forall t \in [0, t^*] \cap [T_s, T], \end{cases}$$

which completes the proof by observing that the bound on the right hand side can be expressed as

$$\Delta_{\mathcal{U}}(\mathbf{p}, t^*, \omega, T_s) = \Delta^-(\mathbf{p}, \omega, T_s) + \mathbb{1}_{\geq T_s}(t^*) \sum_{k=1}^3 \Delta_k^+(\mathbf{p}, t^*, \omega, T_s).$$

□

We next derive the bound on the moments of $\Xi_{\mathcal{U}}$ in Lemma 4.5 using the results from Proposition D.10 to Lemma D.2 above.

Lemma D.3 *If the stopping time τ^* , defined in (104), Lemma 4.5, satisfies $\tau^* = t^*$, then the term $\Xi_{\mathcal{U}}(\tau(t), Z_N; \omega)$ defined in (107b), Lemma 4.5, satisfies the following bounds:*

$$\begin{aligned} &\mathbb{E} \left[e^{-(2\lambda + \omega)\tau(t)} \Xi_{\mathcal{U}}(\tau(t), Z_N; \omega) \right] \\ &\leq \frac{1}{|2\lambda - \omega|} \left(\Delta_{\mathcal{U}_1}(1, t^*) + \sqrt{\omega} \Delta_{\mathcal{U}_2}(1, t^*) + \omega \Delta_{\mathcal{U}_3}(1, t^*) + \frac{\omega}{\lambda} \Delta_{\mathcal{U}}(1, t^*, \omega, T_s) \right), \quad \forall t \in [0, t^*], \quad (\text{D.98}) \end{aligned}$$

and

$$\begin{aligned} &\left\| e^{-(2\lambda + \omega)\tau(t)} \Xi_{\mathcal{U}}(\tau(t), Z_N; \omega) \right\|_{L_p} \\ &\leq \frac{1}{|2\lambda - \omega|} \left(\Delta_{\mathcal{U}_1}(\mathbf{p}, t^*) + \sqrt{\omega} \Delta_{\mathcal{U}_2}(\mathbf{p}, t^*) + \omega \Delta_{\mathcal{U}_3}(\mathbf{p}, t^*) + \frac{\omega}{\lambda} \Delta_{\mathcal{U}}(\mathbf{p}, t^*, \omega, T_s) \right), \quad (\text{D.99}) \end{aligned}$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 2}$, where

$$\begin{aligned} \Delta_{\mathcal{U}_1}(\mathbf{p}, t^*) &= \sqrt{\lambda} \Delta_g \mathbf{p}(\mathbf{p}) \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \sup_{\nu \in [0, t^*]} \left\| \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) \right\|_{L_{2p}} \\ &\quad + \left(\frac{\Delta_{\mathcal{P}_1}(\mathbf{p}, t^*)}{\sqrt{\lambda}} + 2\Delta_g \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \right. \\ &\quad \left. + \frac{\Delta_g^2}{\lambda} \sup_{\nu \in [0, t^*]} \left\| \dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) \right\|_{L_{2p}} \right) \sup_{\nu \in [0, t^*]} \left\| \dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) \right\|_{L_{2p}}, \end{aligned}$$

$$\begin{aligned}\Delta_{\mathcal{U}_2}(\mathbf{p}, t^*) &= \left(\frac{\Delta_{\mathcal{P}_2}(\mathbf{p}, t^*)}{\sqrt{\lambda}} + \Delta_g \mathbf{p}'(\mathbf{p}) \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N, \nu})\|_{L_{2p}} + 2 \frac{\Delta_g^2}{\lambda} \mathbf{p}'(\mathbf{p}) \sup_{\nu \in [0, t^*]} \left\| \dot{\Lambda}_\mu^\parallel(\nu, Z_{N, \nu}) \right\|_{L_{2p}} \right) \\ &\quad \times \sup_{\nu \in [0, t^*]} \left\| \dot{F}_\sigma^\parallel(\nu, Z_{N, \nu}) \right\|_{L_{2p}}, \\ \Delta_{\mathcal{U}_3}(\mathbf{p}, t^*) &= \left(\frac{\Delta_{\mathcal{P}_3}(\mathbf{p}, t^*)}{\lambda} + \frac{\Delta_g^2}{\lambda} \mathbf{p}'(\mathbf{p})^2 \sup_{\nu \in [0, t^*]} \left\| \dot{F}_\sigma^\parallel(\nu, Z_{N, \nu}) \right\|_{L_{2p}} \right) \sup_{\nu \in [0, t^*]} \left\| \dot{F}_\sigma^\parallel(\nu, Z_{N, \nu}) \right\|_{L_{2p}},\end{aligned}$$

and where $\{\Delta_{\mathcal{P}_1}(\mathbf{p}, t^*), \Delta_{\mathcal{P}_2}(\mathbf{p}, t^*), \Delta_{\mathcal{P}_3}(\mathbf{p}, t^*)\}$, and $\Delta_{\mathcal{U}}(\mathbf{p}, t^*, \omega, \mathbf{T}_s)$ are defined in Proposition D.8 and Lemma D.2, respectively.

Proof. We begin by recalling the definition of $\psi(\tau(t), \nu, Z_N)$ from (110)

$$\begin{aligned}\psi(\tau(t), \nu, Z_N) &= \frac{\omega}{2\lambda - \omega} \left(e^{\omega(\tau(t)+\nu)} \mathcal{P}(\tau(t), \nu) - e^{(2\lambda\tau(t)+\omega\nu)} \partial \mathcal{V}(Z_{N, \tau(t)})^\top g(\tau(t)) \right) \\ &\quad + \frac{2\lambda}{2\lambda - \omega} e^{(\omega\tau(t)+2\lambda\nu)} \partial \mathcal{V}(Z_{N, \nu})^\top g(\nu) \in \mathbb{R}^{1 \times m},\end{aligned}$$

which, upon using the decomposition (D.35) in Proposition D.6, can be re-written as

$$\psi(\tau(t), \nu, Z_N) = \sum_{i=1}^3 \psi_i(\tau(t), \nu, Z_N) + \psi_{ad}(\tau(t), \nu, Z_N) + \tilde{\psi}(\tau(t), \nu, Z_N) \in \mathbb{R}^{1 \times m}, \quad (\text{D.100})$$

where

$$\begin{aligned}\psi_1(\tau(t), \nu, Z_N) &= \frac{\omega}{2\lambda - \omega} e^{\omega(\tau(t)+\nu)} \mathcal{P}_o(\tau(t), \nu), \\ \psi_2(\tau(t), \nu, Z_N) &= -\frac{\omega}{2\lambda - \omega} e^{(2\lambda\tau(t)+\omega\nu)} \partial \mathcal{V}(Z_{N, \tau(t)})^\top g(\tau(t)), \\ \psi_3(\tau(t), \nu, Z_N) &= \frac{2\lambda}{2\lambda - \omega} e^{(\omega\tau(t)+2\lambda\nu)} \partial \mathcal{V}(Z_{N, \nu})^\top g(\nu),\end{aligned}$$

and

$$\psi_{ad}(\tau(t), \nu, Z_N) = \frac{\omega}{2\lambda - \omega} e^{\omega(\tau(t)+\nu)} \mathcal{P}_{ad}(\tau(t), \nu), \quad \tilde{\psi}(\tau(t), \nu, Z_N) = \frac{\omega}{2\lambda - \omega} e^{\omega(\tau(t)+\nu)} \tilde{\mathcal{P}}(\tau(t), \nu).$$

Using the decomposition in (D.100), we re-write the terms $\mathcal{U}_\mu(\tau(t), \nu, Z_N; \omega)$ and $\mathcal{U}_\sigma(\tau(t), \nu, Z_N; \omega)$ defined in (109), as follows:

$$\begin{aligned}\mathcal{U}_\mu(\tau(t), \nu, Z_N; \omega) &= \left(\sum_{i=1}^3 \psi_i(\tau(t), \nu, Z_N) + \psi_{ad}(\tau(t), \nu, Z_N) + \tilde{\psi}(\tau(t), \nu, Z_N) \right) \dot{\Lambda}_\mu^\parallel(\nu, Z_{N, \nu}) \in \mathbb{R}, \\ \mathcal{U}_\sigma(\tau(t), \nu, Z_N; \omega) &= \left(\sum_{i=1}^3 \psi_i(\tau(t), \nu, Z_N) + \psi_{ad}(\tau(t), \nu, Z_N) + \tilde{\psi}(\tau(t), \nu, Z_N) \right) \dot{F}_\sigma^\parallel(\nu, Z_{N, \nu}) \in \mathbb{R}^{1 \times d}.\end{aligned}$$

Thus, using the definition of $\Xi_{\mathcal{U}}(\tau(t), Z_N; \omega)$ in (107), we may write

$$\begin{aligned}\Xi_{\mathcal{U}}(\tau(t), Z_N; \omega) &= \int_0^{\tau(t)} (\mathcal{U}_\mu(\tau(t), \nu, Z_N; \omega) d\nu + \mathcal{U}_\sigma(\tau(t), \nu, Z_N; \omega) dW_\nu) \\ &= \sum_{i=1}^3 \int_0^{\tau(t)} \psi_i(\tau(t), \nu, Z_N) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N, \nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N, \nu}) dW_\nu \right] \\ &\quad + \int_0^{\tau(t)} \psi_{ad}(\tau(t), \nu, Z_N) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N, \nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N, \nu}) dW_\nu \right] \\ &\quad + \int_0^{\tau(t)} \tilde{\psi}(\tau(t), \nu, Z_N) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N, \nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N, \nu}) dW_\nu \right].\end{aligned}$$

Hence, exploiting the linearity of the expectation operator, along with the Minkowski's inequality, we obtain

$$\begin{aligned}
& \mathbb{E} \left[e^{-(2\lambda + \omega)\tau(t)} \Xi_{\mathcal{U}}(\tau(t), Z_N; \omega) \right] \\
&= e^{-(2\lambda + \omega)t} \sum_{i=1}^3 \mathbb{E} \left[\int_0^t \psi_i(t, \nu, Z_N) \left[\dot{\Lambda}_{\mu}^{\parallel}(\nu, Z_{N,\nu}) d\nu + \dot{F}_{\sigma}^{\parallel}(\nu, Z_{N,\nu}) dW_{\nu} \right] \right. \\
&\quad \left. + e^{-(2\lambda + \omega)t} \mathbb{E} \left[\int_0^t \psi_{ad}(t, \nu, Z_N) \left[\dot{\Lambda}_{\mu}^{\parallel}(\nu, Z_{N,\nu}) d\nu + \dot{F}_{\sigma}^{\parallel}(\nu, Z_{N,\nu}) dW_{\nu} \right] \right] \right. \\
&\quad \left. + e^{-(2\lambda + \omega)t} \mathbb{E} \left[\int_0^t \tilde{\psi}(t, \nu, Z_N) \left[\dot{\Lambda}_{\mu}^{\parallel}(\nu, Z_{N,\nu}) d\nu + \dot{F}_{\sigma}^{\parallel}(\nu, Z_{N,\nu}) dW_{\nu} \right] \right] \right], \quad (\text{D.101})
\end{aligned}$$

for all $t \in [0, t^*]$, and

$$\begin{aligned}
& \left\| e^{-(2\lambda + \omega)\tau(t)} \Xi_{\mathcal{U}}(\tau(t), Z_N; \omega) \right\|_{L_p} \\
&\leq e^{-(2\lambda + \omega)t} \sum_{i=1}^3 \left\| \int_0^t \psi_i(t, \nu, Z_N) \left[\dot{\Lambda}_{\mu}^{\parallel}(\nu, Z_{N,\nu}) d\nu + \dot{F}_{\sigma}^{\parallel}(\nu, Z_{N,\nu}) dW_{\nu} \right] \right\|_{L_p} \\
&\quad + e^{-(2\lambda + \omega)t} \left\| \int_0^t \psi_{ad}(t, \nu, Z_N) \left[\dot{\Lambda}_{\mu}^{\parallel}(\nu, Z_{N,\nu}) d\nu + \dot{F}_{\sigma}^{\parallel}(\nu, Z_{N,\nu}) dW_{\nu} \right] \right\|_{L_p} \\
&\quad + e^{-(2\lambda + \omega)t} \left\| \int_0^t \tilde{\psi}(t, \nu, Z_N) \left[\dot{\Lambda}_{\mu}^{\parallel}(\nu, Z_{N,\nu}) d\nu + \dot{F}_{\sigma}^{\parallel}(\nu, Z_{N,\nu}) dW_{\nu} \right] \right\|_{L_p}, \quad (\text{D.102})
\end{aligned}$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 2}$. In the two expressions above, we have used the fact that since $\tau^* = t^*$, the definition of $\tau(t)$ in (34) leads to the conclusion that $\tau(t) = t$, for $t \in [0, t^*]$. Hence, the stopping time $\tau(t)$ is a deterministic function of t since t^* is a constant, and thus

$$\mathbb{E} \left[e^{-(2\lambda + \omega)\tau(t)} \right] = e^{-(2\lambda + \omega)\tau(t)} = e^{-(2\lambda + \omega)t}, \quad \forall t \in [0, t^*].$$

Now, using the definition of ψ_1 in (D.100), we get

$$\begin{aligned}
& \int_0^t \psi_1(t, \nu, Z_N) \left[\dot{\Lambda}_{\mu}^{\parallel}(\nu, Z_{N,\nu}) d\nu + \dot{F}_{\sigma}^{\parallel}(\nu, Z_{N,\nu}) dW_{\nu} \right] \\
&= \frac{\omega e^{\omega t}}{2\lambda - \omega} \int_0^t e^{\omega \nu} \mathcal{P}_{\circ}(t, \nu) \left[\dot{\Lambda}_{\mu}^{\parallel}(\nu, Z_{N,\nu}) d\nu + \dot{F}_{\sigma}^{\parallel}(\nu, Z_{N,\nu}) dW_{\nu} \right], \quad t \in [0, t^*]. \quad (\text{D.103})
\end{aligned}$$

Using the definition of \mathcal{P}_{\circ} in (D.36a), one sees that

$$\int_0^t e^{\omega \nu} \mathcal{P}_{\circ}(t, \nu) \left[\dot{\Lambda}_{\mu}^{\parallel}(\nu, Z_{N,\nu}) d\nu + \dot{F}_{\sigma}^{\parallel}(\nu, Z_{N,\nu}) dW_{\nu} \right] = \sum_{i=1}^4 N_i(t), \quad t \in [0, T],$$

where the processes $N_i(t)$, $i \in \{1, \dots, 4\}$, are defined in (D.50) in the statement of Proposition D.8. Hence, we may write (D.103) as

$$\int_0^t \psi_1(t, \nu, Z_N) \left[\dot{\Lambda}_{\mu}^{\parallel}(\nu, Z_{N,\nu}) d\nu + \dot{F}_{\sigma}^{\parallel}(\nu, Z_{N,\nu}) dW_{\nu} \right] = \frac{\omega e^{\omega t}}{2\lambda - \omega} \sum_{i=1}^4 N_i(t), \quad t \in [0, t^*].$$

It then follows from the linearity of the expectation operator and the Minkowski's inequality that

$$\mathbb{E} \left[\int_0^t \psi_1(t, \nu, Z_N) \left[\dot{\Lambda}_{\mu}^{\parallel}(\nu, Z_{N,\nu}) d\nu + \dot{F}_{\sigma}^{\parallel}(\nu, Z_{N,\nu}) dW_{\nu} \right] \right] \leq \frac{\omega e^{\omega t}}{|2\lambda - \omega|} \sum_{i=1}^4 |\mathbb{E}[N_i(t)]|, \quad \forall t \in [0, t^*],$$

and

$$\begin{aligned}
& \left\| \int_0^t \psi_1(t, \nu, Z_N) \left[\dot{\Lambda}_{\mu}^{\parallel}(\nu, Z_{N,\nu}) d\nu + \dot{F}_{\sigma}^{\parallel}(\nu, Z_{N,\nu}) dW_{\nu} \right] \right\|_{L_p} \\
&\leq \frac{\omega e^{\omega t}}{|2\lambda - \omega|} \sum_{i=1}^4 \|N_i(t)\|_{L_p}, \quad \forall (t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 2}.
\end{aligned}$$

Hence, using Proposition D.8 leads to

$$\begin{aligned} & \mathbb{E} \left[\int_0^t \psi_1(t, \nu, Z_N) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right] \right] \\ & \leq \frac{e^{(2\lambda+\omega)t}}{\sqrt{\lambda}|2\lambda-\omega|} \Delta_{\mathcal{P}_1}(1, t^*) \sup_{\nu \in [0, t^*]} \left\| \dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) \right\|_{L_2} \\ & \quad + \frac{e^{(2\lambda+\omega)t}}{|2\lambda-\omega|} \left(\frac{\sqrt{\omega}}{\sqrt{\lambda}} \Delta_{\mathcal{P}_2}(1, t^*) + \frac{\omega}{\lambda} \Delta_{\mathcal{P}_3}(1, t^*) \right) \sup_{\nu \in [0, t^*]} \left\| \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) \right\|_{L_2}, \quad (\text{D.104}) \end{aligned}$$

for all $t \in [0, t^*]$, and

$$\begin{aligned} & \left\| \int_0^t \psi_1(t, \nu, Z_N) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right] \right\|_{L_p} \\ & \leq \frac{e^{(2\lambda+\omega)t}}{\sqrt{\lambda}|2\lambda-\omega|} \Delta_{\mathcal{P}_1}(\mathbf{p}, t^*) \sup_{\nu \in [0, t^*]} \left\| \dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) \right\|_{L_{2p}} \\ & \quad + \frac{e^{(2\lambda+\omega)t}}{|2\lambda-\omega|} \left(\frac{\sqrt{\omega}}{\sqrt{\lambda}} \Delta_{\mathcal{P}_2}(\mathbf{p}, t^*) + \frac{\omega}{\lambda} \Delta_{\mathcal{P}_3}(\mathbf{p}, t^*) \right) \sup_{\nu \in [0, t^*]} \left\| \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) \right\|_{L_{2p}}, \quad (\text{D.105}) \end{aligned}$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 2}$.

Next, using the definitions of ψ_2 and ψ_3 in (D.100) we see that

$$\begin{aligned} & \int_0^t \psi_2(t, \nu, Z_N) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right] \\ & = -\frac{\omega e^{2\lambda t}}{2\lambda-\omega} \partial \mathcal{V}(Z_{N,\tau(t)})^\top g(\tau(t)) \int_0^t e^{\omega \nu} \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right], \end{aligned}$$

and

$$\begin{aligned} & \int_0^t \psi_3(t, \nu, Z_N) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right] \\ & = \frac{2\lambda e^{\omega t}}{2\lambda-\omega} \int_0^t e^{2\lambda \nu} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right], \end{aligned}$$

for $t \in [0, t^*]$. Following the derivations of (C.56), (C.62), and (C.63) in the proof of Lemma C.2, we obtain

$$\begin{aligned} & \left\| \int_0^t \psi_2(t, \nu, Z_N) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right] \right\|_{L_p} \\ & \leq \frac{e^{(2\lambda+\omega)t}}{|2\lambda-\omega|} \Delta_g \sup_{\nu \in [0, t^*]} \left\| \partial \mathcal{V}(Z_{N,\nu}) \right\|_{L_{2p}} \sup_{\nu \in [0, t^*]} \left\| \dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) \right\|_{L_{2p}} \\ & \quad + \frac{\sqrt{\omega} e^{(2\lambda+\omega)t}}{|2\lambda-\omega|} \Delta_g \mathbf{p}'(\mathbf{p}) \sup_{\nu \in [0, t^*]} \left\| \partial \mathcal{V}(Z_{N,\nu}) \right\|_{L_{2p}} \sup_{\nu \in [0, t^*]} \left\| \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) \right\|_{L_{2p}}, \quad (\text{D.106}) \end{aligned}$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 1}$, along with

$$\begin{aligned} & \mathbb{E} \left[\int_0^t \psi_3(t, \nu, Z_N) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right] \right] \\ & \leq \frac{e^{(2\lambda+\omega)t}}{|2\lambda-\omega|} \Delta_g \sup_{\nu \in [0, t^*]} \left\| \partial \mathcal{V}(Z_{N,\nu}) \right\|_{L_2} \sup_{\nu \in [0, t^*]} \left\| \dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) \right\|_{L_2}, \quad \forall t \in [0, t^*], \quad (\text{D.107}) \end{aligned}$$

and

$$\begin{aligned} & \left\| \int_0^t \psi_3(t, \nu, Z_N) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right] \right\|_{L_p} \\ & \leq \frac{e^{(2\lambda+\omega)t}}{|2\lambda-\omega|} \Delta_g \sup_{\nu \in [0, t^*]} \left\| \partial \mathcal{V}(Z_{N,\nu}) \right\|_{L_{2p}} \sup_{\nu \in [0, t^*]} \left\| \dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) \right\|_{L_{2p}} \end{aligned}$$

$$+ \frac{\sqrt{\lambda}e^{(2\lambda+\omega)t}}{|2\lambda-\omega|} \Delta_g \mathbf{p}(\mathbf{p}) \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2\mathbf{p}}} \sup_{\nu \in [0, t^*]} \left\| \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) \right\|_{L_{2\mathbf{p}}}, \quad (\text{D.108})$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 2}$.

For the derivation of the penultimate component of the desired bound, we use the definition of ψ_{ad} in (D.100) to obtain

$$\begin{aligned} & \int_0^t \psi_{ad}(t, \nu, Z_N) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right] \\ &= \frac{\omega e^{\omega t}}{2\lambda - \omega} \int_0^t e^{\omega \nu} \mathcal{P}_{ad}(t, \nu) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right] \\ &= \frac{\omega e^{\omega t}}{2\lambda - \omega} \int_0^t e^{\omega \nu} \left(\int_\nu e^{(2\lambda-\omega)\beta} \mathcal{P}_\mathcal{U}(\beta)^\top d\beta \right) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right], \quad t \in [0, t^*], \end{aligned}$$

where, in the last expression we have used the definition of \mathcal{P}_{ad} in (D.36b), Proposition D.6. Observe that the expression above can be alternatively expressed as

$$\int_0^t \psi_{ad}(t, \nu, Z_N) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right] = \frac{\omega e^{\omega t}}{2\lambda - \omega} N_\mathcal{U}(t), \quad t \in [0, t^*],$$

where the process $N_\mathcal{U}$ is defined in (D.53), Proposition D.9. Hence,

$$\left\| \int_0^t \psi_{ad}(t, \nu, Z_N) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right] \right\|_{L_{\mathbf{p}}} \leq \frac{\omega e^{\omega t}}{|2\lambda - \omega|} \|N_\mathcal{U}(t)\|_{L_{\mathbf{p}}},$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 1}$. Thus, we invoke Proposition D.9 and obtain

$$\begin{aligned} & \left\| \int_0^t \psi_{ad}(t, \nu, Z_N) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right] \right\|_{L_{\mathbf{p}}} \\ & \leq \frac{e^{(2\lambda+\omega)t}}{\lambda|2\lambda-\omega|} \Delta_g^2 \left(\sup_{\nu \in [0, t^*]} \left\| \dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) \right\|_{L_{2\mathbf{p}}} + \sqrt{\omega} \mathbf{p}'(\mathbf{p}) \sup_{\nu \in [0, t^*]} \left\| \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) \right\|_{L_{2\mathbf{p}}} \right)^2, \quad (\text{D.109}) \end{aligned}$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 1}$.

Finally, using the definition of $\tilde{\psi}$ in (D.100) we see that

$$\begin{aligned} & \int_0^t \tilde{\psi}(\tau(t), \nu, Z_N) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right] \\ &= \frac{\omega e^{\omega t}}{2\lambda - \omega} \int_0^t e^{\omega \nu} \tilde{\mathcal{P}}(t, \nu) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right] \\ &= \frac{\omega e^{\omega t}}{2\lambda - \omega} \int_0^t e^{\omega \nu} \left(\int_\nu e^{(2\lambda-\omega)\beta} \tilde{\mathcal{P}}_\mathcal{U}(\beta)^\top d\beta \right) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right], \end{aligned}$$

for $t \in [0, t^*]$, where, in the last expression we have used the definition of $\tilde{\mathcal{P}}$ in (D.36b), Proposition D.6. Since the process on the right hand side is identical to $N_\mathcal{U}(t)$ defined in the statement of Lemma D.2, we can write

$$\int_0^t \tilde{\psi}(\tau(t), \nu, Z_N) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right] = \frac{\omega e^{\omega t}}{2\lambda - \omega} N_\mathcal{U}(t), \quad t \in [0, t^*],$$

which further implies that

$$\left\| \int_0^t \tilde{\psi}(\tau(t), \nu, Z_N) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right] \right\|_{L_{\mathbf{p}}} \leq \frac{\omega e^{\omega t}}{|2\lambda - \omega|} \|N_\mathcal{U}(t)\|_{L_{\mathbf{p}}},$$

for all $(t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 1}$. It then follows from Lemma D.2 that

$$\left\| \int_0^t \tilde{\psi}(\tau(t), \nu, Z_N) \left[\dot{\Lambda}_\mu^\parallel(\nu, Z_{N,\nu}) d\nu + \dot{F}_\sigma^\parallel(\nu, Z_{N,\nu}) dW_\nu \right] \right\|_{L_{\mathbf{p}}}$$

$$\leq \frac{e^{(2\lambda+\omega)t}}{\lambda|2\lambda-\omega|} \omega \Delta_{\mathcal{U}}(\mathbf{p}, t^*, \omega, \mathbf{T}_s), \quad \forall (t, \mathbf{p}) \in [0, t^*] \times \mathbb{N}_{\geq 1}. \quad (\text{D.110})$$

Substituting (D.104), (D.106), (D.107), (D.109), and (D.110) into (D.101), followed by grouping terms that are $\propto \{\omega, \sqrt{\omega}, 1\} / |2\lambda - \omega|$ leads to (D.98). Similarly, substituting (D.105), (D.106), (D.108), (D.109), and (D.110) into (D.102) leads to (D.99), thus concluding the proof. \square

The next set of results below will allow us to bound the moments of $\tilde{\Xi}_{\mathcal{U}}$ in Lemma 4.5. We begin with a result which is similar to Proposition D.12.

Proposition D.13 *Let*

$$\mathfrak{J}_j(t) = \int_{\mathbf{T}_s}^t e^{(2\lambda-\omega)\nu} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \left(\int_0^\nu e^{\omega\beta} \mathcal{J}_j(\beta, X_N) dW_\beta \right) d\nu, \quad (j, t) \in \{1, \dots, 4\} \times [\mathbf{T}_s, t^*], \quad (\text{D.111})$$

where the piecewise functions \mathcal{J}_j are defined in (D.73), Proposition D.12, for $\beta \in [0, \nu]$, $\nu \in [\mathbf{T}_s, t]$, $t \geq \mathbf{T}_s$. If the stopping time τ^* , defined in (104), Lemma 4.5, satisfies $\tau^* = t^*$ and $t^* \geq \mathbf{T}_s$, then

$$\sum_{j=1}^4 \|\mathfrak{J}_j(t)\|_{L_p} \leq \frac{e^{2\lambda t}}{2\lambda\sqrt{\omega}} \Delta_g \gamma''(\mathbf{p}, \omega, \mathbf{T}_s) \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{4p}}, \quad (\text{D.112})$$

for all $(t, \mathbf{p}) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$, where $\mathbf{p}'(\mathbf{p})$ and $\mathbf{p}''(\mathbf{p})$ are defined in (A.1a), and $\gamma_2(\omega, \mathbf{T}_s)$, $\gamma_2'(\omega, \mathbf{T}_s)$, and $\gamma''(\mathbf{p}, \omega, \mathbf{T}_s)$ are defined in (A.13).

Proof. We begin by using the bound on $g(t)$ in Assumption 1 to obtain

$$\|\mathfrak{J}_j(t)\|_{L_p} \leq \Delta_g \left\| \int_{\mathbf{T}_s}^t e^{(2\lambda-\omega)\nu} \|\partial \mathcal{V}(Z_{N,\nu})\| \left\| \int_0^\nu e^{\omega\beta} \mathcal{J}_j(\beta, X_N) dW_\beta \right\| d\nu \right\|_{L_p}, \quad j \in \{1, \dots, 4\},$$

for all $(t, \mathbf{p}) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$. We now bound each of the processes above.

Bound for $\mathfrak{J}_1(t)$: Invoking Lemma B.3 by setting

$$\xi = \mathbf{T}_s, \quad \theta_1 = 2\lambda, \quad \theta_2 = \omega, \quad \{\zeta_1, \zeta_2\} = \{0, t\}, \quad S(\nu) = \partial \mathcal{V}(Z_{N,\nu}), \quad L(\beta) = \mathcal{J}_1(\beta, X_N),$$

we obtain

$$\begin{aligned} \|\mathfrak{J}_1(t)\|_{L_p} &\leq \Delta_g \left\| \int_{\mathbf{T}_s}^t e^{(2\lambda-\omega)\nu} \|\partial \mathcal{V}(Z_{N,\nu})\| \left\| \int_0^\nu e^{\omega\beta} \mathcal{J}_1(\beta, X_N) dW_\beta \right\| d\nu \right\|_{L_p} \\ &\leq \frac{e^{2\lambda t}}{2\lambda\sqrt{\omega}} \mathbf{p}'(\mathbf{p}) \Delta_g \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \sup_{\nu \in [0, t^*]} \|\mathcal{J}_1(\nu, X_N)\|_{L_{2p}}, \end{aligned} \quad (\text{D.113})$$

for all $(t, \mathbf{p}) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$. Recall from (D.76) that

$$\|\mathcal{J}_1(\nu, X_N)\|_{L_{2p}} \leq \gamma_2'(\omega, \mathbf{T}_s) \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{2p}}, \quad \forall (t, \mathbf{p}) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}.$$

Substituting the above into (D.113) leads to

$$\|\mathfrak{J}_1(t)\|_{L_p} \leq \frac{e^{2\lambda t}}{2\lambda\sqrt{\omega}} \mathbf{p}'(\mathbf{p}) \Delta_g \gamma_2'(\omega, \mathbf{T}_s) \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{4p}}, \quad (\text{D.114})$$

for all $(t, \mathbf{p}) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$. Note that we have also used the bound $\|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{2p}} \leq \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{4p}}$, which is a consequence of the Jensen's inequality.

Bound for $\mathfrak{J}_2(t)$: The definition of \mathcal{J}_2 allows us to write

$$\begin{aligned} &\|\mathfrak{J}_2(t)\|_{L_p} \\ &\leq \Delta_g \left\| \int_{\mathbf{T}_s}^t e^{(2\lambda-\omega)\nu} \|\partial \mathcal{V}(Z_{N,\nu})\| \left\| \int_0^\nu e^{\omega\beta} \mathcal{J}_2(\beta, X_N) dW_\beta \right\| d\nu \right\|_{L_p} \end{aligned}$$

$$= \Delta_g \left\| \int_{\mathbf{T}_s}^t e^{(2\lambda - \omega)\nu} \|\partial \mathcal{V}(Z_{N,\nu})\| \left\| F_\sigma^\parallel((i^*(\nu) - 1)\mathbf{T}_s, X_{N,(i^*(\nu)-1)\mathbf{T}_s}) \int_{(i^*(\nu)-1)\mathbf{T}_s}^{i^*(\nu)\mathbf{T}_s} e^{\omega\beta} dW_\beta \right\| d\nu \right\|_{L_p},$$

for all $(t, p) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$, which can be cast as the process $N_a(t)$ in Corollary B.2 by setting

$$\theta_1 = 2\lambda, \quad \theta_2 = \omega, \quad Q_t = W_t, \\ S(\nu) = \partial \mathcal{V}(Z_{N,\nu}), \quad L(\nu) = F_\sigma^\parallel((i^*(\nu) - 1)\mathbf{T}_s, X_{N,(i^*(\nu)-1)\mathbf{T}_s}), \quad \varsigma_1(\nu) = \varsigma_2(\beta) = 1.$$

Hence, by (B.17) in Corollary B.2 we have

$$\|\mathfrak{J}_2(t)\|_{L_p} \leq \frac{e^{2\lambda t}}{2\lambda\sqrt{\omega}} \mathbf{p}''(\mathbf{p}) \Delta_g (1 - e^{-2\omega\mathbf{T}_s})^{\frac{1}{2}} \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{4p}}, \quad (\text{D.115})$$

for all $(t, p) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$, where we have used the fact that

$$\begin{aligned} & \sup_{\nu \in [\mathbf{T}_s, t]} \|F_\sigma^\parallel((i^*(\nu) - 1)\mathbf{T}_s, X_{N,(i^*(\nu)-1)\mathbf{T}_s})\|_{L_{4p}} \\ & \leq \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel((i^*(\nu) - 1)\mathbf{T}_s, X_{N,(i^*(\nu)-1)\mathbf{T}_s})\|_{L_{4p}} \leq \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{4p}}, \quad \forall t \in [\mathbf{T}_s, t^*]. \end{aligned}$$

Bound for $\mathfrak{J}_3(t)$: Similar to the above, we use the definition of \mathcal{J}_3 to obtain the bound:

$$\begin{aligned} & \|\mathfrak{J}_3(t)\|_{L_p} \\ & \leq \Delta_g \left\| \int_{\mathbf{T}_s}^t e^{(2\lambda - \omega)\nu} \|\partial \mathcal{V}(Z_{N,\nu})\| \left\| \int_0^\nu e^{\omega\beta} \mathcal{J}_3(\beta, X_N) dW_\beta \right\| d\nu \right\|_{L_p} \\ & = \Delta_g \left\| \int_{\mathbf{T}_s}^t e^{(2\lambda - \omega)\nu} \|\partial \mathcal{V}(Z_{N,\nu})\| \left\| F_\sigma^\parallel(i^*(\nu)\mathbf{T}_s, X_{N,i^*(\nu)\mathbf{T}_s}) \tilde{\gamma}^*(\nu, \mathbf{T}_s) \int_{(i^*(\nu)-1)\mathbf{T}_s}^{i^*(\nu)\mathbf{T}_s} e^{\omega\beta} e^{(\lambda_s - \omega)\beta} dW_\beta \right\| d\nu \right\|_{L_p}, \end{aligned}$$

for all $(t, p) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$, which can be cast as the process $N_a(t)$ in Corollary B.2 by setting

$$\theta_1 = 2\lambda, \quad \theta_2 = \omega, \quad Q_t = W_t, \quad \varsigma_1(\nu) = \tilde{\gamma}^*(\nu, \mathbf{T}_s), \quad \varsigma_2(\beta) = e^{(\lambda_s - \omega)\beta}, \\ S(\nu) = \partial \mathcal{V}(Z_{N,\nu}), \quad L(\nu) = F_\sigma^\parallel(i^*(\nu)\mathbf{T}_s, X_{N,i^*(\nu)\mathbf{T}_s}).$$

Hence, by (B.17) in Corollary B.2 we have

$$\begin{aligned} \|\mathfrak{J}_3(t)\|_{L_p} & \leq \frac{e^{2\lambda t}}{2\lambda\sqrt{\omega}} \mathbf{p}''(\mathbf{p}) \Delta_g (1 - e^{-2\omega\mathbf{T}_s})^{\frac{1}{2}} \sup_{\nu \in [\mathbf{T}_s, t]} \left(|\tilde{\gamma}^*(\nu, \mathbf{T}_s)| \sup_{\beta \in [(i^*(\nu)-1)\mathbf{T}_s, i^*(\nu)\mathbf{T}_s]} e^{(\lambda_s - \omega)\beta} \right) \\ & \quad \times \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{4p}}, \quad (\text{D.116}) \end{aligned}$$

for all $(t, p) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$, where we have used the fact that

$$\begin{aligned} & \sup_{\nu \in [\mathbf{T}_s, t]} \|F_\sigma^\parallel(i^*(\nu)\mathbf{T}_s, X_{N,i^*(\nu)\mathbf{T}_s})\|_{L_{4p}} \\ & \leq \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(i^*(\nu)\mathbf{T}_s, X_{N,i^*(\nu)\mathbf{T}_s})\|_{L_{4p}} \leq \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{4p}}, \quad \forall t \in [\mathbf{T}_s, t^*]. \quad (\text{D.117}) \end{aligned}$$

Once again we recall the following from (D.81):

$$|\tilde{\gamma}^*(\nu, \mathbf{T}_s)| \sup_{\beta \in [(i^*(\nu)-1)\mathbf{T}_s, i^*(\nu)\mathbf{T}_s]} e^{(\lambda_s - \omega)\beta} \leq \max \left\{ e^{(\omega - \lambda_s)\mathbf{T}_s}, 1 \right\} \frac{\lambda_s}{\omega} \frac{e^{\omega\mathbf{T}_s} - 1}{e^{\lambda_s\mathbf{T}_s} - 1} \doteq \gamma_2(\omega, \mathbf{T}_s),$$

for all $\nu \in [\mathbf{T}_s, t^*]$. Substituting the above into (D.116) then leads to

$$\begin{aligned} \|\mathfrak{J}_3(t)\|_{L_p} & \leq \frac{e^{2\lambda t}}{2\lambda\sqrt{\omega}} \mathbf{p}''(\mathbf{p}) \Delta_g \gamma_2(\omega, \mathbf{T}_s) (1 - e^{-2\omega\mathbf{T}_s})^{\frac{1}{2}} \\ & \quad \times \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{4p}}, \quad (\text{D.118}) \end{aligned}$$

for all $(t, p) \in [T_s, t^*] \times \mathbb{N}_{\geq 1}$.

Bound for $\mathfrak{J}_4(t)$: Finally, we have

$$\begin{aligned} \|\mathfrak{J}_4(t)\|_{L_p} &\leq \Delta_g \left\| \int_{T_s}^t e^{(2\lambda - \omega)\nu} \|\partial \mathcal{V}(Z_{N,\nu})\| \left\| \int_0^\nu e^{\omega\beta} \mathcal{J}_4(\beta, X_N) dW_\beta \right\| d\nu \right\|_{L_p} \\ &= \Delta_g \left\| \int_{T_s}^t e^{(2\lambda - \omega)\nu} \|\partial \mathcal{V}(Z_{N,\nu})\| \left\| F_\sigma^\parallel(i^*(\nu) T_s, X_{N,i^*(\nu) T_s}) \int_{i^*(\nu) T_s}^\nu e^{\omega\beta} dW_\beta \right\| d\nu \right\|_{L_p}, \end{aligned}$$

for all $(t, p) \in [T_s, t^*] \times \mathbb{N}_{\geq 1}$. The term inside the norm can be cast as the process $N_b(t)$ in Corollary B.2 by setting

$$\begin{aligned} \theta_1 &= 2\lambda, \quad \theta_2 = \omega, \quad Q_t = W_t, \\ S(\nu) &= \partial \mathcal{V}(Z_{N,\nu}), \quad L(\nu) = F_\sigma^\parallel(i^*(\nu) T_s, X_{N,i^*(\nu) T_s}). \end{aligned}$$

Hence, by (B.18) in Corollary B.2

$$\|\mathfrak{J}_4(t)\|_{L_p} \leq \frac{e^{2\lambda t}}{2\lambda\sqrt{\omega}} p''(p) \Delta_g (1 - e^{-2\omega T_s})^{\frac{1}{2}} \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{4p}}, \quad (\text{D.119})$$

for all $(t, p) \in [T_s, t^*] \times \mathbb{N}_{\geq 1}$, where we have also used (D.117). The proof is then concluded by adding together the bounds in (D.114), (D.115), (D.118), and (D.119). \square

We now derive a bound on the moments of $\tilde{\Xi}_{\mathcal{U}}$.

Lemma D.4 *If the stopping time τ^* , defined in (104), Lemma 4.5, satisfies $\tau^* = t^*$, then the term $\tilde{\Xi}_{\mathcal{U}}(\tau(t), Z_N; \omega)$ defined in (107c), Lemma 4.5, satisfies the following bound for all $(t, p) \in [0, t^*] \times \mathbb{N}_{\geq 1}$:*

$$\left\| e^{-2\lambda\tau(t)} \tilde{\Xi}_{\mathcal{U}}(\tau(t), Z_N; \omega) \right\|_{L_p} \leq \tilde{\Delta}_{\mathcal{U}}(p, t^*, \omega, T_s) \doteq \tilde{\Delta}^-(p, t^*, \omega, T_s) + 1_{\geq T_s}(t^*) \sum_{k=1}^3 \tilde{\Delta}_k^+(p, t^*, \omega, T_s), \quad (\text{D.120})$$

where

$$\begin{aligned} &\tilde{\Delta}^-(p, t^*, \omega, T_s) \\ &= \frac{1 - e^{-2\lambda T_s}}{2\lambda} \Delta_g \sup_{\nu \in [0, t^* \wedge T_s]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \\ &\quad \times \left(\sup_{\nu \in [0, t^* \wedge T_s]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu})\|_{L_{2p}} + \sqrt{\omega} p'(p) (1 - e^{-2\omega T_s})^{\frac{1}{2}} \sup_{\nu \in [0, t^* \wedge T_s]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{2p}} \right), \end{aligned}$$

and

$$\begin{aligned} \tilde{\Delta}_1^+(p, t^*, \omega, T_s) &= \frac{1}{2\lambda} \Delta_g \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \\ &\quad \times \left(\Delta_{\tilde{\mathcal{H}}_\mu}(T_s, p, t^*) + \Delta_{\mathcal{G}_\mu}(T_s, t^*) + \sqrt{\omega} p'(p) \left[\Delta_{\tilde{\mathcal{H}}_\sigma}(T_s, p, t^*) + \Delta_{\tilde{\mathcal{G}}_\sigma}(T_s, p, t^*) \right] \right), \\ \tilde{\Delta}_2^+(p, t^*, \omega, T_s) &= \sqrt{\omega} \frac{1}{2\lambda} \Delta_g \gamma''(p, \omega, T_s) \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \sup_{\nu \in [0, t^*]} \|F_\sigma^\parallel(\nu, X_{N,\nu})\|_{L_{4p}}, \\ \tilde{\Delta}_3^+(p, t^*, \omega, T_s) &= \frac{1}{|2\lambda - \omega|} (e^{\omega T_s} - 1) \Delta_g \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \sup_{\nu \in [0, t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu})\|_{L_{2p}}, \end{aligned}$$

and where the terms $\Delta_{\tilde{\mathcal{H}}_\mu}$, $\Delta_{\tilde{\mathcal{H}}_\sigma}$, $\Delta_{\tilde{\mathcal{G}}_\sigma}$, and $\Delta_{\mathcal{G}_\mu}$ are defined in Proposition D.11.

Proof. We begin by using the respective definitions of $\tilde{\Xi}_{\mathcal{U}}(\tau(t), Z_N; \omega)$ and $\tilde{\mathcal{U}}$ in (107c) and (108c) to write

$$\tilde{\Xi}_{\mathcal{U}}(\tau(t), Z_N) = \int_0^{\tau(t)} e^{2\lambda\nu} \tilde{\mathcal{U}}(\nu, Z_{N,\nu}) d\nu = \int_0^{\tau(t)} e^{2\lambda\nu} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) (\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r)(X_N)(\nu) d\nu$$

$$= \omega \int_0^{\tau(t)} e^{(2\lambda - \omega)\nu} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \tilde{\mathcal{F}}(X_N, \nu) d\nu,$$

where the last expression is due to the conclusion in Proposition D.1 that $(\mathcal{F}_{\mathcal{L}_1} - \mathcal{F}_r)(X_N)(\nu) = \omega e^{-\omega\nu} \tilde{\mathcal{F}}(X_N, \nu)$. Note that $\tau(t) = t$, for all $t \in [0, t^*]$, due to the hypothesis that $\tau^* = t^*$. We can thus write the last expression as

$$\tilde{\Xi}_{\mathcal{U}}(\tau(t), Z_N) = \omega \int_0^t e^{(2\lambda - \omega)\nu} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \tilde{\mathcal{F}}(X_N, \nu) d\nu, \quad t \in [0, t^*]. \quad (\text{D.121})$$

To obtain the expression for $\tilde{\mathcal{F}}(X_N, \nu)$, we invoke Proposition D.3 with $z = X_N$ to obtain:

$$\tilde{\mathcal{F}}(X_N, \nu) = \int_0^\nu e^{\omega\beta} \Sigma_\beta^\parallel(X_N), \quad \nu < T_s,$$

and for $\nu \geq T_s$,

$$\begin{aligned} \tilde{\mathcal{F}}(X_N, \nu) &= \int_0^{T_s} e^{\omega\beta} \Lambda_\mu^\parallel(\beta, X_{N,\beta}) d\beta \\ &\quad + \int_0^\nu e^{\omega\beta} \left[\left(\tilde{\mathcal{H}}_\mu^\parallel(\beta, X_N) + e^{(\lambda_s - \omega)\beta} \tilde{\mathcal{H}}_\mu(\beta, X_N) \right) d\beta + \left(\tilde{\mathcal{H}}_\sigma^\parallel(\beta, X_N) + e^{(\lambda_s - \omega)\beta} \tilde{\mathcal{H}}_\sigma(\beta, X_N) \right) dW_\beta \right] \\ &\quad + \int_0^\nu e^{\omega\beta} \left(\tilde{\mathcal{G}}_{\sigma_1}(\beta, X_N) + e^{(\lambda_s - \omega)\beta} \tilde{\mathcal{G}}_{\sigma_2}(\beta, X_N) \right) dW_\beta \\ &\quad + \int_0^\nu e^{\omega\beta} [\mathcal{G}_\mu(\beta, X_N) d\beta + (\mathcal{G}_{\sigma_1}(\beta, X_N) + \mathcal{G}_{\sigma_2}(\beta, X_N)) dW_\beta]. \end{aligned}$$

Substituting the above expression for $\tilde{\mathcal{F}}(X_N, \nu)$ into (D.121):

$$\tilde{\Xi}_{\mathcal{U}}(\tau(t), Z_N) = \begin{cases} \omega \int_0^t e^{(2\lambda - \omega)\nu} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \mathcal{N}_\star(\nu) d\nu, & \text{if } t < T_s, \\ \omega \int_0^{T_s} e^{(2\lambda - \omega)\nu} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \mathcal{N}_\star(\nu) d\nu \\ + \omega \int_{T_s}^t e^{(2\lambda - \omega)\nu} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \\ \quad \times (\mathcal{N}_1(\nu) + \mathcal{N}_2(\nu) + \mathcal{N}_3(\nu)) d\nu, & \text{if otherwise } t \geq T_s, \end{cases} \quad (\text{D.122})$$

where $\mathcal{N}_\star, \mathcal{N}_1, \mathcal{N}_2$, and \mathcal{N}_3 are defined in (D.87) as follows:

$$\begin{aligned} \mathcal{N}_\star(\nu) &= \int_0^\nu e^{\omega\beta} \Sigma_\beta^\parallel(X_N) = \int_0^\nu e^{\omega\beta} [\Lambda_\mu^\parallel(\beta, X_{N,\beta}) d\beta + F_\sigma^\parallel(\beta, X_{N,\beta}) dW_\beta], \\ \mathcal{N}_1(\nu) &= \int_0^\nu e^{\omega\beta} \left[\left(\tilde{\mathcal{H}}_\mu^\parallel(\beta, X_N) + e^{(\lambda_s - \omega)\beta} \tilde{\mathcal{H}}_\mu(\beta, X_N) + \mathcal{G}_\mu(\beta, X_N) \right) d\beta \right. \\ &\quad \left. + \left(\tilde{\mathcal{H}}_\sigma^\parallel(\beta, X_N) + e^{(\lambda_s - \omega)\beta} \tilde{\mathcal{H}}_\sigma(\beta, X_N) + \tilde{\mathcal{G}}_{\sigma_1}(\beta, X_N) + e^{(\lambda_s - \omega)\beta} \tilde{\mathcal{G}}_{\sigma_2}(\beta, X_N) \right) dW_\beta \right], \\ \mathcal{N}_2(\nu) &= \int_0^\nu e^{\omega\beta} (\mathcal{G}_{\sigma_1}(\beta, X_N) + \mathcal{G}_{\sigma_2}(\beta, X_N)) dW_\beta, \\ \mathcal{N}_3 &= \int_0^{T_s} e^{\omega\beta} \Lambda_\mu^\parallel(\beta, X_{N,\beta}) d\beta. \end{aligned}$$

We can further write (D.122) as

$$\tilde{\Xi}_{\mathcal{U}}(\tau(t), Z_N) = \begin{cases} \omega \tilde{\mathfrak{N}}_\star(t), & \text{if } t < T_s, \\ \omega \left(\tilde{\mathfrak{N}}_\star(T_s) + \sum_{k=1}^3 \tilde{\mathfrak{N}}_k(t) \right), & \text{if otherwise } t \geq T_s, \end{cases} \quad (\text{D.123})$$

where we have defined

$$\tilde{\mathfrak{N}}_\star(t) = \int_0^t e^{(2\lambda - \omega)\nu} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \mathcal{N}_\star(\nu) d\nu,$$

$$\tilde{\mathfrak{N}}_k(t) = \int_{\mathbf{T}_s}^t e^{(2\lambda - \omega)\nu} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \mathcal{N}_k(\nu) d\nu, \quad k \in \{1, 2, 3\}.$$

Using the Minkowski's inequality along with the fact that t^* and \mathbf{T}_s are constants leads to the following bounds for all $\mathbf{p} \in \mathbb{N}_{\geq 1}$:

$$\left\| e^{-2\lambda\tau(t)} \tilde{\Xi}_{\mathcal{U}}(t, Z_N) \right\|_{L_{\mathbf{p}}} \leq \begin{cases} \omega e^{-2\lambda t} \left\| \tilde{\mathfrak{N}}_{\star}(t) \right\|_{L_{\mathbf{p}}}, & \forall t \in [0, t^*] \cap [0, \mathbf{T}_s], \\ \omega e^{-2\lambda t} \left(\left\| \tilde{\mathfrak{N}}_{\star}(\mathbf{T}_s) \right\|_{L_{\mathbf{p}}} + \sum_{k=1}^3 \left\| \tilde{\mathfrak{N}}_k(t) \right\|_{L_{\mathbf{p}}} \right), & \forall t \in [0, t^*] \cap [\mathbf{T}_s, T]. \end{cases} \quad (\text{D.124})$$

We now bound the processes $\tilde{\mathfrak{N}}_{\star}(t)$ and $\tilde{\mathfrak{N}}_k(t)$, $k \in \{1, 2, 3\}$.

Bound for $\tilde{\mathfrak{N}}_{\star}(t)$: Using the definitions in (D.122)-(D.123), we have

$$\begin{aligned} \tilde{\mathfrak{N}}_{\star}(t) &= \int_0^t e^{(2\lambda - \omega)\nu} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \mathcal{N}_{\star}(\nu) d\nu \\ &= \int_0^t e^{(2\lambda - \omega)\nu} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \left(\int_0^\nu e^{\omega\beta} [\Lambda_\mu^\parallel(\beta, X_{N,\beta}) d\beta + F_\sigma^\parallel(\beta, X_{N,\beta}) dW_\beta] \right) d\nu, \end{aligned}$$

which, upon using Minkowski's inequality, leads to

$$\begin{aligned} \left\| \tilde{\mathfrak{N}}_{\star}(t) \right\|_{L_{\mathbf{p}}} &\leq \Delta_g \left\| \int_0^t e^{(2\lambda - \omega)\nu} \left\| \partial \mathcal{V}(Z_{N,\nu}) \right\| \left(\int_0^\nu e^{\omega\beta} \left\| \Lambda_\mu^\parallel(\beta, X_{N,\beta}) \right\| d\beta \right) d\nu \right\|_{L_{\mathbf{p}}} \\ &\quad + \Delta_g \left\| \int_0^t e^{(2\lambda - \omega)\nu} \left\| \partial \mathcal{V}(Z_{N,\nu}) \right\| \left\| \int_0^\nu e^{\omega\beta} F_\sigma^\parallel(\beta, X_{N,\beta}) dW_\beta \right\| d\nu \right\|_{L_{\mathbf{p}}}, \end{aligned}$$

for all $(t, \mathbf{p}) \in [0, t^* \wedge \mathbf{T}_s] \times \mathbb{N}_{\geq 1}$, where we have used the uniform bound on $g(t)$ in Assumption 1. We now invoke Proposition B.2 to bound the first term by setting

$$\xi = 0, \quad \theta_1 = 2\lambda, \quad \theta_2 = \omega, \quad S_1(\nu) = \partial \mathcal{V}(Z_{N,\nu}), \quad S_2(\beta) = \Lambda_\mu^\parallel(\beta, X_{N,\beta}),$$

and we invoke Lemma B.3 to bound the second term by setting

$$\xi = 0, \quad \theta_1 = 2\lambda, \quad \theta_2 = \omega, \quad \{\zeta_1, \zeta_2\} = \{0, t\}, \quad S(\nu) = \partial \mathcal{V}(Z_{N,\nu}), \quad L(\beta) = F_\sigma^\parallel(\beta, X_{N,\beta}),$$

and obtain

$$\begin{aligned} \left\| \tilde{\mathfrak{N}}_{\star}(t) \right\|_{L_{\mathbf{p}}} &\leq \frac{e^{2\lambda t}}{2\lambda\sqrt{\omega}} (1 - e^{-2\lambda t}) \Delta_g \sup_{\nu \in [0, t]} \left\| \partial \mathcal{V}(Z_{N,\nu}) \right\|_{L_{2\mathbf{p}}} \\ &\quad \times \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t]} \left\| \Lambda_\mu^\parallel(\nu, X_{N,\nu}) \right\|_{L_{2\mathbf{p}}} + \mathbf{p}'(\mathbf{p}) (1 - e^{-2\omega t})^{\frac{1}{2}} \sup_{\nu \in [0, t]} \left\| F_\sigma^\parallel(\nu, X_{N,\nu}) \right\|_{L_{2\mathbf{p}}} \right), \end{aligned}$$

for all $(t, \mathbf{p}) \in [0, t^* \wedge \mathbf{T}_s] \times \mathbb{N}_{\geq 1}$. Since $t \leq (\mathbf{T}_s \wedge t^*) \leq \mathbf{T}_s$, we can develop the bound above into

$$\begin{aligned} \left\| \tilde{\mathfrak{N}}_{\star}(t) \right\|_{L_{\mathbf{p}}} &\leq \frac{e^{2\lambda t}}{2\lambda\sqrt{\omega}} (1 - e^{-2\lambda \mathbf{T}_s}) \Delta_g \sup_{\nu \in [0, t^* \wedge \mathbf{T}_s]} \left\| \partial \mathcal{V}(Z_{N,\nu}) \right\|_{L_{2\mathbf{p}}} \\ &\quad \times \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^* \wedge \mathbf{T}_s]} \left\| \Lambda_\mu^\parallel(\nu, X_{N,\nu}) \right\|_{L_{2\mathbf{p}}} + \mathbf{p}'(\mathbf{p}) (1 - e^{-2\omega \mathbf{T}_s})^{\frac{1}{2}} \sup_{\nu \in [0, t^* \wedge \mathbf{T}_s]} \left\| F_\sigma^\parallel(\nu, X_{N,\nu}) \right\|_{L_{2\mathbf{p}}} \right), \end{aligned}$$

for all $(t, \mathbf{p}) \in [0, t^* \wedge \mathbf{T}_s] \times \mathbb{N}_{\geq 1}$. Hence, we get

$$\omega e^{-2\lambda t} \left\| \tilde{\mathfrak{N}}_{\star}(t) \right\|_{L_{\mathbf{p}}} \leq \tilde{\Delta}^-(\mathbf{p}, t^*, \omega, \mathbf{T}_s), \quad \forall (t, \mathbf{p}) \in [0, t^* \wedge \mathbf{T}_s] \times \mathbb{N}_{\geq 1}. \quad (\text{D.125})$$

Bound for $\tilde{\mathfrak{N}}_1(t)$: The process $\tilde{\mathfrak{N}}_1$ in (D.123) can be written as

$$\tilde{\mathfrak{N}}_1(t) = \int_{\mathbf{T}_s}^t e^{(2\lambda - \omega)\nu} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \mathcal{N}_1(\nu) d\nu$$

$$= \int_{\mathbf{T}_s}^t e^{(2\lambda - \omega)\nu} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \left(\int_0^\nu e^{\omega\beta} [\mathcal{K}_\mu(\beta, X_N) d\beta + \mathcal{K}_\sigma(\beta, X_N) dW_\beta] \right) d\nu,$$

where we have defined

$$\mathcal{K}_\mu(\beta, X_N) = \tilde{\mathcal{H}}_\mu^\parallel(\beta, X_N) + e^{(\lambda_s - \omega)\beta} \tilde{\mathcal{H}}_\mu(\beta, X_N) + \mathcal{G}_\mu(\beta, X_N), \quad (\text{D.126a})$$

$$\mathcal{K}_\sigma(\beta, X_N) = \tilde{\mathcal{H}}_\sigma^\parallel(\beta, X_N) + e^{(\lambda_s - \omega)\beta} \tilde{\mathcal{H}}_\sigma(\beta, X_N) + \tilde{\mathcal{G}}_{\sigma_1}(\beta, X_N) + e^{(\lambda_s - \omega)\beta} \tilde{\mathcal{G}}_{\sigma_2}(\beta, X_N). \quad (\text{D.126b})$$

Therefore, using Minkowski's inequality and the bound on $g(t)$ in Assumption 1, we obtain

$$\begin{aligned} \|\tilde{\mathfrak{N}}_1(t)\|_{L_p} &\leq \Delta_g \left\| \int_{\mathbf{T}_s}^t e^{(2\lambda - \omega)\nu} \|\partial \mathcal{V}(Z_{N,\nu})\| \left(\int_0^\nu e^{\omega\beta} \|\mathcal{K}_\mu(\beta, X_N)\| d\beta \right) d\nu \right\|_{L_p} \\ &\quad + \Delta_g \left\| \int_{\mathbf{T}_s}^t e^{(2\lambda - \omega)\nu} \|\partial \mathcal{V}(Z_{N,\nu})\| \left\| \int_0^\nu e^{\omega\beta} \mathcal{K}_\sigma(\beta, X_N) dW_\beta \right\| d\nu \right\|_{L_p}, \end{aligned}$$

for all $(t, p) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$. We now invoke Proposition B.2 to bound the first term by setting

$$\xi = \mathbf{T}_s, \quad \theta_1 = 2\lambda, \quad \theta_2 = \omega, \quad S_1(\nu) = \partial \mathcal{V}(Z_{N,\nu}), \quad S_2(\beta) = \mathcal{K}_\mu(\beta, X_N),$$

and we invoke Lemma B.3 to bound the second term by setting

$$\xi = \mathbf{T}_s, \quad \theta_1 = 2\lambda, \quad \theta_2 = \omega, \quad \{\zeta_1, \zeta_2\} = \{0, t\}, \quad S(\nu) = \partial \mathcal{V}(Z_{N,\nu}), \quad L(\beta) = \mathcal{K}_\sigma(\beta, X_N),$$

and obtain

$$\begin{aligned} \|\tilde{\mathfrak{N}}_1(t)\|_{L_p} &\leq \frac{e^{2\lambda t} - e^{2\lambda \mathbf{T}_s}}{2\lambda\sqrt{\omega}} \Delta_g \sup_{\nu \in [\mathbf{T}_s, t]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \\ &\quad \times \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [\mathbf{T}_s, t]} \|\mathcal{K}_\mu(\nu, X_N)\|_{L_{2p}} + \mathbf{p}'(\mathbf{p}) (1 - e^{-2\omega t})^{\frac{1}{2}} \sup_{\nu \in [0, t]} \|\mathcal{K}_\sigma(\beta, X_N)\|_{L_{2p}} \right), \end{aligned}$$

for all $(t, p) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$. Since $t \in [\mathbf{T}_s, t^*]$, we can further develop the bound above into

$$\begin{aligned} \|\tilde{\mathfrak{N}}_1(t)\|_{L_p} &\leq \frac{e^{2\lambda t}}{2\lambda\sqrt{\omega}} \Delta_g \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \\ &\quad \times \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^*]} \|\mathcal{K}_\mu(\nu, X_N)\|_{L_{2p}} + \mathbf{p}'(\mathbf{p}) \sup_{\nu \in [0, t^*]} \|\mathcal{K}_\sigma(\beta, X_N)\|_{L_{2p}} \right), \end{aligned}$$

for all $(t, p) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$. Therefore,

$$\begin{aligned} \omega e^{-2\lambda t} \|\tilde{\mathfrak{N}}_1(t)\|_{L_p} &\leq \frac{1}{2\lambda} \Delta_g \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \\ &\quad \times \left(\sup_{\nu \in [0, t^*]} \|\mathcal{K}_\mu(\nu, X_N)\|_{L_{2p}} + \sqrt{\omega} \mathbf{p}'(\mathbf{p}) \sup_{\nu \in [0, t^*]} \|\mathcal{K}_\sigma(\beta, X_N)\|_{L_{2p}} \right), \quad (\text{D.127}) \end{aligned}$$

for all $(t, p) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$. Using the definitions of \mathcal{K}_μ and \mathcal{K}_σ in (D.126), we use Minkowski's inequality and the bounds derived in Proposition D.11 to obtain the following:

$$\|\mathcal{K}_\mu(\nu, X_N)\|_{L_{2p}} \leq \left\| \tilde{\mathcal{H}}_\mu^\parallel(\nu, X_N) + e^{(\lambda_s - \omega)\nu} \tilde{\mathcal{H}}_\mu(\nu, X_N) \right\|_{L_{2p}} + \|\mathcal{G}_\mu(\nu, X_N)\|_{L_{2p}} \leq \Delta_{\tilde{\mathcal{H}}_\mu}(\mathbf{T}_s, \mathbf{p}, t^*) + \Delta_{\mathcal{G}_\mu}(\mathbf{T}_s, t^*),$$

and

$$\begin{aligned} \|\mathcal{K}_\sigma(\nu, X_N)\|_{L_{2p}} &\leq \left\| \tilde{\mathcal{H}}_\sigma^\parallel(\nu, X_N) + e^{(\lambda_s - \omega)\nu} \tilde{\mathcal{H}}_\sigma(\nu, X_N) \right\|_{L_{2p}} + \left\| \tilde{\mathcal{G}}_{\sigma_1}(\nu, X_N) + e^{(\lambda_s - \omega)\nu} \tilde{\mathcal{G}}_{\sigma_2}(\nu, X_N) \right\|_{L_{2p}} \\ &\leq \Delta_{\tilde{\mathcal{H}}_\sigma}(\mathbf{T}_s, \mathbf{p}, t^*) + \Delta_{\tilde{\mathcal{G}}_\sigma}(\mathbf{T}_s, \mathbf{p}, t^*), \end{aligned}$$

for all $(\nu, p) \in [0, t^*] \times \mathbb{N}_{\geq 1}$. Substituting the above bounds into (D.127) then leads to

$$\omega e^{-2\lambda t} \|\tilde{\mathfrak{N}}_1(t)\|_{L_p} \leq \tilde{\Delta}_1^+(\mathbf{p}, t^*, \omega, \mathbf{T}_s), \quad \forall (t, p) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}. \quad (\text{D.128})$$

Bound for $\tilde{\mathfrak{N}}_2(t)$: Recall the definition of the process $\tilde{\mathfrak{N}}_2$ in (D.123):

$$\begin{aligned}\tilde{\mathfrak{N}}_2(t) &= \int_{\mathbf{T}_s}^t e^{(2\lambda - \omega)\nu} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \mathcal{N}_2(\nu) d\nu \\ &= \int_{\mathbf{T}_s}^t e^{(2\lambda - \omega)\nu} \partial \mathcal{V}(Z_{N,\nu})^\top g(\nu) \left(\int_0^\nu e^{\omega\beta} (\mathcal{G}_{\sigma_1}(\beta, X_N) + \mathcal{G}_{\sigma_2}(\beta, X_N)) dW_\beta \right) d\nu, \quad t \in [\mathbf{T}_s, t^*],\end{aligned}$$

where we have the following due to the definitions of \mathcal{G}_{σ_1} and \mathcal{G}_{σ_2} in Proposition D.3:

$$\begin{aligned}\int_0^\nu e^{\omega\beta} \mathcal{G}_{\sigma_1}(\beta, X_N) dW_\beta &= \int_0^\nu e^{\omega\beta} \left(\sum_{i=1}^{i^*(\nu)-1} \mathbb{1}_{[(i-1)\mathbf{T}_s, i\mathbf{T}_s)}(\beta) F_\sigma^\parallel((i-1)\mathbf{T}_s, X_{N,(i-1)\mathbf{T}_s}) \left(1 - \tilde{\gamma}_i(\mathbf{T}_s) e^{(\lambda_s - \omega)\beta}\right) \right) dW_\beta \\ &\quad + \int_0^\nu e^{\omega\beta} \mathbb{1}_{[(i^*(\nu)-1)\mathbf{T}_s, i^*(\nu)\mathbf{T}_s)}(\beta) F_\sigma^\parallel((i^*(\nu)-1)\mathbf{T}_s, X_{N,(i^*(\nu)-1)\mathbf{T}_s}) dW_\beta,\end{aligned}$$

and

$$\begin{aligned}\int_0^\nu e^{\omega\beta} \mathcal{G}_{\sigma_2}(\beta, X_N) dW_\beta &= - \int_0^\nu e^{\omega\beta} F_\sigma^\parallel(i^*(\nu)\mathbf{T}_s, X_{N,i^*(\nu)\mathbf{T}_s}) \mathbb{1}_{[(i^*(\nu)-1)\mathbf{T}_s, i^*(\nu)\mathbf{T}_s)}(\beta) \tilde{\gamma}^*(\nu, \mathbf{T}_s) e^{(\lambda_s - \omega)\beta} dW_\beta \\ &\quad + \int_0^\nu e^{\omega\beta} F_\sigma^\parallel(i^*(\nu)\mathbf{T}_s, X_{N,i^*(\nu)\mathbf{T}_s}) \mathbb{1}_{\geq i^*(\nu)\mathbf{T}_s}(\beta) dW_\beta.\end{aligned}$$

Using the above, it is straightforward to establish that

$$\tilde{\mathfrak{N}}_2(t) = \sum_{j=1}^4 \mathfrak{J}_j(t), \quad t \in [\mathbf{T}_s, t^*],$$

where the processes \mathfrak{J}_j are defined in (D.111) in the statement of Proposition D.13. Hence, we obtain the following by the Minkowski's inequality and Proposition D.13:

$$\left\| \tilde{\mathfrak{N}}_2(t) \right\|_{L_p} \leq \sum_{j=1}^4 \left\| \mathfrak{J}_j(t) \right\|_{L_p} \leq \frac{e^{2\lambda t}}{2\lambda\sqrt{\omega}} \Delta_g \gamma''(\mathbf{p}, \omega, \mathbf{T}_s) \sup_{\nu \in [0, t^*]} \left\| \partial \mathcal{V}(Z_{N,\nu}) \right\|_{L_{2p}} \sup_{\nu \in [0, t^*]} \left\| F_\sigma^\parallel(\nu, X_{N,\nu}) \right\|_{L_{4p}}$$

for all $(t, p) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$, which further leads to

$$\omega e^{-2\lambda t} \left\| \tilde{\mathfrak{N}}_2(t) \right\|_{L_p} \leq \tilde{\Delta}_2^+(\mathbf{p}, t^*, \omega, \mathbf{T}_s), \quad \forall (t, \mathbf{p}) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}. \quad (\text{D.129})$$

Bound for $\tilde{\mathfrak{N}}_3(t)$: Using the definition of $\tilde{\mathfrak{N}}_3$ from (D.123), we obtain

$$\begin{aligned}\left\| \tilde{\mathfrak{N}}_3(t) \right\|_{L_p} &\leq \Delta_g \left\| \left(\int_{\mathbf{T}_s}^t e^{(2\lambda - \omega)\nu} \left\| \partial \mathcal{V}(Z_{N,\nu}) \right\| d\nu \right) \left(\int_0^{\mathbf{T}_s} e^{\omega\beta} \left\| \Lambda_\mu^\parallel(\beta, X_{N,\beta}) \right\| d\beta \right) \right\|_{L_p} \\ &\leq \Delta_g \left\| \int_{\mathbf{T}_s}^t e^{(2\lambda - \omega)\nu} \left\| \partial \mathcal{V}(Z_{N,\nu}) \right\| d\nu \right\|_{L_{2p}} \left\| \int_0^{\mathbf{T}_s} e^{\omega\beta} \left\| \Lambda_\mu^\parallel(\beta, X_{N,\beta}) \right\| d\beta \right\|_{L_{2p}},\end{aligned}$$

for all $(t, \mathbf{p}) \in [\mathbf{T}_s, t^*] \times \mathbb{N}_{\geq 1}$, where we have used the bound on $g(t)$ from Assumption 1 and applied the Cauchy-Schwarz inequality. Bounding both the integrals on the right hand side via Proposition B.1 leads to

$$\begin{aligned}\left\| \tilde{\mathfrak{N}}_3(t) \right\|_{L_p} &\leq \Delta_g \frac{e^{(2\lambda - \omega)t} - e^{(2\lambda - \omega)\mathbf{T}_s}}{|2\lambda - \omega|} \frac{e^{\omega\mathbf{T}_s} - 1}{\omega} \sup_{\nu \in [\mathbf{T}_s, t]} \left\| \partial \mathcal{V}(Z_{N,\nu}) \right\|_{L_{2p}} \sup_{\nu \in [0, \mathbf{T}_s]} \left\| \Lambda_\mu^\parallel(\nu, X_{N,\nu}) \right\|_{L_{2p}} \\ &\leq \Delta_g \frac{e^{2\lambda t}}{|2\lambda - \omega|} \frac{e^{\omega\mathbf{T}_s} - 1}{\omega} \sup_{\nu \in [0, t^*]} \left\| \partial \mathcal{V}(Z_{N,\nu}) \right\|_{L_{2p}} \sup_{\nu \in [0, t^*]} \left\| \Lambda_\mu^\parallel(\nu, X_{N,\nu}) \right\|_{L_{2p}},\end{aligned}$$

for all $(t, p) \in [T_s, t^*] \times \mathbb{N}_{\geq 1}$. Therefore,

$$\omega e^{-2\lambda t} \left\| \tilde{\mathfrak{N}}_3(t) \right\|_{L_p} \leq \tilde{\Delta}_3^+(p, t^*, \omega, T_s), \quad \forall (t, p) \in [T_s, t^*] \times \mathbb{N}_{\geq 1}, \quad (\text{D.130})$$

for all $(t, p) \in [T_s, t^*] \times \mathbb{N}_{\geq 1}$. The desired bound in (D.120) is then established by substituting (D.125), (D.128), (D.129), and (D.130) into (D.124). \square

We now further develop a few bounds that will eventually allow us to consolidate the results in Lemmas D.1, D.3, and D.4.

Proposition D.14 Consider the stopping times τ^* and t^* , as defined in (104), Lemma 4.5, and assume that $\tau^* = t^*$. Define the following for $p \in \{1, \dots, p^*\}$, where p^* is defined in Assumption 3:

$$\tilde{\mathcal{E}}(2p) \doteq \sup_{\nu \in [0, t^*]} \|X_{N,\nu} - X_{N,\nu}^r\|_{L_{2p}}, \quad \mathcal{E}(2p) \doteq \sup_{\nu \in [0, t^*]} \|X_{N,\nu}\|_{L_{2p}}, \quad \mathcal{E}_r(2p) \doteq \sup_{\nu \in [0, t^*]} \|X_{N,\nu}^r\|_{L_{2p}}. \quad (\text{D.131})$$

If Assumptions 1 and 6 hold, then we have the following bounds for all $p \in \{1, \dots, p^*\}$:

$$\Delta_{\mathcal{P}_1}(p, t^*) \leq \hat{\Delta}_1(p) + \hat{\Delta}_2(p) \sqrt{\tilde{\mathcal{E}}(2p)} + \hat{\Delta}_3(\mathcal{E}(2p) + \mathcal{E}_r(2p)) + \hat{\Delta}_4 \tilde{\mathcal{E}}(2p), \quad (\text{D.132a})$$

$$\Delta_{\mathcal{P}_3}(p, t^*) \leq \hat{\Delta}_5 \sqrt{\tilde{\mathcal{E}}(2p)}, \quad (\text{D.132b})$$

where $\Delta_{\mathcal{P}_1}(p, t^*)$ and $\Delta_{\mathcal{P}_3}(p, t^*)$ are defined in the statement of Proposition D.8, and the constants $\hat{\Delta}_1(p)$, $\hat{\Delta}_2(p)$, and $\hat{\Delta}_{\{3, \dots, 6\}}$ are defined in (A.8).

Proof. We begin with the term $\Delta_{\mathcal{P}_\mu}$ defined in Proposition D.7 and the definitions $\mathring{f}(\nu, Z_{N,\nu}) = f(\nu, X_{N,\nu}) - f(\nu, X_{N,\nu}^r)$ and $\mathring{\Lambda}_\mu(\nu, Z_{N,\nu}) = \Lambda_\mu(\nu, X_{N,\nu}) - \Lambda_\mu(\nu, X_{N,\nu}^r)$ from (103) to get

$$\begin{aligned} \Delta_{\mathcal{P}_\mu}(p) &= \sup_{\nu \in [0, t^*]} \left(\left\| \mathring{f}(\nu, Z_{N,\nu}) \right\|_{L_{2p}} + \left\| \mathring{\Lambda}_\mu(\nu, Z_{N,\nu}) \right\|_{L_{2p}} \right) \\ &= \sup_{\nu \in [0, t^*]} \left(\left\| f(\nu, X_{N,\nu}) - f(\nu, X_{N,\nu}^r) \right\|_{L_{2p}} + \left\| \Lambda_\mu(\nu, X_{N,\nu}) - \Lambda_\mu(\nu, X_{N,\nu}^r) \right\|_{L_{2p}} \right), \end{aligned} \quad (\text{D.133a})$$

$$\leq \sup_{\nu \in [0, t^*]} \left(\left\| f(\nu, X_{N,\nu}) \right\|_{L_{2p}} + \left\| f(\nu, X_{N,\nu}^r) \right\|_{L_{2p}} + \left\| \Lambda_\mu(\nu, X_{N,\nu}) - \Lambda_\mu(\nu, X_{N,\nu}^r) \right\|_{L_{2p}} \right), \quad (\text{D.133b})$$

where the second inequality is due to the Minkowski inequality. Using Assumptions 1 and 6, the bound in (D.133b) can be developed into

$$\Delta_{\mathcal{P}_\mu}(p) \leq 2\Delta_f + \Delta_f(\mathcal{E}(2p) + \mathcal{E}_r(2p)) + L_\mu \tilde{\mathcal{E}}(2p), \quad (\text{D.134})$$

where we have also used the subadditivity of the square root function. If additionally Assumption 8 holds, then we obtain a tighter bound by developing (D.133a) to get

$$\Delta_{\mathcal{P}_\mu}(p) \leq (L_f + L_\mu) \tilde{\mathcal{E}}(2p). \quad (\text{D.135})$$

We thus combine the bounds in (D.134) and (D.135) by writing

$$\Delta_{\mathcal{P}_\mu}(p) \leq 2\Delta_f(1 - \mathbb{1}_{\text{Lip}}(f)) + \Delta_f(1 - \mathbb{1}_{\text{Lip}}(f))(\mathcal{E}(2p) + \mathcal{E}_r(2p)) + (L_\mu + L_f \mathbb{1}_{\text{Lip}}(f)) \tilde{\mathcal{E}}(2p), \quad (\text{D.136})$$

where $\mathbb{1}_{\text{Lip}}(f)$ is defined in (A.1b).

Next, we have the following from Proposition D.7:

$$\begin{aligned} \Delta_{\mathcal{P}_\sigma}(p) &= \sup_{\nu \in [0, t^*]} \left\| \mathring{F}_\sigma(\nu, Z_{N,\nu}) \right\|_{L_{2p}} \\ &= \sup_{\nu \in [0, t^*]} \left\| F_\sigma(\nu, X_{N,\nu}) - F_\sigma(\nu, X_{N,\nu}^r) \right\|_{L_{2p}} \\ &\leq \sup_{\nu \in [0, t^*]} \left(\left\| p(\nu, X_{N,\nu}) - p(\nu, X_{N,\nu}^r) \right\|_{L_{2p}} + \left\| \Lambda_\sigma(\nu, X_{N,\nu}) - \Lambda_\sigma(\nu, X_{N,\nu}^r) \right\|_{L_{2p}} \right) \end{aligned}$$

$$\leq \sup_{\nu \in [0, t^*]} \left(\|p(\nu, X_{N,\nu})\|_{L_{2p}} + \|p(\nu, X_{N,\nu}^r)\|_{L_{2p}} + \|\Lambda_\sigma(\nu, X_{N,\nu}) - \Lambda_\sigma(\nu, X_{N,\nu}^r)\|_{L_{2p}} \right),$$

where we have used the definition $\bar{F}_\sigma(\nu, X_{N,\nu}) = F_\sigma(\nu, X_{N,\nu}) - F_\sigma(\nu, X_{N,\nu}^r)$ from (103), and $F_\sigma = \bar{F}_\sigma + \Lambda_\sigma = p + \Lambda_\sigma$. Then, as above, it follows from Assumptions 1 and 6 that

$$\Delta_{\mathcal{P}_\sigma}(\mathbf{p}) \leq (L_p + L_\sigma) \sqrt{\tilde{\mathcal{E}}(2\mathbf{p})}, \quad (\text{D.137})$$

where $\mathbb{1}_{\text{Lip}}(p)$ is defined in (A.1b) and we have used manipulations similar to (C.72). Using (D.136) and (D.137), we can now bound the terms in (D.41) as follows:

$$\begin{aligned} \sup_{\nu \in [0, t^*]} \|\mathcal{P}_\mu(\nu)\|_{L_{2p}} &\leq 4\Delta_g \Delta_f (1 - \mathbb{1}_{\text{Lip}}(f)) + 2\Delta_g \Delta_f (1 - \mathbb{1}_{\text{Lip}}(f)) (\mathcal{E}(2\mathbf{p}) + \mathcal{E}_r(2\mathbf{p})) \\ &\quad + 2(\Delta_g (L_\mu + L_f \mathbb{1}_{\text{Lip}}(f)) + \Delta_{\dot{g}}) \tilde{\mathcal{E}}(2\mathbf{p}), \end{aligned} \quad (\text{D.138a})$$

$$\sup_{\nu \in [0, t^*]} \|\mathcal{P}_\sigma(\nu)\|_{L_q} \leq 2\Delta_g (L_p + L_\sigma) \sqrt{\tilde{\mathcal{E}}(2\mathbf{p})}, \quad (\text{D.138b})$$

where we have used the definition $\partial\mathcal{V}(Z_{N,\nu}) = 2(X_{N,\nu} - X_{N,\nu}^r)$ from the statement of Lemma 4.5. Now, recall the following definitions of $\Delta_{\mathcal{P}_1}(\mathbf{p}, t^*)$ and $\Delta_{\mathcal{P}_3}(\mathbf{p}, t^*)$ from the statement of Proposition D.8:

$$\begin{aligned} \Delta_{\mathcal{P}_1}(\mathbf{p}, t^*) &= \frac{1}{2\sqrt{\lambda}} \sup_{\nu \in [0, t^*]} \|\mathcal{P}_\mu(\nu)\|_{L_{2p}} + \frac{\mathbf{p}(\mathbf{p})}{2} \sup_{\nu \in [0, t^*]} \|\mathcal{P}_\sigma(\nu)\|_{L_{2p}}, \\ \Delta_{\mathcal{P}_3}(\mathbf{p}, t^*) &= \frac{\sqrt{m}}{2} \sup_{\nu \in [0, t^*]} \|\mathcal{P}_\sigma(\nu)\|_{L_{2p}}. \end{aligned}$$

Substituting (D.138) into the above then leads to (D.132). □

The next result bounds the terms derived in Proposition D.11.

Proposition D.15 Consider the stopping times τ^* and t^* , as defined in (104), Lemma 4.5, and assume that $\tau^* = t^*$. Define the following for $\mathbf{p} \in \{1, \dots, \mathbf{p}^*\}$, where \mathbf{p}^* is defined in Assumption 3:

$$\mathcal{E}(2\mathbf{p}) \doteq \sup_{\nu \in [0, t^*]} \|X_{N,\nu}\|_{L_{2p}}. \quad (\text{D.139})$$

If Assumptions 1 and 4 hold, then we have the following bounds for all $\mathbf{p} \in \{1, \dots, \mathbf{p}^*\}$:

$$\begin{aligned} &\Delta_{\tilde{\mathcal{H}}_\mu}(\mathbf{T}_s, \mathbf{p}, t^*) + \Delta_{\mathcal{G}_\mu}(\mathbf{T}_s, t^*) + \sqrt{\omega} \mathbf{p}'(\mathbf{p}) \left[\Delta_{\tilde{\mathcal{H}}_\sigma}(\mathbf{T}_s, \mathbf{p}, t^*) + \Delta_{\tilde{\mathcal{G}}_\sigma}(\mathbf{T}_s, \mathbf{p}, t^*) \right] \\ &\leq \Delta_{\mu_1}(\mathbf{p}, \omega, \mathbf{T}_s) + \sqrt{\omega} \mathbf{p}'(\mathbf{p}) \Delta_{\sigma_1}(\mathbf{p}, \omega, \mathbf{T}_s) + \sqrt{\omega} \mathbf{p}'(\mathbf{p}) \Delta_{\sigma_2}(\mathbf{p}, \omega, \mathbf{T}_s) (\mathcal{E}(2\mathbf{p}))^{\frac{1}{4}} \\ &\quad + (\Delta_{\mu_2}(\mathbf{p}, \omega, \mathbf{T}_s) + \sqrt{\omega} \mathbf{p}'(\mathbf{p}) \Delta_{\sigma_3}(\mathbf{p}, \omega, \mathbf{T}_s)) \sqrt{\mathcal{E}(2\mathbf{p})} + \Delta_{\mu_3}(\mathbf{p}, \omega, \mathbf{T}_s) \mathcal{E}(2\mathbf{p}), \end{aligned} \quad (\text{D.140})$$

where $\Delta_{\tilde{\mathcal{H}}_\mu}$, $\Delta_{\tilde{\mathcal{H}}_\sigma}$, $\Delta_{\tilde{\mathcal{G}}_\sigma}$, and $\Delta_{\mathcal{G}_\mu}$ are defined in Proposition D.11, and the constants $\Delta_{\mu_{\{1,2,3\}}}$ and $\Delta_{\sigma_{\{1,2,3\}}}$ are defined in (A.16) and (A.17), respectively.

Proof. We begin by using $\bar{F}_\mu(\nu, \cdot) = f(\nu, \cdot)$ and $F_\sigma(\nu, \cdot) = p(\nu, \cdot) + \Lambda_\sigma(\nu, \cdot)$ from Definition 1, along with Assumptions 1 and 4 to arrive at the following bounds:

$$\sup_{\nu \in [0, t^*]} \|\bar{F}_\mu(\nu, X_{N,\nu})\|_{L_{2p}} \leq \Delta_f + \Delta_f \mathcal{E}(2\mathbf{p}), \quad \sup_{\nu \in [0, t^*]} \|F_\sigma(\nu, X_{N,\nu})\|_{L_{2p}} \leq \Delta_p + \Delta_\sigma + \Delta_\sigma \sqrt{\mathcal{E}(2\mathbf{p})}, \quad (\text{D.141a})$$

$$\sup_{\nu \in [0, t^*]} \|\Lambda_\mu(\nu, X_{N,\nu})\|_{L_{2p}} \leq \Delta_\mu + \Delta_\mu \mathcal{E}(2\mathbf{p}), \quad (\text{D.141b})$$

where we have further used the subadditivity of the square root function and performed manipulations similar to (C.72). Now, recall the following from Proposition D.10 for $\tau^* = t^*$:

$$\begin{aligned} \tilde{\delta}_1(\mathbf{T}_s, \mathbf{p}, t^*) &= 2 \sup_{\nu \in [0, t^*]} \|\bar{F}_\mu(\nu, X_{N,\nu})\|_{L_{2p}} + 2(1 + \Delta_g \Delta_\Theta e^{-\lambda_s \mathbf{T}_s}) \sup_{\nu \in [0, t^*]} \|\Lambda_\mu(\nu, X_{N,\nu})\|_{L_{2p}} \\ &\quad + 2\mathbf{p}'(\mathbf{p}) \Delta_g \Delta_\Theta e^{-\lambda_s \mathbf{T}_s} \gamma_1(\mathbf{T}_s) \sup_{\nu \in [0, t^*]} \|F_\sigma(\nu, X_{N,\nu})\|_{L_{2p}}, \end{aligned}$$

$$\tilde{\delta}_2(\mathbf{p}, t^*) = 2\sqrt{2}\mathbf{p}'(\mathbf{p}) \sup_{\nu \in [0, t^*]} \|F_\sigma(\nu, X_{N,\nu})\|_{L_{2p}}.$$

Substituting (D.141) into the above leads to

$$\begin{aligned} \tilde{\delta}_1(\mathbf{T}_s, \mathbf{p}, t^*) &\leq 2\Delta_f + 2(1 + \Delta_g \Delta_\Theta e^{-\lambda_s \mathbf{T}_s}) \Delta_\mu + 2\mathbf{p}'(\mathbf{p}) \Delta_g \Delta_\Theta e^{-\lambda_s \mathbf{T}_s} \gamma_1(\mathbf{T}_s) (\Delta_p + \Delta_\sigma) \\ &\quad + 2\mathbf{p}'(\mathbf{p}) \Delta_g \Delta_\Theta e^{-\lambda_s \mathbf{T}_s} \gamma_1(\mathbf{T}_s) \Delta_\sigma \sqrt{\mathcal{E}(2\mathbf{p})} \\ &\quad + 2(\Delta_f + (1 + \Delta_g \Delta_\Theta e^{-\lambda_s \mathbf{T}_s}) \Delta_\mu) \mathcal{E}(2\mathbf{p}), \\ \tilde{\delta}_2(\mathbf{p}, t^*) &\leq 2\sqrt{2}\mathbf{p}'(\mathbf{p}) (\Delta_p + \Delta_\sigma) + 2\sqrt{2}\mathbf{p}'(\mathbf{p}) \Delta_\sigma \sqrt{\mathcal{E}(2\mathbf{p})}. \end{aligned}$$

We further substitute these bounds into $\Delta_{\tilde{\mathcal{H}}_\mu}$, $\Delta_{\tilde{\mathcal{H}}_\sigma}$, $\Delta_{\tilde{\mathcal{G}}_\sigma}$, and $\Delta_{\mathcal{G}_\mu}$ that are defined in Proposition D.11 to obtain:

$$\Delta_{\tilde{\mathcal{H}}_\mu}(\mathbf{T}_s, \mathbf{p}, t^*) \leq \mathbf{T}_s \hat{\gamma}_\mu(\omega, \mathbf{T}_s) + \bar{\Delta}_1(\mathbf{T}_s, \mathbf{p}) \gamma_\mu(\omega, \mathbf{T}_s) + \bar{\Delta}_2(\mathbf{T}_s, \mathbf{p}) \gamma_\mu(\omega, \mathbf{T}_s) \sqrt{\mathcal{E}(2\mathbf{p})} + \gamma_\mu(\omega, \mathbf{T}_s) \bar{\Delta}_3(\mathbf{T}_s) \mathcal{E}(2\mathbf{p}), \quad (\text{D.142a})$$

$$\begin{aligned} \Delta_{\tilde{\mathcal{H}}_\sigma}(\mathbf{T}_s, \mathbf{p}, t^*) &\leq \gamma'_\sigma(\omega, \mathbf{T}_s) + \gamma_\sigma(\omega, \mathbf{T}_s) (\bar{\Delta}_1(\mathbf{T}_s, \mathbf{p}))^{\frac{1}{2}} + \mathbf{T}_s \hat{\gamma}_\sigma(\omega, \mathbf{T}_s) \\ &\quad + \gamma_\sigma(\omega, \mathbf{T}_s) (\bar{\Delta}_2(\mathbf{T}_s, \mathbf{p}))^{\frac{1}{2}} (\mathcal{E}(2\mathbf{p}))^{\frac{1}{4}} + \gamma_\sigma(\omega, \mathbf{T}_s) (\bar{\Delta}_3(\mathbf{T}_s))^{\frac{1}{2}} \sqrt{\mathcal{E}(2\mathbf{p})}, \end{aligned} \quad (\text{D.142b})$$

$$\begin{aligned} \Delta_{\tilde{\mathcal{G}}_\sigma}(\mathbf{T}_s, \mathbf{p}, t^*) &\leq (1 + \gamma_2(\omega, \mathbf{T}_s)) \left[(L_p^\parallel + L_\sigma^\parallel) (\bar{\Delta}_1(\mathbf{T}_s, \mathbf{p}))^{\frac{1}{2}} + \mathbf{T}_s (\hat{L}_p^\parallel + \hat{L}_\sigma^\parallel) \right] \\ &\quad + (1 + \gamma_2(\omega, \mathbf{T}_s)) (L_p^\parallel + L_\sigma^\parallel) (\bar{\Delta}_2(\mathbf{T}_s, \mathbf{p}))^{\frac{1}{2}} (\mathcal{E}(2\mathbf{p}))^{\frac{1}{4}} \\ &\quad + (1 + \gamma_2(\omega, \mathbf{T}_s)) (L_p^\parallel + L_\sigma^\parallel) (\bar{\Delta}_3(\mathbf{T}_s))^{\frac{1}{2}} \sqrt{\mathcal{E}(2\mathbf{p})}, \end{aligned} \quad (\text{D.142c})$$

$$\Delta_{\mathcal{G}_\mu}(\mathbf{T}_s, t^*) \leq \Delta_\mu^\parallel (1 - e^{-\lambda_s \mathbf{T}_s}) + \Delta_\mu^\parallel (1 - e^{-\lambda_s \mathbf{T}_s}) \mathcal{E}(2\mathbf{p}), \quad (\text{D.142d})$$

where the constants $\bar{\Delta}_{\{1,2,3\}}$ are defined in (A.15) and we have used the following due to Assumption 4 and the subadditivity of the square root function:

$$\sup_{\nu \in [0, t^* \wedge \mathbf{T}_s]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu})\|_{L_{2p}} \leq \sup_{\nu \in [0, t^*]} \|\Lambda_\mu^\parallel(\nu, X_{N,\nu})\|_{L_{2p}} \leq \Delta_\mu^\parallel + \Delta_\mu^\parallel \mathcal{E}(2\mathbf{p}). \quad (\text{D.143})$$

Using all of the derived bounds above, we obtain the following:

$$\Delta_{\tilde{\mathcal{H}}_\mu}(\mathbf{T}_s, \mathbf{p}, t^*) + \Delta_{\mathcal{G}_\mu}(\mathbf{T}_s, t^*) \leq \Delta_{\mu_1}(\mathbf{p}, \omega, \mathbf{T}_s) + \Delta_{\mu_2}(\mathbf{p}, \omega, \mathbf{T}_s) \sqrt{\mathcal{E}(2\mathbf{p})} + \Delta_{\mu_3}(\mathbf{p}, \omega, \mathbf{T}_s) \mathcal{E}(2\mathbf{p}),$$

and

$$\Delta_{\tilde{\mathcal{H}}_\sigma}(\mathbf{T}_s, \mathbf{p}, t^*) + \Delta_{\tilde{\mathcal{G}}_\sigma}(\mathbf{T}_s, \mathbf{p}, t^*) \leq \Delta_{\sigma_1}(\mathbf{p}, \omega, \mathbf{T}_s) + \Delta_{\sigma_2}(\mathbf{p}, \omega, \mathbf{T}_s) (\mathcal{E}(2\mathbf{p}))^{\frac{1}{4}} + \Delta_{\sigma_3}(\mathbf{p}, \omega, \mathbf{T}_s) \sqrt{\mathcal{E}(2\mathbf{p})}.$$

The expression in (D.140) then follows in a straightforward manner. □

We conclude the appendix by collecting and developing the bounds from Lemmas D.1, D.3, and D.4.

Lemma D.5 Consider the stopping times τ^* and t^* , as defined in (104), and assume that $\tau^* = t^*$. As in (D.131), define the following for $\mathbf{p} \in \{1, \dots, \mathbf{p}^*\}$, where \mathbf{p}^* is defined in Assumption 3:

$$\tilde{\mathcal{E}}(2\mathbf{p}) \doteq \sup_{\nu \in [0, t^*]} \|X_{N,\nu} - X_{N,\nu}^r\|_{L_{2p}}, \quad \mathcal{E}(2\mathbf{p}) \doteq \sup_{\nu \in [0, t^*]} \|X_{N,\nu}\|_{L_{2p}}, \quad \mathcal{E}_r(2\mathbf{p}) \doteq \sup_{\nu \in [0, t^*]} \|X_{N,\nu}^r\|_{L_{2p}}. \quad (\text{D.144})$$

If Assumptions 1, 4, 5 and 6 hold, then we have the following for all $(t, \mathbf{p}) \in \mathbb{R}_{\geq 0} \times \{2, \dots, \mathbf{p}^*\}$:

$$\begin{aligned} &\left\| e^{-2\lambda\tau(t)} \Xi(\tau(t), Z_N) \right\|_{L_p} + \left\| e^{-(2\lambda + \omega)\tau(t)} \Xi_{\mathcal{U}}(\tau(t), Z_N) \right\|_{L_p} + \left\| e^{-2\lambda\tau(t)} \tilde{\Xi}_{\mathcal{U}}(\tau(t), Z_N) \right\|_{L_p} \\ &\leq \Delta_{\odot}(\mathbf{p}, \omega) \left(\tilde{\mathcal{E}}(2\mathbf{p}) \right)^{\frac{1}{2}} + \Delta_{\odot}(\mathbf{p}, \omega) \tilde{\mathcal{E}}(2\mathbf{p}) + \left(\tilde{\Delta}_{\otimes}(\mathbf{p}, \omega) (\mathcal{E}(2\mathbf{p}) + \mathcal{E}_r(2\mathbf{p})) + \Delta_{\otimes}(\mathbf{p}, \omega) \tilde{\mathcal{E}}(2\mathbf{p}) \right) \left(\tilde{\mathcal{E}}(2\mathbf{p}) \right)^{\frac{1}{2}} \\ &\quad + \left(\tilde{\Delta}_{\otimes}(\omega) (\mathcal{E}(2\mathbf{p}) + \mathcal{E}_r(2\mathbf{p})) + \Delta_{\otimes}(\omega) \tilde{\mathcal{E}}(2\mathbf{p}) \right) \tilde{\mathcal{E}}(2\mathbf{p}) + \sum_{i=1}^3 \Upsilon_{a_i} \left(\mathcal{E}(2\mathbf{p}), \tilde{\mathcal{E}}(2\mathbf{p}); \mathbf{p}, \omega, \mathbf{T}_s \right), \end{aligned} \quad (\text{D.145})$$

where

$$\begin{aligned}\Delta_{\odot}(\mathbf{p}, \omega) &= \frac{1}{|2\lambda - \omega|} \sqrt{\omega} \Delta_{\odot_1}(\mathbf{p}), \\ \Delta_{\odot}(\mathbf{p}, \omega) &= \frac{\Delta_{\odot_1}}{2\lambda} + \frac{1}{|2\lambda - \omega|} (\Delta_{\odot_2}(\mathbf{p}) + \sqrt{\omega} \Delta_{\odot_3}(\mathbf{p}) + \omega \Delta_{\odot_4}(\mathbf{p})), \\ \check{\Delta}_{\otimes}(\mathbf{p}, \omega) &= \sqrt{\omega} \frac{\Delta_{\otimes_3}(\mathbf{p})}{|2\lambda - \omega|}, \quad \Delta_{\otimes}(\mathbf{p}, \omega) = \frac{\Delta_{\otimes_1}(\mathbf{p})}{2\sqrt{\lambda}} + \frac{1}{|2\lambda - \omega|} (\Delta_{\otimes_2}(\mathbf{p}) + \sqrt{\omega} \Delta_{\otimes_4}(\mathbf{p})), \\ \check{\Delta}_{\oplus}(\omega) &= \frac{\Delta_{\oplus_2}}{|2\lambda - \omega|}, \quad \Delta_{\oplus}(\omega) = \left(\frac{\Delta_{\oplus_1}}{2\lambda} + \frac{\Delta_{\oplus_3}}{|2\lambda - \omega|} \right),\end{aligned}$$

and where the constants $\Delta_i(\mathbf{p})$, $i \in \{\odot_1, \odot_1, \dots, \odot_4, \otimes_1, \dots, \otimes_4, \oplus_1, \dots, \oplus_3\}$, are defined in (A.9) - (A.12) in Section A.2, and $\Upsilon_{a_1 \dots a_3}$ are defined in (25) in Section 3.2.2.

Lastly, the right hand side of (D.145), evaluated at $\mathbf{p} = 1$, is an upper bound for the following term, for all $t \in \mathbb{R}_{\geq 0}$:

$$\mathbb{E} \left[e^{-2\lambda\tau(t)} \Xi(\tau(t), Z_N) \right] + \mathbb{E} \left[e^{-(2\lambda + \omega)\tau(t)} \Xi_{\mathcal{U}}(\tau(t), Z_N) \right] + \mathbb{E} \left[e^{-2\lambda\tau(t)} \check{\Xi}_{\mathcal{U}}(\tau(t), Z_N) \right].$$

Proof. We begin by combining the results from Lemmas D.1, D.3, and D.4 to obtain the following bounds that hold for all $t \in [0, t^*]$:

$$\begin{aligned}\mathbb{E} \left[e^{-2\lambda\tau(t)} \Xi(\tau(t), Z_N) \right] &+ \mathbb{E} \left[e^{-(2\lambda + \omega)\tau(t)} \Xi_{\mathcal{U}}(\tau(t), Z_N) \right] + \mathbb{E} \left[e^{-2\lambda\tau(t)} \check{\Xi}_{\mathcal{U}}(\tau(t), Z_N) \right] \\ &\leq \frac{\Delta_{\Xi_1}(1, t^*)}{2\lambda} + \check{\Delta}_{\mathcal{U}}(1, t^*, \omega, \mathbf{T}_s) \\ &\quad + \frac{1}{|2\lambda - \omega|} \left(\Delta_{\mathcal{U}_1}(1, t^*) + \sqrt{\omega} \Delta_{\mathcal{U}_2}(1, t^*) + \omega \Delta_{\mathcal{U}_3}(1, t^*) + \frac{\omega}{\lambda} \Delta_{\mathcal{U}}(1, t^*, \omega, \mathbf{T}_s) \right),\end{aligned}\tag{D.146a}$$

$$\begin{aligned}\left\| e^{-2\lambda\tau(t)} \Xi(\tau(t), Z_N) \right\|_{L_p} &+ \left\| e^{-(2\lambda + \omega)\tau(t)} \Xi_{\mathcal{U}}(\tau(t), Z_N) \right\|_{L_p} + \left\| e^{-2\lambda\tau(t)} \check{\Xi}_{\mathcal{U}}(\tau(t), Z_N) \right\|_{L_p} \\ &\leq \frac{\Delta_{\Xi_1}(\mathbf{p}, t^*)}{2\lambda} + \frac{\Delta_{\Xi_2}(\mathbf{p}, t^*)}{2\sqrt{\lambda}} + \check{\Delta}_{\mathcal{U}}(\mathbf{p}, t^*, \omega, \mathbf{T}_s) \\ &\quad + \frac{1}{|2\lambda - \omega|} \left(\Delta_{\mathcal{U}_1}(\mathbf{p}, t^*) + \sqrt{\omega} \Delta_{\mathcal{U}_2}(\mathbf{p}, t^*) + \omega \Delta_{\mathcal{U}_3}(\mathbf{p}, t^*) + \frac{\omega}{\lambda} \Delta_{\mathcal{U}}(\mathbf{p}, t^*, \omega, \mathbf{T}_s) \right), \quad \forall \mathbf{p} \in \mathbb{N}_{\geq 2}.\end{aligned}\tag{D.146b}$$

We now obtain the expressions for each of the bounds.

Bounds for Δ_{Ξ_1} and Δ_{Ξ_2} : We have from Definitions 1 and 3 that $F_{\sigma}(\nu, \cdot) = p(\nu, \cdot) + \Lambda_{\sigma}(\nu, \cdot)$ and $F_{\sigma}^{\perp}(\nu, \cdot) = p^{\perp}(\nu, \cdot) + \Lambda_{\sigma}^{\perp}(\nu, \cdot)$, respectively. Furthermore, $\partial \mathcal{V}(Z_{N,\nu}) = 2(X_{N,\nu} - X_{N,\nu}^r)$ as defined in the statement of Lemma 4.5. Hence, using Assumptions 1, 5, and 6, we obtain the following bounds:

$$\sup_{\nu \in [0, t^*] \wedge \mathbf{T}_s} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \leq \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N,\nu})\|_{L_{2p}} \leq 2\tilde{\mathcal{E}}(2\mathbf{p}),\tag{D.147a}$$

$$\sup_{\nu \in [0, t^*]} \left\| \dot{\Lambda}_{\mu}^{\perp}(\nu, Z_{N,\nu}) \right\|_{L_{2p}} = \sup_{\nu \in [0, t^*]} \left\| \Lambda_{\mu}^{\perp}(\nu, X_{N,\nu}) - \Lambda_{\mu}^{\perp}(\nu, X_{N,\nu}^r) \right\|_{L_{2p}} \leq L_{\mu}^{\perp} \tilde{\mathcal{E}}(2\mathbf{p}),\tag{D.147b}$$

$$\begin{aligned}\sup_{\nu \in [0, t^*]} \left\| \dot{F}_{\sigma}(\nu, Z_{N,\nu}) \right\|_{L_{2p}} &= \sup_{\nu \in [0, t^*]} \left\| F_{\sigma}(\nu, X_{N,\nu}) - F_{\sigma}(\nu, X_{N,\nu}^r) \right\|_{L_{2p}} \\ &\leq \sup_{\nu \in [0, t^*]} \left(\left\| p(\nu, X_{N,\nu}) - p(\nu, X_{N,\nu}^r) \right\|_{L_{2p}} + \left\| \Lambda_{\sigma}(\nu, X_{N,\nu}) - \Lambda_{\sigma}(\nu, X_{N,\nu}^r) \right\|_{L_{2p}} \right) \\ &\leq (L_p + L_{\sigma}) \sqrt{\tilde{\mathcal{E}}(2\mathbf{p})},\end{aligned}\tag{D.147c}$$

where we have also used the definitions in (103). We similarly obtain

$$\sup_{\nu \in [0, t^*]} \left\| \dot{F}_{\sigma}^{\perp}(\nu, Z_{N,\nu}) \right\|_{L_{2p}} \leq (L_p^{\perp} + L_{\sigma}^{\perp}) \sqrt{\tilde{\mathcal{E}}(2\mathbf{p})}.\tag{D.148}$$

Substituting the bounds above into the following from Lemma D.1:

$$\begin{aligned}\Delta_{\Xi_1}(\mathbf{p}, t^*) &= \sup_{\nu \in [0, t^*]} \left(\Delta_g^\perp \|\partial \mathcal{V}(Z_{N, \nu})\|_{L_{2p}} \left\| \dot{\Lambda}_\mu^\perp(\nu, Z_{N, \nu}) \right\|_{L_{2p}} + \left(\left\| \dot{F}_\sigma^\perp(\nu, Z_{N, \nu}) \right\|_{L_{2p}} \right)^2 \right), \\ \Delta_{\Xi_2}(\mathbf{p}, t^*) &= \Delta_g^\perp \mathbf{p}(\mathbf{p}) \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N, \nu})\|_{L_{2p}} \left\| \dot{F}_\sigma^\perp(\nu, Z_{N, \nu}) \right\|_{L_{2p}},\end{aligned}$$

leads to the following for $\mathbf{p} \in \{1, \dots, \mathbf{p}^*\}$:

$$\Delta_{\Xi_1}(\mathbf{p}, t^*) \leq \Delta_{\odot_1} \tilde{\mathcal{E}}(2\mathbf{p}) + \Delta_{\otimes_1} \tilde{\mathcal{E}}(2\mathbf{p})^2, \quad \Delta_{\Xi_2}(\mathbf{p}, t^*) \leq \Delta_{\otimes_1}(\mathbf{p}) \left(\tilde{\mathcal{E}}(2\mathbf{p}) \right)^{\frac{3}{2}}. \quad (\text{D.149})$$

Bounds for $\Delta_{\mathcal{U}_1}$, $\Delta_{\mathcal{U}_2}$, $\Delta_{\mathcal{U}_3}$: Recall the following from Lemma D.3:

$$\begin{aligned}\Delta_{\mathcal{U}_1}(\mathbf{p}, t^*) &= \sqrt{\lambda} \Delta_g \mathbf{p}(\mathbf{p}) \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N, \nu})\|_{L_{2p}} \sup_{\nu \in [0, t^*]} \left\| \dot{F}_\sigma^\parallel(\nu, Z_{N, \nu}) \right\|_{L_{2p}} \\ &\quad + \left(\frac{\Delta_{\mathcal{P}_1}(\mathbf{p}, t^*)}{\sqrt{\lambda}} + 2\Delta_g \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N, \nu})\|_{L_{2p}} \right. \\ &\quad \left. + \frac{\Delta_g^2}{\lambda} \sup_{\nu \in [0, t^*]} \left\| \dot{\Lambda}_\mu^\parallel(\nu, Z_{N, \nu}) \right\|_{L_{2p}} \right) \sup_{\nu \in [0, t^*]} \left\| \dot{\Lambda}_\mu^\parallel(\nu, Z_{N, \nu}) \right\|_{L_{2p}}, \\ \Delta_{\mathcal{U}_2}(\mathbf{p}, t^*) &= \mathbf{p}'(\mathbf{p}) \left(\frac{\Delta_{\mathcal{P}_1}(\mathbf{p}, t^*)}{\sqrt{\lambda}} + \Delta_g \sup_{\nu \in [0, t^*]} \|\partial \mathcal{V}(Z_{N, \nu})\|_{L_{2p}} + 2 \frac{\Delta_g^2}{\lambda} \sup_{\nu \in [0, t^*]} \left\| \dot{\Lambda}_\mu^\parallel(\nu, Z_{N, \nu}) \right\|_{L_{2p}} \right) \\ &\quad \times \sup_{\nu \in [0, t^*]} \left\| \dot{F}_\sigma^\parallel(\nu, Z_{N, \nu}) \right\|_{L_{2p}}, \\ \Delta_{\mathcal{U}_3}(\mathbf{p}, t^*) &= \left(\frac{\Delta_{\mathcal{P}_3}(\mathbf{p}, t^*)}{\lambda} + \frac{\Delta_g^2}{\lambda} \mathbf{p}'(\mathbf{p})^2 \sup_{\nu \in [0, t^*]} \left\| \dot{F}_\sigma^\parallel(\nu, Z_{N, \nu}) \right\|_{L_{2p}} \right) \sup_{\nu \in [0, t^*]} \left\| \dot{F}_\sigma^\parallel(\nu, Z_{N, \nu}) \right\|_{L_{2p}},\end{aligned}$$

where we have further used the definition $\Delta_{\mathcal{P}_2}(\mathbf{p}, t^*) = \mathbf{p}'(\mathbf{p}) \Delta_{\mathcal{P}_1}(\mathbf{p}, t^*)$ from Proposition D.8. To bound these terms, in addition to the results in Proposition D.14, we also use (D.147a), and the following bounds that can be derived similarly to (D.147b) and (D.148):

$$\sup_{\nu \in [0, t^*]} \left\| \dot{\Lambda}_\mu^\parallel(\nu, Z_{N, \nu}) \right\|_{L_{2p}} = \sup_{\nu \in [0, t^*]} \left\| \Lambda_\mu^\parallel(\nu, X_{N, \nu}) - \Lambda_\mu^\parallel(\nu, X_{N, \nu}^r) \right\|_{L_{2p}} \leq L_\mu^\parallel \tilde{\mathcal{E}}(2\mathbf{p}), \quad (\text{D.150a})$$

$$\sup_{\nu \in [0, t^*]} \left\| \dot{F}_\sigma^\parallel(\nu, Z_{N, \nu}) \right\|_{L_{2p}} \leq (L_p^\parallel + L_\sigma^\parallel) \sqrt{\tilde{\mathcal{E}}(2\mathbf{p})}. \quad (\text{D.150b})$$

Using the above then leads to the following for $\mathbf{p} \in \{1, \dots, \mathbf{p}^*\}$:

$$\Delta_{\mathcal{U}_1}(\mathbf{p}, t^*) \leq \Delta_{\odot_2}(\mathbf{p}) \tilde{\mathcal{E}}(2\mathbf{p}) + \Delta_{\otimes_2}(\mathbf{p}) \left(\tilde{\mathcal{E}}(2\mathbf{p}) \right)^{\frac{3}{2}} + \left(\Delta_{\otimes_2}(\mathcal{E}(2\mathbf{p}) + \mathcal{E}_r(2\mathbf{p})) + \Delta_{\otimes_3} \tilde{\mathcal{E}}(2\mathbf{p}) \right) \tilde{\mathcal{E}}(2\mathbf{p}), \quad (\text{D.151a})$$

$$\begin{aligned}\Delta_{\mathcal{U}_2}(\mathbf{p}, t^*) &\leq \Delta_{\odot_1}(\mathbf{p}) \sqrt{\tilde{\mathcal{E}}(2\mathbf{p})} + \Delta_{\odot_3}(\mathbf{p}) \tilde{\mathcal{E}}(2\mathbf{p}) \\ &\quad + \left(\Delta_{\otimes_3}(\mathbf{p}) (\mathcal{E}(2\mathbf{p}) + \mathcal{E}_r(2\mathbf{p})) + \Delta_{\otimes_4}(\mathbf{p}) \tilde{\mathcal{E}}(2\mathbf{p}) \right) \sqrt{\tilde{\mathcal{E}}(2\mathbf{p})}, \quad (\text{D.151b})\end{aligned}$$

$$\Delta_{\mathcal{U}_3}(\mathbf{p}, t^*) \leq \Delta_{\odot_4}(\mathbf{p}) \tilde{\mathcal{E}}(2\mathbf{p}). \quad (\text{D.151c})$$

Bounds for $\tilde{\Delta}_{\mathcal{U}}$: Now, we have from Definitions 1 and 3 that $F_\sigma(\nu, \cdot) = p(\nu, \cdot) + \Lambda_\sigma(\nu, \cdot)$ and $F_\sigma^\parallel(\nu, \cdot) = p^\parallel(\nu, \cdot) + \Lambda_\sigma^\parallel(\nu, \cdot)$, respectively. Thus, we obtain the following bound due to Assmptions 1 and 4:

$$\begin{aligned}\sup_{\nu \in [0, t^* \wedge T_s]} \left\| F_\sigma^\parallel(\nu, X_{N, \nu}) \right\|_{L_{2p}} &\leq \sup_{\nu \in [0, t^*]} \left\| F_\sigma^\parallel(\nu, X_{N, \nu}) \right\|_{L_{2p}} \leq \sup_{\nu \in [0, t^*]} \left\| F_\sigma^\parallel(\nu, X_{N, \nu}) \right\|_{L_{4p}} \\ &\leq \Delta_p^\parallel + \Delta_\sigma^\parallel + \Delta_\sigma^\parallel \sqrt{\mathcal{E}(2\mathbf{p})}.\end{aligned} \quad (\text{D.152})$$

Then, using the bounds (D.142) and (D.143) in the proof of Proposition D.15, and the bounds (D.147a) and (D.152) above, we derive the following expressions for the terms $\tilde{\Delta}^-(p, t^*, \omega, T_s)$ and $\tilde{\Delta}_k^+(p, t^*, \omega, T_s)$, $k = \{1, \dots, 3\}$, defined in Lemma D.4:

$$\tilde{\Delta}^-(p, t^*, \omega, T_s) \leq \tilde{\Upsilon}^-(\mathcal{E}(2p); p, \omega, T_s) \tilde{\mathcal{E}}(2p), \quad (\text{D.153a})$$

$$\tilde{\Delta}_1^+(p, t^*, \omega, T_s) \leq \frac{1}{\lambda} \Delta_g \Upsilon'_1(\mathcal{E}(2p); p, \omega, T_s) \tilde{\mathcal{E}}(2p), \quad (\text{D.153b})$$

$$\tilde{\Delta}_2^+(p, t^*, \omega, T_s) \leq \sqrt{\omega} \frac{1}{\lambda} \Delta_g \Upsilon'_2(\mathcal{E}(2p); p, \omega, T_s) \tilde{\mathcal{E}}(2p), \quad (\text{D.153c})$$

$$\tilde{\Delta}_3^+(p, t^*, \omega, T_s) \leq \frac{2}{|2\lambda - \omega|} \Delta_g \Upsilon'_3(\mathcal{E}(2p); p, \omega, T_s) \tilde{\mathcal{E}}(2p), \quad (\text{D.153d})$$

for all $p \in \{1, \dots, p^*\}$:

Recall the definition of $\tilde{\Delta}_{\mathcal{U}}(p, t^*, \omega, T_s)$ from (D.120) in Lemma D.4

$$\tilde{\Delta}_{\mathcal{U}}(p, t^*, \omega, T_s) = \tilde{\Delta}^-(p, t^*, \omega, T_s) + \mathbb{1}_{\geq T_s}(t^*) \sum_{k=1}^3 \tilde{\Delta}_k^+(p, t^*, \omega, T_s),$$

which can be trivially bounded as

$$\tilde{\Delta}_{\mathcal{U}}(p, t^*, \omega, T_s) \leq \tilde{\Delta}^-(p, t^*, \omega, T_s) + \sum_{k=1}^3 \tilde{\Delta}_k^+(p, t^*, \omega, T_s).$$

Substituting the bounds from (D.153) then leads to

$$\begin{aligned} \tilde{\Delta}_{\mathcal{U}}(p, t^*, \omega, T_s) &\leq \tilde{\Upsilon}^-(\mathcal{E}(2p); p, \omega, T_s) \tilde{\mathcal{E}}(2p) \\ &\quad + \frac{1}{\lambda} \Delta_g (\Upsilon'_1(\mathcal{E}(2p); p, \omega, T_s) + \sqrt{\omega} \Upsilon'_2(\mathcal{E}(2p); p, \omega, T_s)) \tilde{\mathcal{E}}(2p) \\ &\quad + \frac{2}{|2\lambda - \omega|} \Delta_g \Upsilon'_3(\mathcal{E}(2p); p, \omega, T_s) \tilde{\mathcal{E}}(2p), \end{aligned} \quad (\text{D.154})$$

for all $p \in \{1, \dots, p^*\}$.

Bounds for $\Delta_{\mathcal{U}}$: We next use (D.143), (D.152), and Proposition D.15, to derive the following expressions for the terms $\Delta^-(p, t^*, \omega, T_s)$ and $\Delta_k^+(p, t^*, \omega, T_s)$, $k = \{1, \dots, 3\}$, defined in Lemma D.2:

$$\Delta^-(p, t^*, \omega, T_s) \leq \mathring{\Delta}(p, t^*, \omega) \Upsilon^-(\mathcal{E}(2p); p, \omega, T_s), \quad (\text{D.155a})$$

$$\Delta_1^+(p, t^*, \omega, T_s) \leq \mathring{\Delta}(p, t^*, \omega) \frac{1}{\sqrt{\omega}} \Upsilon'_1(\mathcal{E}(2p); p, \omega, T_s), \quad (\text{D.155b})$$

$$\Delta_2^+(p, t^*, \omega, T_s) \leq \mathring{\Delta}(p, t^*, \omega) \Upsilon'_2(\mathcal{E}(2p); p, \omega, T_s), \quad (\text{D.155c})$$

$$\Delta_3^+(p, t^*, \omega, T_s) \leq \mathring{\Delta}(p, t^*, \omega) \frac{1}{\sqrt{\omega}} \Upsilon'_3(\mathcal{E}(2p); p, \omega, T_s). \quad (\text{D.155d})$$

As before, the definition of $\Delta_{\mathcal{U}}(p, t^*, \omega, T_s)$ in Lemma D.2 implies the following straightforward bound:

$$\Delta_{\mathcal{U}}(p, t^*, \omega, T_s) \leq \Delta^-(p, t^*, \omega, T_s) + \sum_{k=1}^3 \Delta_k^+(p, t^*, \omega, T_s).$$

Substituting (D.155) then leads to:

$$\begin{aligned} \Delta_{\mathcal{U}}(p, t^*, \omega, T_s) &\leq \mathring{\Delta}(p, t^*, \omega) \left(\Upsilon^-(\mathcal{E}(2p); p, \omega, T_s) + \frac{1}{\sqrt{\omega}} \Upsilon'_1(\mathcal{E}(2p); p, \omega, T_s) \right. \\ &\quad \left. + \Upsilon'_2(\mathcal{E}(2p); p, \omega, T_s) + \frac{1}{\sqrt{\omega}} \Upsilon'_3(\mathcal{E}(2p); p, \omega, T_s) \right). \end{aligned} \quad (\text{D.156})$$

We next recall the definition of the term $\mathring{\Delta}(p, t^*, \omega)$ from the statement of Lemma D.2 given by

$$\mathring{\Delta}(p, t^*, \omega) = \Delta_g^2 \left(\frac{1}{\sqrt{\omega}} \sup_{\nu \in [0, t^*]} \left\| \mathring{\Delta}_{\mu}^{\parallel}(\nu, Z_{N, \nu}) \right\|_{L_{2p}} + p'(p) \sup_{\nu \in [0, t^*]} \left\| \mathring{F}_{\sigma}^{\parallel}(\nu, Z_{N, \nu}) \right\|_{L_{2p}} \right),$$

which can be bounded by using (D.150) to obtain the following:

$$\mathring{\Delta}(\mathbf{p}, t^*, \omega) \leq \mathring{\Upsilon}(\tilde{\mathcal{E}}(2\mathbf{p}); \mathbf{p}, \omega) \doteq \Delta_g^2 \left(\mathbf{p}'(\mathbf{p}) (L_p^\parallel + L_\sigma^\parallel) \sqrt{\tilde{\mathcal{E}}(2\mathbf{p})} + \frac{1}{\sqrt{\omega}} L_\mu^\parallel \tilde{\mathcal{E}}(2\mathbf{p}) \right).$$

Substituting the above into (D.156) produces the following for $\mathbf{p} \in \{1, \dots, \mathbf{p}^*\}$:

$$\begin{aligned} \Delta_{\mathcal{U}}(\mathbf{p}, t^*, \omega, \mathbf{T}_s) \leq \mathring{\Upsilon}(\tilde{\mathcal{E}}(2\mathbf{p}); \mathbf{p}, \omega) & \left(\Upsilon^-(\mathcal{E}(2\mathbf{p}); \mathbf{p}, \omega, \mathbf{T}_s) + \frac{1}{\sqrt{\omega}} \Upsilon'_1(\mathcal{E}(2\mathbf{p}); \mathbf{p}, \omega, \mathbf{T}_s) \right. \\ & \left. + \Upsilon'_2(\mathcal{E}(2\mathbf{p}); \mathbf{p}, \omega, \mathbf{T}_s) + \frac{1}{\sqrt{\omega}} \Upsilon'_3(\mathcal{E}(2\mathbf{p}); \mathbf{p}, \omega, \mathbf{T}_s) \right). \quad (\text{D.157}) \end{aligned}$$

The proof is then concluded by substituting (D.149), (D.151), (D.154), and (D.157) into (D.146).

□